Optimising Outcomes For Potentially Resectable Pancreatic Cancer Through Personalised Predictive Medicine:
The application of complexity theory to probabilistic statistical modeling

PhD Thesis

Alison Bradley
Health Research Group
Management Science Department
University of Strathclyde Business School Glasgow
10th March 2020

Abstract

Survival outcomes for pancreatic cancer remain poor. Surgical resection with adjuvant therapy is the only potentially curative treatment, but for many people surgery is of limited benefit. Neoadjuvant therapy has emerged as an alternative treatment pathway however the evidence base surrounding the treatment of potentially resectable pancreatic cancer is highly heterogeneous and fraught with uncertainty and controversy.

This research seeks to engage with conjunctive theorising by avoiding simplification and abstraction to draw on different kinds of data from multiple sources to move research towards a theory that can build a rich picture of pancreatic cancer management pathways as a complex system. The overall aim is to move research towards personalised realistic medicine by using personalised predictive modeling to facilitate better decision making to achieve the optimisation of outcomes.

This research is theory driven and empirically focused from a complexity perspective. Combining operational and healthcare research methodology, and drawing on influences from complementary paradigms of critical realism and systems theory, then enhancing their impact by using Cilliers' complexity theory 'lean ontology', an open-world ontology is held and both epistemic reality and judgmental relativity are accepted. The use of imperfect data within statistical simulation models is explored to attempt to expand our capabilities for handling the emergent and uncertainty and to

find other ways of relating to complexity within the field of pancreatic cancer research.

Markov and discrete-event simulation modelling uncovered new insights and added a further dimension to the current debate by demonstrating that superior treatment pathway selection depended on individual patient and tumour factors. A Bayesian Belief Network was developed that modelled the dynamic nature of this complex system to make personalised prognostic predictions across competing treatments pathways throughout the patient journey to facilitate better shared clinical decision making with an accuracy exceeding existing predictive models.

Contents

List of Figures

pathway would be the superior pathway. The blue area depicts the range over which neoadjuvant (NAT) pathway would be the superior pathway.

Figure 37 b): predicted QALMs given altering probabilities of receiving R0 resection in SF pathway (y-axis) and resection in NAT pathway (x-axis)

Figure 38: Outline of Discrete Event Simulation Model

Figure 39: Monte Carlo Strategy Selection Frequency. The y-axis depicts the probability that each treatment option along the x-axis is selected as the optimal treatment pathway after 10000patients are simulated through the model. NAT means neoadjuvant pathway and SF means upfront surgery pathway.

Figure 40. Threshold Analysis. This figure shows the probability threshold for resection in the neoadjuvant pathway displayed on the x-axis that must be reached for this pathway to be superior. NAT means neoadjuvant therapy. SF means upfront surgery pathway. Expected value on y-axis is in life months.

Figure 41: Threshold analysis. Probability of resection in neoadjuvant pathway (x-axis) against expected outcome in months (y-axis). NAT is neoadjuvant pathway. SF is surgery first pathway

Figure 42. Threshold Analysis for Unresectable Disease in the Neoadjuvant Pathway. This figure shows the probability threshold for resection in the upfront surgery pathway, displayed on the x-axis, that must be reached for this pathway to be

superior. NAT means neoadjuvant therapy. SF means upfront surgery pathway. Expected value on y-axis is in life months.

Figure 43: Threshold analysis for borderline resectable cases treated in neoadjuvant and surgery-first pathways. Probability of resection in neoadjuvant pathway (x-axis) against expected outcome in months (y-axis). Red line represents surgery first pathway, blue line represents neoadjuvant pathway, and yellow line represents unresectable disease.

Figure 44: Overview of disease model

Figure 45a: Markov Two-way sensitivity analysis. Y-axis shows probability of receiving adjuvant therapy in upfront surgery (SF) pathway and x-axis shows probability of receiving resection in neoadjuvant (NAT) pathway. The red area depicts the range whereby upfront surgery pathway would be the superior pathway. The blue area depicts the range over which neoadjuvant pathway would be the superior pathway.

Figure 45b: Markov Two-way sensitivity analysis. Y-axis shows probability of receiving R0 resection in upfront surgery (SF) pathway and x-axis shows probability of receiving resection in neoadjuvant (NAT) pathway. The red area depicts the range whereby upfront surgery pathway would be the superior pathway. The blue area depicts the range over which neoadjuvant (NAT) pathway would be the superior pathway.

Figure 46: Results of Paired T-test. Comparison of expected outcomes from Markov and discrete-event simulation models with observed survival times from study population.

Figure 47: Overview of models

Figure Kviii: League table based on results of fixed effects and random effects (vague prior) models. Where odds ratio (O.R.) is greater than 1 treatment at top left is superior

Figure Kix: Rankogram summarising surface under the cumulative ranking (SUCRA).

Figure Kx: Results of fixed effects and random effects (vague prior) models

Figure Kxi: League table based on results of fixed effects and random effects (vague prior) models. Where odds ratio (O.R.) is greater than 1 treatment at top left is superior

Figure Kxii: Rankogram summarising surface under the cumulative ranking (SUCRA).

Figure Kxiii: Results of fixed effects and random effects (vague prior) models

Figure Kxiv: League table based on results of fixed effects and random effects (vague prior) models. Where odds ratio (O.R.) is greater than 1 treatment at top left is superior

Figure Kxv: Rankogram summarising surface under the cumulative ranking (SUCRA).

Figure Kxvi: League table based on results of fixed effects and random effects (vague prior) models. Where odds ratio (O.R.) is greater than 1 treatment at top left is superior

Figure Kxvii: Rankogram summarising surface under the cumulative ranking (SUCRA).

Figure Kxviii: Bayesian Network Meta-analysis of Neoadjuvant therapy versus Upfront surgery plus adjuvant therapy versus surgery only

Figure Kxix: Results of fixed effects and random effects (vague prior) models

Figure Kxx: League table based on results of fixed effects and random effects (vague prior) models. Where odds ratio (O.R.) is greater than 1 treatment at top left is superior

Figure Kxxi: Rankogram summarising surface under the cumulative ranking (SUCRA).

Figure Kxxii: League table based on results of fixed effects and random effects (vague prior) models. Where odds ratio (O.R.) is greater than 1 treatment at top left is superior

Figure Kxxiii: Rankogram summarising surface under the cumulative ranking (SUCRA).

Figure Kxxiv: Results of fixed effects and random effects (vague prior) models

Figure Kxxv: League table based on results of fixed effects and random effects (vague prior) models. Where odds ratio (O.R.) is greater than 1 treatment at top left is superior

Figure Kxxvi: Rankogram summarising surface under the cumulative ranking (SUCRA).

Figure Kxxvii: Bayesian Network Meta-analysis of Neoadjuvant therapy versus Upfront surgery plus adjuvant therapy versus surgery only

Figure Kxxviii: Results of fixed effects and random effects (vague prior) models

Figure Kxxix: League table based on results of fixed effects and random effects (vague prior) models. Where odds ratio (O.R.) is greater than 1 treatment at top left is superior

Figure Kxxx: Rankogram summarising surface under the cumulative ranking (SUCRA).

Figure Kxxxi: Bayesian Network Meta-analysis of Neoadjuvant therapy versus Upfront surgery plus adjuvant therapy versus surgery only

Figure Kxxxii: League table based on results of fixed effects and random effects (vague prior) models. Where odds ratio (O.R.) is greater than 1 treatment at top left is superior

Figure Kxxxiii: Rankogram summarising surface under the cumulative ranking (SUCRA).

Figure Kxxxiv: Results of fixed effects and random effects (vague prior) models

Figure Kxxxv: League table based on results of fixed effects and random effects (vague prior) models. Where odds ratio (O.R.) is greater than 1 treatment at top left is superior

Figure Kxxxvi: Rankogram summarising surface under the cumulative ranking (SUCRA).

Figure Kxxxvii: Results of fixed effects and random effects (vague prior) models

Figure Kxxxviii: League table based on results of fixed effects and random effects (vague prior) models. Where odds ratio (O.R.) is greater than 1 treatment at top left is superior

Figure Kxxxxix: Rankogram summarising surface under the cumulative ranking (SUCRA).

Figure Kxl: Bayesian Network Meta-analysis of Neoadjuvant therapy versus Upfront surgery plus adjuvant therapy versus surgery only

Figure Kxlviii: Results of fixed effects and random effects (vague prior) models

Figure Kxli: League table based on results of fixed effects and random effects (vague prior) models. Where odds ratio (O.R.) is greater than 1 treatment at top left is superior

Figure Kxlii: Rankogram summarising surface under the cumulative ranking (SUCRA).

Figure Kxliii: Bayesian Network Meta-analysis of Neoadjuvant therapy versus Upfront surgery plus adjuvant therapy versus surgery only

Figure Kxliv: League table based on results of fixed effects and random effects (vague prior) models. Where odds ratio (O.R.) is greater than 1 treatment at top left is superior

Figure Kxlv: Rankogram summarising surface under the cumulative ranking (SUCRA).

Figure Kxlvi: Results of fixed effects and random effects (vague prior) models

Figure Kxlvii: League table based on results of fixed effects and random effects (vague prior) models. Where odds ratio (O.R.) is greater than 1 treatment at top left is superior

Figure Kxlviii: Rankogram summarising surface under the cumulative ranking (SUCRA).

Figure Kxlix: Results of fixed effects and random effects (vague prior) models

Figure Kl: League table based on results of fixed effects and random effects (vague prior) models. Where odds ratio (O.R.) is greater than 1 treatment at top left is superior

Figure Kli: Rankogram summarising surface under the cumulative ranking (SUCRA).

Figure Klii: Results of fixed effects and random effects (vague prior) models

Figure Kliii: League table based on results of fixed effects and random effects (vague prior) models. Where odds ratio (O.R.) is greater than 1 treatment at top left is superior

Figure Kliv: Rankogram summarising surface under the cumulative ranking (SUCRA).

Figure Klv: Results of fixed effects and random effects (vague prior) models

Figure Klvi: League table based on results of fixed effects and random effects (vague prior) models. Where odds ratio (O.R.) is greater than 1 treatment at top left is superior

Figure Klvii: Rankogram summarising surface under the cumulative ranking (SUCRA).

Figure Klviii: Bayesian Network Meta-analysis of Neoadjuvant therapy versus Upfront surgery plus adjuvant therapy versus surgery only

Figure Klix: League table based on results of fixed effects and random effects (vague prior) models. Where odds ratio (O.R.) is greater than 1 treatment at top left is superior

Figure Klx: Rankogram summarising surface under the cumulative ranking (SUCRA).

Figure Klxi: Bayesian Network Meta-analysis of Neoadjuvant therapy versus Upfront surgery plus adjuvant therapy versus surgery only

Figure Klxii: Results of fixed effects and random effects (vague prior) models

Figure Klxiii: League table based on results of fixed effects and random effects (vague prior) models. Where odds ratio (O.R.) is greater than 1 treatment at top left is superior

Figure Klxiv: Rankogram summarising surface under the cumulative ranking (SUCRA).

Figure Klxv: Assessment of Risk of Bias for each included trials of neoadjuvant versus upfront surgery plus adjuvant therapy

Figure Klxvi: Assessment of Risk of Bias of Randomised Controlled Trials comparing Upfront Surgery plus Adjuvant Therapy versus Surgery Only.

Figure Klxviia: 1-year Survival Model Including Phase II/III Trials and Randomised Controlled Trials

Figure Klxviib: 1-year Survival Model Including Phase II/III Trials, Randomised Controlled Trials and Cohort Studies

Figure Klxviiia: 2-year Survival Model Including Phase II/III Trials and Randomised Controlled Trials

Figure Klxviiib: 2-year Survival Model Including Phase II/III Trials, Randomised Controlled Trials and Cohort Studies

Figure Klxixa: 3-year Survival Model Including Phase II/III Trials and Randomised Controlled Trials

Figure Klxixb: 3-year Survival Model Including Phase II/III Trials, Randomised Controlled Trials and Cohort Studies

Figure Klxxa: 5-year Survival Model Including Phase II/III Trials and Randomised Controlled Trials

Figure Klxxb: 5-year Survival Model Including Phase II/III Trials, Randomised Controlled Trials and Cohort Studies

Figure Li: Results of fixed effects and random effects (vague prior) models

Figure Lii: League table based on results of fixed effects and random effects (vague prior) models. Where odds ratio (O.R.) is greater than 1 treatment at top left is superior

Figure Liii: Rankogram summarising surface under the cumulative ranking (SUCRA).

Figure Liv: League table based on results of fixed effects and random effects (vague prior) models. Where odds ratio (O.R.) is greater than 1 treatment at top left is superior

Figure Lv: Rankogram summarising surface under the cumulative ranking (SUCRA).

Figure Lvi: Results of fixed effects and random effects (vague prior) models

Figure Lvii: League table based on results of fixed effects and random effects (vague prior) models. Where odds ratio (O.R.) is greater than 1 treatment at top left is superior

Figure Lviii: Rankogram summarising surface under the cumulative ranking (SUCRA).

Figure Lix: Bayesian Network Meta-analysis of Neoadjuvant therapy versus Upfront surgery plus adjuvant therapy versus surgery only

Figure Lx: League table based on results of fixed effects and random effects (vague prior) models. Where odds ratio (O.R.) is greater than 1 treatment at top left is superior

Figure Lxi: Rankogram summarising surface under the cumulative ranking (SUCRA).

Figure Lxii: Results of fixed effects and random effects (vague prior) models

List of Tables

Table Oi: One-Way ANOVA Analysis: Results of overall survival time equal or greater than 36months between groups contained within each pre-operative variable.

Table Oii: Bayesian Log Linear Regression: Pre-operative Variables: overall survival grouped as 36months or above or not

Table Oiii: Pearson Correlation Coefficient: results of each variable assessed against overall survival defined as 36 months or more (column 2) and overall survival time in months (column 3)

Table 65Oiv: Bayesian Linear Regression

Table Ov: Bayesian Linear Regression

Table Ovi: One-Way ANOVA Analysis: Results of overall survival time equal or less than 12months between groups contained within each pre-operative variable.

Table Ovii: Bayesian Log Linear Regression: Pre-operative Variable: Overall survival grouped as 12 months or less or greater than 12 months

Table Oviii: Pearson Correlation Coefficient: results of each variable assessed against overall survival defined as 12 months or less

Table Oix: One-way ANOVA Neoadjuvant Sub-Group Analysis: Assessment of response to neoadjuvant treatment for all patients treated in neoadjuvant pathway

Table Ox: Pearson Correlation co-efficient Neoadjuvant Sub-Group Analysis

Table Oxi: Bayesian Log Linear Regression Neoadjuvant Sub-Group Analysis

Table Oxii: One-way ANOVA RPDAC Sub-Group Analysis

Table Oxiii: Bayesian Log Linear Regression RPDAC Sub-Group Analysis

Table Oxiv: Pearson Correlation co-efficient RPDAC Sub-Group Analysis

Table Pi. Summary of existing model development studies predicting post-resection prognosis for PDAC

Preface/ Acknowledgements

I would like to thank and acknowledge the much appreciated help, support and encouragement I received throughout from my supervisors Dr Robert Van der Meer and Professor Alec Morton as well as Professor Colin McKay and the staff of the West of Scotland Pancreatic Unit at Glasgow Royal Infirmary.

I would also like to thank my mother for whom 'Deeds not Words' is a way of life that you have shown me through your unwavering support throughout life. I owe you everything.

Finally I would like to dedicate this thesis to the memory of three sorely missed loved ones: my father, granny and dear friend and mentor in surgery Sarah Prince. Thank you for always encouraging me whilst teaching me that the kindness in a person's heart matters more than any letters behind a name.

List of Publications From This Thesis

Bradley, A., Van Der Meer, R. and McKay, C.J. (2019) 'A systematic review of methodological quality of model development studies predicting prognostic outcome for resectable pancreatic cancer'. *BMJ Open,* 9**:**e027192. doi: 10.1136/bmjopen-2018-027192

Bradley, A., Van der Meer, R. and McKay, C. (2019) 'Personalized pancreatic cancer management: a systematic review of how machine learning is supporting decision-making'*. Pancreas,*48 (5). pp. 598-604.

Bradley, A. *et al.* (2018). 'Markov decision analysis of neoadjuvant treatment pathway versus surgery first pathway for resectable pancreatic cancer'. *Journal of Clinical Oncology,* 36 (4). pp. 456-456

Bradley, A. and Van Der Meer, R. (2019). 'Neoadjuvant therapy versus upfront surgery for potentially resectable pancreatic cancer: a Markov decision analysis'. *PLoS One*, 14(2):e0212805. doi:10.1371/journal.pone.0212805

Bradley, A. and Van Der Meer, R. (2019). 'Upfront surgery versus neoadjuvant therapy for resectable pancreatic cancer: systematic review and Bayesian network meta-analysis'. *Nature Scientific Reports*, 9(1):4354. doi:10.1038/s41598-019-40951-6

Bradley, A., Van der Meer, R. and McKay, C.J. (2019). 'A prognostic Bayesian network that makes personalized predictions of poor

prognostic outcome post resection of pancreatic ductal adenocarcinoma'. *PLoS One,* 14(9): e0222270. doi.org/10.1371/journal.pone.0222270

Bradley, A. *et al.* (2019). 'Personalized prognostic Bayesian network for pancreatic cancer: delivering personalized pancreatic cancer management throughout the patient journey'*. Pancreatology*, 19 (S1). pp. S31-S32.

Bradley, A. *et al.* (2019). 'Making personalized predictions of poor outcome post resection of pancreatic ductal adenocarcinoma (PDAC): a prognostic Bayesian network with pre and post-operative application'. *Pancreatology*, 19 (S1), pp. S122. P6-13

Bradley, A., Van Der Meer, R., McKay, C.J. (2020) 'Bayesian network meta-analysis of upfront surgery versus neoadjuvant therapy for potentially resectable pancreatic ductal adenocarcinoma'. *British Journal of Surgery*: (Abstract): accepted

Bradley, A., Van Der Meer, R., McKay, C.J. (2020) 'Optimising outcomes for resectable pancreatic cancer by learning lessons from military strategy and the stock market: creation of a prognostic Bayesian belief network that makes personalised pre and post-operative predictions of outcomes across competing treatment strategies.' *British Journal of Surgery*: (Abstract): accepted

Bradley, A., Van Der Meer, R., McKay, C.J. (2020) 'Computer simulated comparison of neoadjuvant versus upfront surgery for resectable

pancreatic cancer: the application of machine-learning algorithms to support personalised decision-making.' *British Journal of Surgery*: (Abstract): accepted

Chapter 1

Introduction

Pancreatic cancer is arguably the most challenging of all gastrointestinal tumours with 10-year survival remaining at 1% and overall 5-year survival remaining at 4% despite advancement in surgical technique and adjuvant therapy (Pancreatic Cancer United Kingdom (PCUK), 2016). Pancreatic cancer is the twelth most common cancer worldwide (World Cancer Research Fund, 2018) and the fourth and fifth most common cause of cancer deaths in the United States of America (USA) and Europe respectively (Siegel *et al.,* 2015; Ferlay *et al.,* 2012). Within the United Kingdom (UK) 9,400 new cases of pancreatic cancer were diagnosed in 2013, accounting for 3% of all cancer diagnosis, and making pancreatic cancer the tenth most common cancer and fifth most common cause of cancer death in the UK (PCUK, 2016). Overall this represented an increased incidence rate of 10% over the past decade (PCUK, 2016).

In the UK it is estimated that only 9.8% of cases are amenable to surgical resection and 5-year survival for resected cases is reported at between 7% and 25% despite surgical resection being the only potentially curative treatment (Cancer Research UK (CRUK), 2019). For cases that are resectable at presentation current guidelines recommend surgery followed by adjuvant therapy (Khorana *et al*., 2019). However, up to 50% of patients do not receive adjuvant therapy, which nullifies any potential benefit from high-risk costly

surgery (Winter *et al*., 2012). Reasons for this include a combination of factors such as early disease recurrence, decline in health related to pre-existing illnesses, and/or post-operative complications rendering the patient too unwell to withstand further treatment (Evans *et al*., 2018). These factors have contributed to an increasing interest in neoadjuvant therapy as an alternative treatment pathway (Evans *et al*., 2018; Winter *et al*., 2012; Bilimoria *et al*., 2007a). Postulated benefits of neoadjuvant treatment pathway include: increased obtainment of multimodal treatment, converting borderline resectable disease to resectable, and filtering patients with more aggressive disease away from ultimately futile high-risk, high-cost surgery (Evans *et al*., 2018; Asare *et al*., 2016; Lee *et al*., 2016; Abbott *et al*., 2013).

However, there is currently a lack of conclusive level I evidence proving superiority of either treatment pathway for resectable disease (Versteijne et al., 2018). Neoadjuvant therapy for resectable pancreatic cancer is an area of prime controversy. Ambiguity reigns over the existing body of research comparing neoadjuvant and traditional upfront surgery approach for resectable cases. Critics highlight the limitations of drawing overly optimistic conclusions about neoadjuvant therapy from small, underpowered studies with a high degree of heterogeneity and caution against potentially losing the window of resectability for cases that are resectable at presentation (Asare *et al.*, 2016; Lee *et al*., 2016). Currently there is a lack of randomised controlled trials (RCT) offering a direct comparison between treatment approaches (Versteijne *et al*., 2018) with many comparison studies including borderline resectable and

locally advanced cases in the neoadjuvant cohort hence failing to offer a true like-for-like comparison. Preliminary results from the Prep-02/JSAP-05 trial (Unno *et al*., 2019), the first RCT to compare neoadjuvant and upfront surgery approach for resectable cases of pancreatic cancer, reported no statistically significant difference in resection, R0 resection and postoperative complication rates between the two cohorts with an overall survival time of 36.72months in the neoadjuvant arm and 26.65months in the surgery-first arm. However, the PRODIGE 24/CCTG PA.6 RCT that compared adjuvant modified (m)FOLFIRINOX to adjuvant gemcitabine in patients who have had their tumour resected within an upfront surgery pathway, reported overall survival times of 54months and 35months in each arm respectively (Conroy *et al*., 2018). The mFOLFIRINOX arm therefore exceeded the survival time reported in the neoadjuvant arm of the Prep-02/JSAP-05 trial (Unno *et al*., 2019).

In summary current guidelines recommend upfront surgery followed by adjuvant therapy for resectable pancreatic cancer and there is a growing acceptance for the use of neoadjuvant therapy for cases of pancreatic cancer that are borderline resectable or locally advanced at presentation with the aim of conversion to resectability (Khorana *et al*., 2019; Raufi *et al.,* 2019; Evans *et al*., 2018). However, to achieve optimisation of individual patient outcomes for pancreatic cancer we must go beyond adherence to guidelines to engage with the complexity of the system of delivering pancreatic cancer management for the reasons that will now be outlined.

Firstly, upfront surgery pathway for resectable pancreatic cancer has not produced a significant change in survival outcomes over the past three decades and the majority of patients who undergo surgical resection of their tumour with or without adjuvant therapy will develop metastatic disease (Evans *et al.*, 2018). The implications of acknowledging this fact have contributed to a growing recognition of pancreatic cancer as a systemic disease, hence even where the tumour is localised and deemed operable micrometastsitc disease is likely to be present although not clinically apparent (Wolff *et al.*, 2017). Whilst this has led some to champion a move towards early systematic therapy through neoadjuvant therapy and the application of surgery only to those patients most likely to benefit from such major operations (Evans *et al.*, 2018), the evidence base for such a move, particularly for cases of pancreatic cancer that are resectable at presentation, is controversial and heavily contested as previously discussed. Furthermore the optimal combination of treatment agents used within either the neoadjuvant or upfront surgery pathways is ever evolving and contested. This demonstrates that even by framing the question as how to best optimise patient outcomes from pancreatic cancer in the most simplistic terms of 'neoadjuvant *versus* upfront surgery', research must engage with a high degree of uncertainty and complexity.

This brings me to my second argument in favour of acknowledging and engaging with complexity in order to move research forward, specifically in changing the narrative surrounding the treatment of potentially resectable pancreatic cancer to reflect the evolution of our understanding of the disease and improve patient counseling and

shared decision making. Theoretically surgery is the only potentially curative treatment for pancreatic cancer. However, potentially resectable disease can include resectable, borderline resectable and some locally advanced stages of the disease. It has been established that the delivery of multimodal treatment (surgery in combination with either chemotherapy or chemoradiotherapy) within either a neoadjuvant or upfront surgery pathway results in improved survival time (Neoptolemos *et al.*, 2018; Khorana *et al*., 2019; Raufi *et al.,* 2019; Evans *et al*., 2018). Inherent in the decision to deliver multimodal treatment therefore is first the identification of patients with potentially resectable disease, but there is currently incomplete consensus about the working definition of 'operable pancreatic cancer' in terms of both tumour anatomy and patient factors including age and comorbidities (Evans *et al.*, 2018). The reality of the narrative therefore becomes much more complicated as each of these disease stages are likely to have different anticipated outcomes even if they undergo surgical resection. The narrative must therefore evolve beyond surgery being the only potentially curative treatment to an understanding that for some patients with operable disease the potential benefits of surgery in terms of disease free survival may be limited (Evans *et al.*, 2018). This requires a move towards better patient selection across competing treatment pathways to deliver individualised treatment sequencing and stage specific therapy to optimise individual patient outcomes (Evans *et al.*, 2018; McGuigan *et al.*, 2018).

There is a growing narrative and focus within pancreatic cancer research to attempt to achieve the delivery of personalised, targeted

treatment sequencing through biomarker driven treatment sequencing and/or the development of gene targeted therapies (Collisson *et al.*, 2019; Amanam & Chung, 2018; Sato-Dahlman *et al*., 2018). The widely held assumption is that breakthroughs in such areas will result in a move away from uncertainty towards precision medicine. The third point being made is that this current direction in pancreatic cancer research, rather than resulting in the diminution of complexity, could result in its augmentation. As our knowledge of disease at biomolecular and genomic level evolves the clinical decision making process will pullulate with varied and complex datasets from multiple sources. The amalgamation of such large complex databases and the meaningful application of this information to the individual patient to optimise treatment outcomes will be beyond the capabilities of the human mind to handle unaided (Obermeyer & Lee, 2017). To illustrate, as previously discussed the consensus definition of resectability, or lack there of, depends not only on tumour anatomy but also patient factors such as age and comorbidities (Evans *et al.*, 2018). It follows that a tumour with the same anatomy, genomics and biomarker profile would not necessarily follow the same clinical course and this must be acknowledged in order to optimise individual patient outcomes. To illustrate, all tumour factors being equal an active health conscious patient in their twenties is likely to have a different risk profile and anticipated clinical course following major surgery compared to a morbidly obese chain-smoker with numerous pre-existing comorbidities. Biomarkers and genetic profiles therefore provide only some of the picture and this information must be combined and integrated with other clinical data if progress is to be made.

To summarise, the challenge of optimising outcomes for pancreatic cancer goes beyond simply choosing between neoadjuvant therapy *versus* upfront surgery approach. The narrative surrounding the management of pancreatic cancer must also move away from the theoretically possible (surgery is the potentially curative treatment) to the reality for many patients (surgery may be of limited benefit). Whilst this move will in part be driven by better objective definitions of resectability and biomarker and gene targeted treatment sequencing, such anticipated developments, whilst important, will not on their own optimise outcomes for pancreatic cancer and could actually serve to increase uncertainty and complexity in clinical decision making. Therefore if the optimisation of outcomes for pancreatic cancer is to come to fruition through a more personalised approach to the delivery of pancreatic cancer treatment we must develop ways to engage with the complexity, handle uncertainty and the emergent when examining the complex system of delivering pancreatic cancer care including areas of debate, ambiguity and disagreement (Law & Mol, 2002; Fraser & Greenhalgh, 2001; Star, 2002; Greenhalgh & Papoutsi, 2018). As both a demonstration of the much broader complexity of pancreatic cancer management as a complex system, as well as providing a contextual framework for how to address the issue of optimising treatment outcomes, the political context in which this research was undertaken will now be outlined.

1.1 Political Context of This Research: The Move Towards Personalised Realistic Medicine

In March 2015 Scotland's Chief Medical Officer (CMO) launched their annual report: 'Realistic Medicine'. This set out six key challenges that must be met to deliver realistic medicine within the Scottish National Health Service (The Scottish Government, 2016; The Scottish Government, 2017):

1. Build a personalised approach to healthcare

2. Change our style to shared decision making

3. Reduce unnecessary variation in practice and outcomes

4. Reduce harm and waste

5. Manage risk better

6. Become improvers and innovators

This report posed some key questions in health care centering on individualised patient care to achieve realistic medicine by asking:

- how can we reduce the burden, both financial and to the patient experience, of over investigation and treatment?
- How can we reduce risk to patients and variation in practice to optimise treatment outcomes for all patients?
- How can we improve the doctor-patient relationship and combine both patient and professional expertise to focus on

treatment outcomes that matter to the individual patient? (The Scottish Government, 2016)

Finally, in recognition of the gravitas of the task ahead and the need for creative solutions the report asks:

- How can doctors become creative innovators to achieve improved individualised outcomes for their patients? (The Scottish Government, 2016)

More recently the CMO in their annual report rebranded the term 'Realistic Medicine' as 'Personalised Realistic Medicine' to emphasise the emerging importance placed by clinicians and patients on a personalised approach to care (The Scottish Government, 2019). This reemphasis on the personalisation of care appeared to be a direct response to the tendency of healthcare delivery towards a reductionist approach to care (The Scottish Government, 2017) with the drive for efficiency and effectiveness leading to the industrialisation of healthcare with the patient being reduced to a statistic or an object on a conveyer belt (Montori, 2017). Such concerns echo the previously expressed concerns regarding the limitations of reducing pancreatic cancer patients to mere biomarkers and genetic codes to decide treatment delivery.

'Personalised Realistic Medicine' still aims to deliver the right care to the right patient at the right time (The Scottish Government, 2019). However, it goes further in calling for the marrying-up of the delivery of better value healthcare with what Professor Victor Montori of the Mayo Clinic has termed 'careful and kind care' (The Scottish Government, 2019; Montori, 2017). Careful care encompasses the

principles of quality, safety and evidence based practice and considers the patient's biology in the context of their biography; that is their disease and comorbidities in the context of their life situation and priorities. Kind care then respects the patient's resources of time, energy and attention and seeks to minimise the impact of healthcare upon these (Montori, 2017).

To actually achieve this the focus must be on understanding the patient as an individual with their own preferences and values as well as focusing on service provision. Therefore questions must be asked regarding how services can be designed and practices adapted to engage patients in their care without overwhelming them (The Scottish Government, 2019). In practical terms this will mean reviewing how resources (physical, monetary and time) could be better distributed and more effectively targeted to support such a move. Secondly the paternalistic culture of communication with patients must change towards a collaborative partnership of shared decision making that fosters a trusting relationship through a dialogue of openness and honesty (The Scottish Government, 2019). This will provide a particular challenge as to how complicated and conflicting information based on population level data can be clearly conveyed and discussed with patients during the decision making process so that patients can ask questions and be provided with honest, realistic answers (The Scottish Government, 2019). This means taking a lead from the House of Lords Science and Technology Committee report in 2000 (UK House of Lords, 2000) in rejecting the 'deficit model' whereby the provision of information by experts is

expected to make the patient, or public, agree with the expert. This report therefore concluded that many issues faced by decision makers and treated as science issues may in fact involve many other non-science factors. The implication for delivering realistic medicine as recognised by the CMO is that:

> "In the same way, we must accept that, to deliver Realistic Medicine, we need to consider many factors besides medicine" (The Scottish Government, 2019, p.17).

This acknowledgement by both the House of Lords Science and Technology Committee and the CMO, who throughout their most recent report discusses patients, the public and healthcare systems as complex interacting systems, mirrors that of recent moves by the Medical Research Council to acknowledge the need to engage with, rather than simplify or deny, complexity (Moore *et al*., 2015).

1.2 Using Complexity Theory As A Lens Through Which To Focus The Research Question

This research seeks to engage with what Tsoukas (2017) called conjunctive theorising by avoiding simplification and abstraction (or disjunctive theorising) and instead draws on different kinds of data from multiple sources to move research towards a theory that can build a rich picture of pancreatic cancer management pathways as a complex system. Combining operational and healthcare research and drawing on influences from complementary paradigms of critical realism and systems theory then enhancing their impact by using Cilliers' complexity theory 'lean ontology', an open-world ontology is held (Cilliers, 1998; Kruger *et al*., 2019). This posits that the interplay

of causal powers or tendencies of domains of 'the real' leads to particular events, 'the actual' (Mingers, 2005). These domains may be physical, social or conceptual and these events may be observable or experienced by people and therefore become empirical, but that as a whole the world is open to multiple interacting influences and to ignore such layers of influence serves no analytical benefit (Cilliers, 1998; Cilliers, 2010; Mingers, 2005). Epistemologically in recognising that all knowledge, whilst provisional, is historically and culturally relative both epistemic reality (observer-independent access as a fallacy) and judgmental relativity (rational grounds for theory preference) are accepted (Mingers, 2005). By amalgamating operational and healthcare research disciplines in this way this research seeks to be theory driven and empirically focused from a complexity perspective. Through a 'systems mindset' methodological pluralism is embraced to expand the methodological repertoire (Cilliers, 1998; Kruger *et al*., 2019). Specifically how imperfect data can be better utilised within statistical simulation models will be explored so that, as Long *et al*. (2018) have suggested, the potential for simulation modelling in the study of complexity in healthcare can be explored to attempt to expand capabilities for handling the emergent and uncertainty. Methods of statistical modeling and their ability to cope with uncertainty and capture system complexity will also be explored.

1.3 Research Aims and Objectives

The aim of this research is to facilitate the advancement of personalised realistic medicine in the delivery of pancreatic cancer services through statistical modelling that will facilitate better shared

decision making with patients and the entire multi-disciplinary team to optimise individual patient outcomes as determined by the individual patient. The impact of this research in reducing unnecessary investigations and treatments will be assessed through cost-effectiveness analysis of treatment pathways.

Particular areas of interest will be:

- Analysis of neoadjuvant *versus* upfront surgery pathways for patients presenting with potentially resectable pancreatic cancer in terms of quality-adjusted health outcomes and cost-effectiveness
- Improved patient selection and risk stratification of patients for pancreatic cancer surgery in both the pre and post-operative phases of the patient journey
- Improved individualised prognostic predictions across competing treatment pathways

Statistical modelling, offering visualisation within a logical framework of a sequence of events resulting from alternative treatment decisions and associated health and cost outcomes, could go some way to achieving these aims (Kuntz *et al.* 2013). Ultimately however such models will seek to assist with decision making rather than make statements about truth (Kuntz *et al*. 2013).

This research seeks to go further and use statistical modelling to give individualised predictions of outcome so that the care delivered to

pancreatic cancer patients can be truly 'realistic' to them as individuals. Perhaps one particular goal highlighted by the CMO holds the key to achieving this: creativity and innovation (The Scottish Government, 2017). The specific objectives of this research are:

1. Perform detailed decision-analysis of competing treatment pathways: upfront surgery *versus* neoadjuvant therapy, to explore thresholds pertaining to individual patient and tumour factors that could determine superiority of competing treatment strategies at an individualised patient level. To achieve this various approaches to statistical modeling including hybrid modeling approaches will be employed
2. Perform cost-effectiveness analysis of competing treatment pathways for potentially resectable cases of pancreatic cancer to assess the wider economic impact of improved individualised treatment pathway selection. The impact of using a variety of modeling techniques in addressing this issue will also be assessed
3. Perform detailed analysis of the West of Scotland Pancreatic Unit 20year prospectively maintained database to identify pre-operative variables that predict survival outcomes. Within this a subgroup analysis of resectable only cases will be conducted to offer a true like-for-like comparison to assess the impact of pathway selection on survival outcomes and the associated cost-effectiveness impact
4. Explore whether better use can be made of existing data and how such data can be combined with available institutional patient level data within a variety of modeling methods to

develop new insights into the ongoing debate and areas of uncertainty concerning pancreatic cancer management

5. Create predictive prognostic models that can make personalised predictions of survival outcome and risk of treatment complications or failure across competing treatment options, based on individual patient and tumour factors, at the pre-operative stage of the patient journey
6. Expand these personalised predictive models to perform prognostic updating at the post-operative stage of the patient journey across alternative post-operative treatment options and potential clinical scenarios.

1.4 Thesis Overview

The thesis will be arranged as follows. Chapter 2 contains the literature review which, after exploring the current evidence base surrounding the treatment of potentially resectable pancreatic cancer and outlining the existing areas of debate and uncertainty, critically analyses how and to what degree of success statistical modeling has been applied to the assessment of the management of pancreatic cancer in terms of both cost-effectiveness analysis and predictive and prognostic modeling including the role of emerging machine learning techniques to support clinical decision-making. This chapter concludes by demonstrating that overall the current application of statistical modeling to support decision making with regards to pancreatic cancer management is limited, not solely by the prevailing flawed, uncertain, proximate and sparse (FUPS) data (Wolpert & Rutter, 2018), but by a failure to acknowledge and attempt to engage with the complexity of the issue. The proceeding

methods chapter therefore builds the case for why and how the system of delivering pancreatic cancer care should be viewed as a complex and adaptive system in order to gain new insights. The case is made that by using complexity theory as a lens through which to view pancreatic cancer management, the existing body of imperfect data can be better utilised within statistical simulation modeling to expand the capabilities of modeling techniques to handle emergent and uncertainty and gain new insights that will facilitate better shared decision making and go some way to making personalised realistic medicine a reality.

The results chapter opens with a series of meta-analyses of the existing body of published data. Conclusive superiority of either upfront surgery or neoadjuvant pathway could not be established therefore the data was further interrogated through Markov decision-analysis. The Markov model was also populated with patient level data from the West of Scotland Pancreatic Unit database. The sensitivity analysis as part of these analysis demonstrated corroborating thresholds that began to indicate how optimal treatment pathway selection could be determined by individual patient and tumour factors. The cost-effectiveness implications of such factors in determining treatment pathway selection and the impact on quality as well as quantity of survival time is then explored. This analysis adds a new dimension to the debate surrounding the treatment of potentially resectable pancreatic cancer by moving the focus away from neoadjuvant *versus* upfront surgery towards the planning and delivery of more personalised care. As Markov modeling performs cohort level analysis, patient level microsimulation was then utilized through discrete event

simulation (DES) modeling. This further corroborated individualised thresholds for optimal pathway selection and added further insight by beginning to explore 'what if' scenarios had, for example, those individual patients who did not progress to surgery within the neoadjuvant pathway been treated within an upfront surgery pathway. This form of patient-level microsimulation analysis also proved to have implications for cost-effectiveness analysis and led to the exploration of the implications of model boundaries on the analysis of such a complex system. This proved to be an important finding as, although Markov modeling is widely used as a modeling method for cost-effectiveness analysis, a comparison of Markov and DES modeling showed that the latter produced survival predictions closer to the actual survival times observed within the institutional database.

By employing statistical modeling techniques that engaged with uncertainty and the research problem as a complex system, emergence occurred which revealed aspects of the system at individual patient level that impacted upon outcomes and could better inform decision making, but that had previously been under appreciated. Attention therefore turned to explore how such emergent properties could be better used to facilitate better shared decision making. Specifically the application of Bayesian network modeling to assess risk and explore the potential of this modeling technique in making personalised predictions of outcome was explored. A Bayesian belief network that could make better use of the existing imperfect data to engage with complexity to address the limitations of pre-existing prediction models to provide personalised predictions of outcomes across competing treatment strategies pre-

operatively and perform prognostic updating at the post-operative stage of the patient journey was consequently created and externally validated.

This thesis then concludes with a discussion regarding the impact and future direction of this research. This centres on the potential of the statistical models developed here to encompass anticipated future breakthroughs in precision medicine by offering a vehicle to integrate such large and complex genetic databases with pre-existing clinical data to make individualised predictions of outcomes and clinically deliver personalised precision medicine. This potential however is also discussed within the context of the need for further research structured around the Non-adoption, Abandonment, and challenges to the Scale-up, Spread and Sustainability (NASSS) framework of health an care technologies (Greenhalgh *et al*., 2017).

Chapter 2

Literature Review

2.1 Current Evidence-Base Underpinning The Management of Potentially Resectable Pancreatic Cancer

Introduction

Pancreatic cancer is a devastating disease associated with aggressive tumour biology, poor survival outcomes and increasing incidence rates (McGuigan *et al*., 2018). Globally it is ranked as the fourteenth most common cancer with an estimated 458,918 new diagnosis made in 2018 (International Agency for Research on Cancer, World Health Organisation (WHO), 2018). It is the seventh most common global cause of cancer death with an estimated 432,242 global pancreatic cancer deaths in 2018 (International Agency for Research on Cancer, WHO, 2018). Age-standardised incidence rates are highest in Europe and Northern America (Ilic & Ilic, 2016) with an increasing trend in incidence rates most marked within the developed Western world (Wong *et al*., 2017; Saad *et al*., 2018). It is estimated that by 2030 pancreatic cancer will have risen from being the fourth to the second most common cause of cancer related death in the USA (Siegel *et al.*, 2017; Rahib *et al*., 2014).

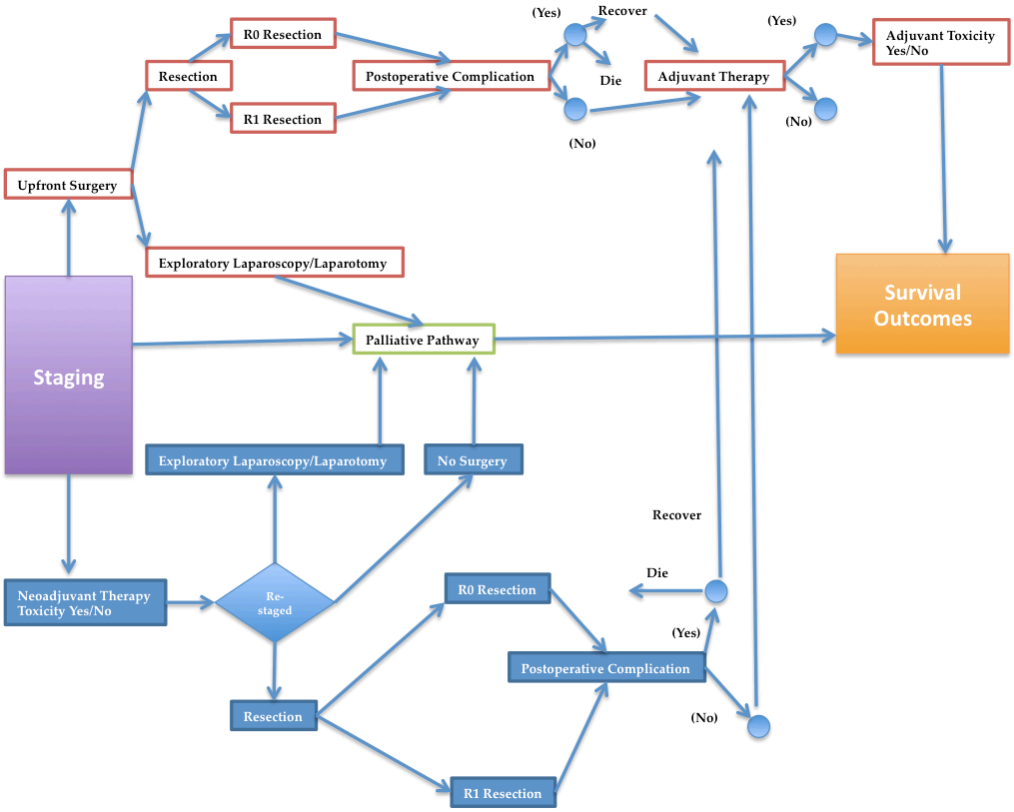
An overview of the treatment pathway options for pancreatic cancer is provided in Figure 1. Surgery is the only potentially curative form of treatment but an estimated 80% to 85% of patients present with inoperable metastatic disease (Vincent *et al*., 2011) with only 10% to 20% of cases deemed to be resectable at presentation (Jemal *et al*., 2010; Levy *et al*., 2016). The most recent estimates from the UK report the percentage of resectable cases to be at only 9.8% (CRUK, 2019). Surgery remains a largely morbid endeavor with potential benefits nullified by local-regional recurrence in 75% of cases, with synchronous distant failure in 50% to 80% of cases, within months of resection (Papavasiliou *et al*., 2014; Iacobuzio-Donahue *et al*., 2009; Hishinuma *et al*., 2006). The reasons for this are multifactorial and include non-specific symptoms resulting in delayed diagnosis and the close anatomical relationship of the pancreas to major blood vessels making tumour invasions and spread a high probability (Evans *et al.*, 2018; McGuigan *et al.,* 2018; Canto *et al*., 2013).

The purpose of this research is to explore ways in which outcomes in pancreatic cancer can be improved through the delivery of more personalised realistic care aided by the application of statistical modeling. Pertinent breakthroughs and controversies in the management of pancreatic cancer will now be explored to contextualise the statistical models developed through the research presented in this thesis. Specific areas that will be covered relate to the key aspects of competing treatment pathways (Figure 1) and include advances and controversies in: 1) the staging of pancreatic cancer and its implications on treatment goals and predicted

outcomes, 2) the application of surgery, 3) adjuvant therapies and 4) neoadjuvant therapies.

As the focus is on management pathways following diagnosis early detection, screening and modifiable risk factors are beyond the scope of this research. This discussion opens with a brief summary of the current understanding of the pathology of pancreatic cancer and is supported by an expanded discussion in appendix A. This is presented not only as background knowledge but also provides context for the discussion within the methods chapter of the disease of pancreatic cancer as a complex system acting within a complex system, the patient, who forms part of a wider complex system, the healthcare system. This also provides context for the discussions regarding the ongoing research focusing on biomarker and gene target therapies. Whilst such developments are in their infancy, and therefore beyond the current scope of inclusion within the statistical models presented within this thesis, they are included to facilitate the discussion within the penultimate chapter of this thesis regarding the future application and impact of the research presented here.

Figure 1: Overview of Treatment Pathways for the Management of Pancreatic Cancer



2.1.1 Pathology of Pancreatic Cancer

Pancreatic ductal adenocarcinoma (PDAC) and its morphological variants account for 90% of all exocrine pancreatic carcinomas (Feldman *et al.*, 2007; Kloppel *et al.*, 2001; Collisson *et al.*, 2019). These variants, recognised by the World Health Organization (WHO) classification of pancreatic tumours, the main ones of which are outlined in table 1, are significant in that they have different histological features, molecular signatures and prognosis (Luchini *et*

al., 2016; Reid *et al*., 2014; Verbeke *et al*., 2016; Hong *et al*., Bosman *et al*., 2018). However, they are currently uninformative with regards to management decisions with many cases of PDAC defined as 'not otherwise specified' (Collisson *et al*., 2019).

Table 1: Summary of variants of PDAC (Bosman *et al*., 2018).

PDAC morphological variant	Characteristics	Prognosis compared to classic pancreatic adenocarcinoma.
Adenosquamous carcinoma	At least 30% component ductal, glandular and squamous differentiation.	Worse
Colloid/mucinous carcinoma	Arise in association with intraductal papillary mucinous neoplasm. Produce excess amounts of extracellular stromal mucin.	Better
Undifferentiated/ anaplastic	Cells appear spindled or sarcomatoid with osteoclast-like giant cells. Very aggressive form with very poor prognosis	Similar
Signet ring cell carcinoma	Very rare. Singularly invasive cell with intracytoplasmic mucin.	
Medullary carcinoma	Pleomorphic epithelial cells with intratumoral lymphoid infiltrate.	Slightly better
Hepatoid carcinoma	Morphology similar to hepatocellular carcinoma. Very rare.	Similar

PDAC develops through a series of step-wise mutations from normal mucosa to precursor lesions (the best characterised of which are: pancreatic intraepithelial neoplasms (the most common), intraductal papillary mucinous neoplasms and mucinous cystic neoplasms) (Esposito *et al.*, 2014) before ultimately becoming invasive malignancy (Mohammed *et al*., 2014). An understanding of this process at molecular level is only just beginning to emerge (Appendix A).

In summary the burgeoning deeper understanding of pancreatic cancer, even when more specifically defined as PDAC, at a molecular level sheds light on the complex and highly heterogeneous nature of this disease (McGuigan *et al*., 2018; Collisson *et al*., 2019). However currently histopathological classifications do not inform clinical decisions as they do in other cancer types (Collisson *et al*., 2019). The purpose of this section and appendix A is firstly to demonstrate how an emerging meaningful and clinically applicable molecular taxonomy could in the near future partly inform clinical decision making (McGuigan *et al*., 2018; Collisson *et al*., 2019) but, as will later be explained, only if this fits within a wider complex system that is the patient within the complex reality of the healthcare system delivering pancreatic cancer care. The second purpose is to demonstrate how the complexity and heterogeneous nature of the disease at molecular level is intrinsically linked with, and contributes to, the uncertainty, ambiguity and complexity that surrounds other key components of the treatment pathway for pancreatic cancer as will now be explored.

2.1.2 Staging Of Pancreatic Cancer

Surgery is the only potentially curative treatment for pancreatic cancer. Furthermore, it has been established that the delivery of multimodal treatment (surgery in combination with either chemotherapy or chemoradiotherapy) within either a neoadjuvant or upfront surgery pathway results in improved survival time (Evans *et al*., 2018). Inherent in the decision to deliver multimodal treatment is first the identification of patients with potentially resectable disease.

Potentially resectable pancreatic cancer can include resectable, borderline resectable and some locally advanced stages of the disease. Whilst diagnostic imaging to determine the tumour relationship to major blood vessels, and specifically in overcoming the challenge of distinguishing tumour vascular invasion from inflammatory changes, has been assisted by structured imaging protocols centering around high resolution computerised tomography (CT) scanning as the first line imaging modality (Al-Hawary *et al*., 2014), staging pancreatic cancer remains problematic as accurate pathological staging can only truly be complete after surgical resection (Wray *et al*., 2005). This matters because the clinical implications of how the staging of pancreatic cancer is defined are potentially mastodonic as will now be illustrated.

An objectively defined staging system allows treatment goals to be more clearly agreed and defined between clinicians and patients resulting in better management of patient expectations (Evans *et al*., 2018). The American Joint Committee on Cancer (AJCC) in

cooperation with the Tumour Node Metastases (TNM) committee of the International Union Against Cancer staging system (table 2) is widely used and is prognostic for overall survival with locally advanced and metastatic disease having a 10 to 12 months and 4 to 6months approximate survival times respectively (Wray *et al*., 2005). However, it has significant limitations in guiding treatment decisions with some, but importantly not all, patients with AJCC stage IVA disease being found to be candidates for surgical resection (Wary *et al*., 2005). This led to clinicians grouping disease as resectable, locally advanced or metastatic based on imaging (Wray *et al*., 2005). As methods of preoperative staging and the understating of the interpretation of these have evolved, a grey zone emerged between tumours that were defined as resectable at presentation and those that were locally advanced (Evans *et al*., 2010). Specifically tumours with short superior mesenteric vein- portal vein (SMV-PV) occlusion and arterial abutment, that were previously defined as locally advanced, were in some cases found to respond to neoadjuvant therapy and therefore considered for surgery. A new definition of borderline resectable disease therefore emerged for such tumours (Evans *et al*., 2010) as importantly these tumours were different from resectable tumours as they carried a higher risk of positive resection margins, required a more complicated surgical resection involving vascular resection and reconstruction, and carried a higher risk of radiologically occult metastatic disease (Evans *et al*., 2010). In 2010 the National Comprehensive Cancer Network (NCCN) adopted a set of guidelines that was established through an expert working group to attempt to establish a universal criteria for resectable, borderline resectable and locally advanced pancreatic disease

(Tempero *et al*., 2017; Callery *et al*., 2009). Whilst this has improved reporting, inter-institutional variations regarding how resectability is defined still exists across trials and continues to impact on decision making (Raufi *et al*., 2019).

Table 2: AJCC Staging and Corresponding TNM Classification

AJCC Stage	TNM Classification
IA	T1 (tumour limited to pancreas and measures<2cm), N0 (no regional lymph node metastasis), M0 (no distant metastases) T1a: ≤0.5cm & <1cm T1b: >0.5cm & <1cm T1c: 1-2cm
IB	T2 (tumour limited to pancreas but measures ≥2cm & ≤4cm), N0, M0
II	T3 (tumour > 4cm extends into duodenum, bile duct or peri pancreatic tissues), N0, M0
III	T1, N1 (regional lymph node metastases), M0; T2, N1, M0; T3, N1, M0
IVA	T4 (tumour extends in to stomach, spleen, colon or celiac axis vessels), any N, M1 (distant metastasis)
IVB	Any T, any N, M1

These issues are further compounded in the neoadjuvant setting where recent studies have called into question the accuracy of current imaging techniques in predicting disease status post neoadjuvant therapy (Katz *et al*., 2012). A review of 122 cases of

borderline resectable pancreatic cancer imaged after neoadjuvant therapy revealed that only one patient had their disease downstaged to resectable yet 85 of these patients actually underwent resection with 81 achieving R0 resection (Katz *et al*., 2012). Although structured reporting strategies have been established in an attempt to facilitate patient selection and mitigate discrepancies, these have yet to be widely adopted (Al-Hawary *et al*., 2014; Raufi *et al*., 2019). More recently there has been a move to subdivide locally advanced tumours into type A, where surgery may be possible after neoadjuvant chemotherapy or chemoradiotherapy, and type B where surgery is unlikely to be possible (Evans *et al*., 2015) (table 3). This has corresponding clinical significance as the likelihood of successful surgical resection following neoadjuvant therapy for resectable, borderline resectable and locally advanced type A and B is estimated as being 90%, 75%, 60% and 25% respectively, although the latter estimate is based on small numbers and could prove to be overly optimistic (Evans *et al*., 2018).

Table 3: Definitions of Resectable, Borderline Resectable and Locally Advanced PDAC Subdivided into Type A and B (Evans *et al*., 2018).

Tumour-Vascular Anatomy				
	Resectable	Borderline Resectable	Locally Advanced Type A	Locally Advanced Type B
Superior Mesenteric Artery (SMA)	No abutment or encasement	≤180° (abutment)	>180° (encasement) but ≤270°	>270° encasement
Celiac Artery	No abutment or encasement	≤180° (abutment)	>180° (encasement), does not extend to aorta, amenable to celiac resection	>180° and abutment/ encasement of aorta
Hepatic Artery	No abutment or encasement	Short-segment abutment or encasement. No extension to celiac artery or hepatic artery bifurcation	>180° (encasement), extends to celiac artery and amenable to vascular reconstruction	>180° encasement with extension beyond bifurcation of hepatic artery into right and left hepatic arteries
Superior Mesenteric Vein-Portal Vein (SMV-PV)	≤50% narrowing	>50% narrowing with distal and proximal target for reconstruction	Occlusion without option for reconstruction	Occlusion without option for reconstruction

Even where surgery is anatomically and technically possible it is the achievement of microscopically negative surgical margins, R0 resection, which determines survival outcome (Kanda *et al*., 2014). This is important as numerous studies have shown that patients who have undergone surgical resection but with microscopically positive resection margins (R1 resection) or macroscopically positive resection margins (R2 resection) have had, outside a neoadjuvant setting, similar survival outcomes to those treated non-operatively (Wray *et al*., 2005). This again highlights the importance of deciding treatment pathway selection at an individual patient level to optimise outcomes.

The survival time for patients with borderline or locally advanced disease that responds to neoadjuvant therapy has improved, as will later be discussed in the section on neoadjuvant therapy. However ambiguity remains across these studies as the definition of R0 and R1 resection is contested with the Union for International Cancer Control and College of American Pathologists defining R1 resection as the presence of microscopic cancer cells at the definite resection margin whilst the Royal College of Pathologists define R1 resection as the presence of tumour within 1millimeter of the resection margin (Kim *et al*., 2017). Different definitions of R0 and R1 resection have therefore been used across studies and the precise definition being used is not always made clear. Despite this the undisputed aim of surgery continues to be R0 resection as this is associated with superior survival outcomes (Kanda *et al*., 2014).

Ultimately however the reality of the narrative becomes much more complicated as each disease stage is likely to have different anticipated outcomes even if they undergo surgical resection with full pathological staging not being possible until after resection. The narrative that surgical resection is the only potentially curative treatment is not realistic for many patients.

There is currently incomplete consensus about the working definition of operable pancreatic cancer in terms of both tumour anatomy, as discussed, and also patient factors (Evans *et al*., 2018). To illustrate, increasing age brings increased co-morbidities but age alone cannot determine operability (Ansari *et al*., 2016a). Many

patients are malnourished at the time of diagnosis, which has implications of an impaired immune system (Argiles, 2005) and low albumin has also been identified as a risk factor for post-operative complications (La Torre *et al*., 2013). More recently it has been recognised that 25% of patients with resectable disease have sarcopenia which has been shown to predict major post-operative complications, longer hospital stay (including intensive care unit stay), and increased risk of infectious and cardiopulmonary complication during post-operative recovery resulting in impaired long-term survival (Joglekar *et al*., 2015). Cardio pulmonary exercise testing (CPET) is widely being used to assist pre-operative assessment but whilst this has helped to identify subgroups at higher operative risk, mortality is low therefore it is not fully adequate as a discriminatory tool (Junejo *et al*., 2014). Obstructive jaundice impairs outcomes following pancreatic resection but more recently it has been discovered that surgery within one week of the diagnosis of obstructive jaundice can actually reduce the overall postoperative morbidity rate (Van Der Gaag *et al*., 2010) making early surgery without biliary drainage the treatment of choice in such circumstances provided serum bilirubin levels are below 300 μmol/l (Tol *et al*., 2015; Sauvanet *et al*., 2015).

In summary the narrative must evolve beyond surgery being the only potentially curative treatment to an understanding that for some patients with operable disease the potential benefits of surgery in terms of disease free survival may be at best limited (Evans *et al*., 2018). Furthermore the decision to operate requires a comprehensive assessment of the individual patient's physical and

mental capacity to cope with the insult of surgery, as well as their personal preferences and treatment goals, to reach a shared decision based on risks and benefit to that patient (Ansari *et al*., 2016a). This requires a move towards better patient selection across competing treatment pathways to deliver individualised treatment sequencing and stage specific therapy to optimise individual patient outcomes. This also poses the challenge as how best to use and then convey such complex information to patients to facilitate the shared decision making process.

2.1.3 Pancreatic Cancer Surgery

In recent years health services have been reorganised so that pancreatic cancer surgery is now mainly performed at high volume centres by experienced surgeons with the resulting increase in surgical expertise being reflected in improved outcomes and morbidity and mortality rates falling to 22.7% and 1.3% respectively (Rohrmann *et al*., 2009; Bliss *et al*., 2014; Gall *et al*., 2015; Evans *et al*., 2018). Enhanced Recovery After Surgery (ERAS) programs have also become part of routine care within these healthcare settings, which has been shown to reduce complications, length of hospital stay and costs (Coolsen *et al*., 2013; Williamsson *et al*., 2015; Ansari *et al*., 2013). However, the ambiguity surrounding disease staging and definitions of operability also permeate decision making even when the decision is taken to proceed with surgical resection as will now be explained.

The majority of PDACs (65%) arise in the head of the pancreas and their removal requires a pancreaticoduodenectomy either in the form of a Whipple procedure or a pylorus preserving pancreaticoduodenectomy (Artinyan *et al*., 2014). Distal Pancreatectomy is performed with splenectomy for PDACs of the body and tail of the pancreas (Gall *et al*., 2015). The main postoperative complications for each procedure are outlined in table 4.

Table 4: Summary of Postoperative Complications

Post-operative Complications of Pancreaticoduodenectomy	Post-operative Complications of Distal Pancreatectomy
Delayed gastric emptying (20-50%) (Lermite *et al*., 2013) Postoperative Pancreatic Fistula (10-15%) (Lermite *et al*., 2013) Wound infection (11%) (Grobmyer *et al*., 2007) Postoperative bleeding (4-16%) (Lermite *et al*., 2013) Anastomotic leaks including biliary fistulae (1-5%)(Lermite *et al*., 2013) Intestinal fistulae (3-8%)(Lermite *et al*., 2013) Pancreatitis (2-3%) (Lermite *et al*., 2013) Ischemic complications (1%) (Lermite *et al*., 2013)	Postoperative Pancreatic Fistula (6-32%) (Pericleous *et al*., 2012) Pancreatic insufficiency with endocrine failure (23.4%) (Iacono *et al*., 2013) Pancreatic insufficiency with exocrine failure (15.6%)(Iacono *et al*., 2013) Risk of sepsis triggered by Streptococcus pneumonia, Neisseria meningitides and Haemophilus influenza 35 times higher than general population (Hansen & Singer, 2001) with incidence of infection 3.2% and mortality 1.4%(Bisharat *et al*., 2001)

As previously discussed tumour anatomy in relation to major vessels determines resectability. The International Study Group of Pancreatic Surgery (ISGPS) presented a consensus statement to prevent borderline resectable tumours being classified as unresectable (Bockhorn *et al*., 2014). Specifically borderline tumours of the head of the pancreas can have venous involvement with narrowing and occlusion of the SMV-PV provided there are no distant metastasis and suitable proximal and distal vessels to allow safe vein resection and reconstruction. The gastroduodenal artery may be encased and there may be short encasement of the hepatic artery but any abutment of the SMA must be bellow 50% with no involvement of the celiac artery (National Comprehensive Cancer Network (NCCN), 2015; Bockhorn *et al*., 2014). For tumours of the body and tail contact with the celiac artery is permitted but only if encasement is less than 50% (NCCN, 2015; Bockhorn *et al*., 2014). Whilst this provides clear guidance on what tumours are now deemed to be technically resectable, the benefit of resecting mesenteric and portal vessels that have been invaded by tumour is controversial (Yu *et al*., 2014; Ansari *et al*., 2016b). Appendix B provides a detailed discussion of the developments and controversies surrounding pancreatic cancer surgery.

To summarise the points being made in more detail in appendix B with regards to optimising outcomes for pancreatic cancer surgery, firstly surgical resection, specifically R0 resection, performed in a high volume specialist centre by experienced surgeons and where an ERAS programme is in place, has been shown to optimise operative outcomes and minimise risk. Secondly arterial resection is not

currently recommended due to the associated risks of morbidity and mortality. Thirdly venous resection has not demonstrated an increased morbidity and mortality risk profile. Whilst venous involvement is not a contraindication to surgery, the impact on survival outcomes of routine venous resection where venous invasion is thought to be probable has not been conclusively established but achieving R0 resection remains the goal of performing resection. Fourthly laparoscopic approach has not been conclusively established as being superior to open approach. Modified approaches to resection have reported increased R0 resection rates but often with higher morbidity and mortality profiles and any reported survival advantage is debatable. The evidence base underpinning these areas of debate (vascular resection, laparoscopic approach, modified techniques) is mainly level III evidence with observational studies potentiating a high degree of selection bias therefore ambiguity prevails.

The importance of establishing this in moving the discussion forward to explore the evidence base underpinning developments in adjuvant and neoadjuvant therapies is that outcomes for such therapeutic approaches are discussed in terms of corresponding resection rates and complications (when the definition of resectability lacks complete consensus and potentially significant exactitudes of the surgical approach are seldom given) and resection margins (the definition of which are debated), all of which impacts on the interpretation of reported survival outcomes.

2.1.4 Adjuvant Therapy

As pancreatic surgery has evolved from a high-risk procedure to a challenging and relatively safe procedure in high volume specialist centres (Kleeff *et al*., 2016; Hartwig *et al*., 2013) it has become evident that surgery alone is not sufficient treatment as greater than 90% of patients who undergo potentially curative surgery will die from disease recurrence without additional therapy (Kleeff *et al*., 2016). The benefits of adjuvant therapy have been established but until recently the optimal treatment regime and modality had not been established with the role of adjuvant chemoradiotherapy remaining controversial (Gall *et al*., 2015; Saif, 2013; Conroy *et al*., 2018). Furthermore the optimal timing and duration of adjuvant therapy has not been established (Gall *et al*., 2015; Saif, 2007; Saif, 2013).

Adjuvant Chemoradiotherapy

Patients with stage I and II disease who undergo surgical resection alone will develop local reoccurrence in 15% of cases and combined local and distal recurrence in 65% of cases (Iacobuzio-Donahue *et al*., 2009). Although the current standard of care is adjuvant chemotherapy, and systematic disease remains the greatest threat to treatment failure, there are patients who could benefit from localised treatment in the form of inclusion of radiotherapy with the rationale

of preventing local recurrence in the pancreatic bed, but who remain difficult to identify (Saif, 2013; Gall *et al*., 2015). Four RTCs have examined the role of adjuvant chemoradiotherapy in cases of resected pancreatic cancer and are summarised in table 5 with a supporting discussion in appendix C critically examining the strength and limitations of these trials.

Table 5: Summary of RCTs of Adjuvant Chemoradiotherapy

Trial	Treatment Arms (n=)	Resection Margin	Median Survival in Months (*P* value)
GITSG	Observation (22)	R0	11
	5-FU radiotherapy + 5-FU (21)		20 (*P* = 0.035)
EORTC	Observation (54)	R0/R1	12.6
	5-FU chemoradiation radiotherapy (60)		17.1
ESPAC-1	No chemoradiotherapy (178)	R0/R1	16.1
	Chemoradiotherapy (175)		15.5 (*P*= 0.235)
	No chemotherapy (235)		14.0
	Chemotherapy (238)		19.7 (*P*= 0.233)
RTOG 97-04	5-FU, 5-FU-based radiation + 5-FU (230)	R0/R1	16.9
	Gemcitabine, 5-FU-based radiation + gemcitabine (221)		20.5 (*P*= 0.9)

Adjuvant Chemotherapy

The results of key RCTs of adjuvant chemotherapy for resected pancreatic cancer are outlined in table 6 and is supported by a critical analysis of these trials in appendix D. Until recently what had emerged from RCTs was a preference for gemcitabine based adjuvant regimes but with controversies surrounding toxicity profiles of regimes and variation in follow-up strategies. However, the PRODIGE24/CCTG trial compared gemcitabine to mFOLFIRINOX (a combination of oxaliplatin, irinotecan and leucovorin) for patients who had undergone R0 and R1 resection (Conroy *et al*., 2018). Interim analysis at 33.6months has demonstrated that mFOLFIRINOX was associated with improved disease-free survival (21.6months *versus* 12.8months) and overall median survival (54.4months v 35month) compared to gemcitabine (Conroy *et al*., 2018). This trial has resulted in the American Society of Clinical Oncology (ASCO) Clinical Practice Guidelines to be updated to recommend the following:

"all patients with resected pancreatic adenocarcinoma who did not receive preoperative therapy should be offered 6 months of adjuvant chemotherapy in the absence of medical or surgical contraindications. The modified combination regimen mFOLFIRINOX as used in the latter part of the PRODIGE 24/CCTG PA.6 trial (oxaliplatin 85 mg/m^2, leucovorin 400 mg/m^2, irinotecan 150

mg/m^2 D1, and 5-fluorouracil 2.4 g/m^{2}over 46 hours every 14 days for 12 cycles) is preferred in the absence of concerns for toxicity or tolerance; alternatively, doublet therapy with gemcitabine and capecitabine or monotherapy with gemcitabine alone or fluorouracil plus folinic acid alone can be offered." (Conroy *et al*., 2018, p.1)

Table 6: Summary of RCTs of Adjuvant Chemotherapy

Trial	Treatment Arms (n=)	Resection Margin	Median Survival in Months (*P* value)
CONKO-001	Observation (175)	R0/R1	10.4
	Gemcitabine (179)		20.7 (*P*=0.01)
CONKO-005	Gemcitabine (217)	R0	26.6
	Gemcitabine + erlotinib (219)		24.6
ESPAC-3	Gemcitabine (539)	R0/R1	23.6
	5-FU (551)		23.0 (*P*=0.39)
JASPAC-01	Gemcitabine (190)	R0	25.5
	S-1 (187)		46.5 (*P*=<0.0001)
ESPAC-4	Gemcitabine (366)	R0/R1	25.5
	Gemcitabine + capecitabine (365)		28.0 (*P*=0.032)
APACT (n=866)	Gemcitabine (NS)	R0/R1	36.2
	Gemcitabine + nab-paclitaxel (NS)		40.5 (*P*=0.045)
PRODIGE24/CCTG	mFOLFIRINOX	R0/R1	54.5

	(247)		
	Gemcitabine (246)		35 (*P*=0.003)

2.1.5 Neoadjuvant Therapy

Although the survival benefit of adjuvant therapy has been established, between 71% and 76% of patients will have disease recurrence within 2years of surgical resection (McGuigan *et al*., 2018). Furthermore up to 50% of patients will not be suitable to progress to adjuvant therapy postoperatively due to early disease recurrence and/or a decline in physiological function following the insult of major surgery with or without treatment complications (Winter *et al*., 2012; Ghosn *et al*., 2016; McGuigan *et al*, 2018). Such figures, coupled with the evidence of the success of neoadjuvant therapy in treating other forms of gastrointestinal cancer (primarily rectal, esophageal and gastric cancers) has resulted in neoadjuvant therapy emerging as a potential strategy for the treatment of borderline resectable pancreatic cancer (Bockhorn *et al*., 2014), and also cases of resectable and locally advanced disease (Neoptolemos *et al*., 2019; Gillen *et al*., 2010; Labori *et al*., 2017).

The rationale for neoadjuvant therapy is that pancreatic cancer is a systemic disease therefore radiographic imaging, whilst providing accurate information on primary tumour to vessel relationship, underestimates radiologically occult micrometastatic disease (Asare *et al.*, 2016) hence systemic treatment should be initiated earlier in the treatment process. It follows that the theoretical benefit of

neoadjuvant therapy is the elimination of micrometastases and shrinkage of the primary tumour which reduces the incidence of recurrence, increases conversion to resectability of borderline and locally advanced cases and increases R0 resection rates (Zhan *et al*., 2017). However, this approach also carries the risks that patients will experience toxicities that could delay surgery, or the tumour may progress making previously resectable disease unresectable (Zhan *et al*., 2017; Lopez *et al*., 2014; Asare *et al*., 2016). This makes neoadjuvant approach for resectable pancreatic cancer an area of prime controversy. Further concerns have been raised that neoadjuvant chemoradiotherapy could cause pancreatic fibrosis which may increase the operative complication rate (Lopez *et al*., 2014).

One of largest meta-analysis of 111 neoadjuvant studies, comprising 4,393 patients with PDAC, was undertaken by Gillen *et al*. (2017) and showed that amongst patients with initially unresectable disease, 46.6% underwent surgical exploration with 69.9% having their tumour resected and 79.2% of these patients having R0 resection. The median overall survival was 20.5months, which was comparable to those presenting with resectable disease. However, this analysis was not performed on an intention-to-treat basis, which potentiated bias in treatment effect as not all patients proceed to surgery (Raufi *et al*., 2019). Furthermore the studies included pre-dated the NCCN definition of resectability, the reporting of which could therefore vary widely across trials. D'Angelo *et al.* (2017) performed one of the first intention-to-treat analysis of 12 prospective neoadjuvant studies that included a total of 624 patients with resectable, borderline resectable

and locally advanced PDAC. It reported a similar median overall survival of 22.78months and a resection rate of 65%.

Neoadjuvant FOLFIRINOX has recently provided some hope with studies reporting improved conversion rates to resectability in borderline resectable and locally advanced PDAC (63.5% and 22.5% respectively of all those patient presenting with these stages of disease) in a meta-analysis of 13 studies comprising 253 patients. (Petrelli *et al.* 2015). Whilst 85% of the 43% of those with either borderline resectable or locally advanced PDAC who underwent surgery achieved R0 resection, this meta-analysis, based on available evidence, could not yet conclude a definite improvement in overall survival with neoadjuvant FOLFIRINOX. Although Suker *et al*. (2016) focused only on cases of locally advanced PDAC in their meta-analysis of 11 studies, comprising 315 patients treated with neoadjuvant FOLFIRINOX, they reported similar resection rates of 25% but a median overall survival of 24.2months.

One of the pivotal unanswered questions is whether neoadjuvant therapy offers a survival advantage over traditional upfront surgery followed by adjuvant therapy for cases of resectable pancreatic cancer. Meta-analysis by both Xu *et al*. (2014) and Andriulli *et al*. (2012) reported only marginal benefit of neoadjuvant chemotherapy in terms of overall and disease free survival in resectable cases. However, neither of these reports focused solely on neoadjuvant therapy therefore omitted significant studies from their meta-analysis (Lee *et al.* 2016). More recently meta-analysis by Versteijine *et al*. (2018) pooled 38 trials comprising 3,484 patients with

resectable and borderline resectable disease in an intention-to-treat analysis. Their findings also reported only marginal benefit with neoadjuvant approach over upfront surgery approach for resectable disease (18.2months *versus* 17.4months) but a greater survival advantage with borderline resectable disease (19.2months *versus* 12.8months). For both resectable and borderline resectable cases the rates of R0 resection were higher with neoadjuvant therapy at 85% versus 71.4% and 88.6% versus 63.9% respectively. Mokdad *et al*. (2017) also reported a survival advantage with neoadjuvant approach in their retrospective analysis of National Cancer Database using propensity score matched analysis of neoadjuvant therapy used to treat 2,005 patients with stage I and II PDAC compared to 6,015 patients who underwent upfront surgical resection of PDAC (26months *versus* 21months). However, this analysis is heavily biased as only those who tolerated neoadjuvant therapy and underwent resection were included in the neoadjuvant group.

The evidence base underpinning neoadjuvant therapy lacks high-quality phase III RCTs and is currently largely based on phase II trials as well as observational cohort studies (which are mainly small, prone to single centre bias, underpowered and with a high degree of heterogeneity) and the meta-analysis of these studies (Neoptolemos *et al*., 2019, McGuigan *et al*., 2018; Versteijine *et al*., 2018). Treatment therapies and dosing regimes vary widely across studies as do definitions of resectability and classification of resection margins, despite the introduction of more established definitions. These factors in addition to how meta-analysis studies group together the analysis of outcomes of resectable, borderline resectable and locally

advanced disease, all of which have different anticipated outcomes that could affect treatment selection and decision making, place limitations on existing data synthesis studies. This mandates a closer critical look at the existing evidence base of prospective phase II drug trials for each disease stage at presentation and is provided in appendix E.

2.1.6 Conclusion: The Ongoing Complexities of Neoadjuvant Therapy for Pancreatic Cancer

The studies discussed in this section and the corresponding appendicies presents the current state-of-play in pancreatic cancer research and highlights the main areas of debate surrounding the treatment of pancreatic cancer. As the disease of pancreatic cancer is beginning to be understood at a molecular level what is emerging is an understating of a highly complex and heterogeneous disease with overlapping defects in genes and signaling pathways between what was previously thought to be clearly defined disease subtypes (Collisson *et al*., 2019). As a molecular taxonomy emerges it is hoped that in future this will help to target treatments. However, this can only ever partly inform clinical decision making and currently our understanding of the disease at this level has not informed clinical decision making in the way an understating of other cancers at molecular level has (Collisson *et al*., 2019). Furthermore it is not known whether aggressive tumour biology or anatomical location close to major vessels, or indeed a combination of both factors, accounts for the propensity of pancreatic cancer for metastatic

spread. This has implications for both staging of the disease and surgery.

Although the NCCN provided a more universal definition of resectable, borderline and locally advanced disease in 2010 there remains a high degree of inter-institutional discrepancy (Evans *et al*., 2018). This definition has also been challenged about its accuracy in re-staging disease after the completion of neoadjuvant therapy (Katz *et al*., 2012). More recently there has also been a call to subdivide locally advanced disease into type A and type B to better manage patient expectations surrounding the anticipated outcomes of neoadjuvant therapy (Evans *et al*., 2018). Operability also depends on patient factors and their physical and mental reserve to cope with major surgery yet despite some advances in patient assessment this aspect of decision making remains largely subjective (Ansari *et al*., 2016).

Staging cannot be completed until after pathological assessment of the resected specimen. However, discrepancies also exist between definitions of R0 and R1 resection margins. The primary goal of surgery is an R0 resection. It has been established that centralising services so that pancreatic cancer surgery is only performed at large volume specialist centres by experienced surgeons with an ERAS programme in place improves operative outcomes. However, the decision of whether to perform resection of the veins, when, and in which patients is debated (McGuigan *et al*., 2018).

In light of findings from the PRODIGE 24/CCTG PA.6 trial guidelines have recently been updated regarding adjuvant therapy to recommend mFOLFIRINOX or, where this is contraindicated, gemcitabine and capecitabine or monotherapy with gemcitabine alone or fluorouracil plus folinic acid alone (Conroy *et al*., 2018). However, a significant proportion of patients who undergo resection will not receive adjuvant therapy. This has raised the possibility of neoadjuvant therapy as an alternative treatment pathway. This is particularly controversial for cases of disease that are defined as resectable at presentation. Existing trials in neoadjuvant therapy are mainly small, unpowered and based on single institution data making comparison of data between trials extremely difficult considering variability in treatment regimes and dosing, and metrics defining resectable, borderline resectable and locally advanced disease as well as R0 and R1 resection (Neoptolemos *et al*., 2018; McGuigan *et al*., 2018). These factors compound the challenge of trying to compare neoadjuvant and upfront surgery pathways for resectable and borderline resectable disease when few existing studies offer such a comparison and those that do tend to combine resectable and borderline resectable cases within the neoadjuvant arm.

Decision making within this arena is eminently challenging, as is the task of delivering personlised realistic medicine. Whilst it is hoped that on-going genomic and drug trials will provide some answers, they alone cannot achieve this. It could be argued that the current problem being presented is that of a complicated system with a high degree of uncertainty. The case for viewing the challenges outlined here in terms of a complex system will be made further in the

methods chapter. In order to build such a case, the existing body of research seeking to assist clinical decision making in pancreatic cancer management will now be examined first in terms of health economics modelling and then in terms of prognostic and predictive modelling.

2.2 Health Economics Modeling To Guide Decision Making In The Management of Potentially Resectable Pancreatic Cancer

Introduction

The cost of cancer care has risen astronomically over recent years (Greenberg *et al.* 2010). Cancer is estimated to cost the European Union economy €126 billion with €51 billion (40%) spent on healthcare alone (Luengo-Fernandez *et al.* 2013). Within the UK estimated annual costs of cancer range from £15.8 billion to £18.33 billion (£7.6billion premature death and work absence, £5.6billion on healthcare, £2.6billion on unpaid care) with this figure predicted to rise to £24.72 billion by 2020 (Department of Health, 2015). Given contemporary financial constraints and subsequent limitations on healthcare resources, it is therefore unsurprising that cost-effectiveness analysis of new treatment approaches is beginning to receive increased attention in medical and health economics literature (Greenberg *et al.,* 2010; Russell, 2016).

Over two decades ago Elixhauser & Halpern (1999) commented on a paucity of literature on the economics of pancreatic cancer. Unfortunately little has changed with many of the initial subsequent studies focusing on specific interventions such as surgery (Lea & Stahlgren, 1987; Brandabur *et al*., 1988; Gudjonsson *et al*., 1995; Holbrook *et al*., 1996; Raikar *et al*., 1996; Topal *et al*., 2007; Enestvedt *et al*., 2008; Jeurnik *et al*., 2010; Waters *et al*., 2010), or chemotherapy and radiotherapy (Glimelius *et al*., 1995; Ishii *et al*.,

2005; Miksad *et al*., 2007; Danese *et al*., 2008; Krzyzanowska *et al*., 2007). As new knowledge of the disease and its treatment has emerged many of these studies are no longer clinically relevant. Of note, disease that was previously thought of as unresectable now has the potential of conversion to resectability, particularly with the emergence of neoadjuvant therapy, yet surgery has not been included as an alternative treatment strategy in studies involving what would now be classified as borderline resectable or locally advanced disease. This is important as not all patients who undergo conversion to resectability see an increased effectiveness in terms of post resection survival time whilst others experience a significant advantage in terms of postoperative survival time. Previous USA based cost analysis have shown that of all pancreatic cancer disease stages, resectable disease carries the highest costs ranging from $65, 335 (Du *et al*., 2000) to $134,700 (O'Neill *et al*., 2012) with the latter study including all costs reimbursed by Medicare as well as an older population. Whilst these studies have several limitations (failure to assess which health services were specifically related to pancreatic cancer, exclusion of costs not covered by Medicare, exclusion of indirect costs and the latter findings only being applicable to an older population) they did provide important insights into the impact of the underutilisation of surgery to their findings. Notably Caucasian patients and those in affluent urban areas were more likely to receive resection, which O'Neill *et al*. (2012) highlighted as suggesting higher costs if all eligible candidates received a resection (Riall & Lillemoe, 2007; Bilimoria *et al*., 2007b; O'Neill *et al.*, 2012). This raises several further questions. Firstly how applicable are cost analysis studies from countries with privatised healthcare systems to countries

where healthcare is free at the point of delivery, such as the UK National Health Service(NHS)? Secondly when synthesising data from drug trials from countries with privatised healthcare systems the inherent bias in these studies must be acknowledged as many patients with pancreatic cancer will never present to such internationally renowned centres due to socioeconomic factors. Thirdly what impact will developments that are aimed at increasing the rate of pancreatic cancer resection (the development of screening programs for earlier disease detection, targeted therapies and the developments of more effective neoadjuvant regimes) have on both survival outcomes and costs, and will they prove to be cost effective? Fourthly where disease is resectable at presentation, is it more effective and/or cost effective to adopt a neoadjuvant or traditional upfront surgery approach to treatment?

Despite the fact that pancreatic cancer is associated with a short life expectancy the costs incurred in a short period of time are substantial (O'Neill *et al.*, 2012). In the current economic climate the ambiguity surrounding many aspects of the treatment pathway for potentially resectable pancreatic cancer, as outlined in the previous section, mandates cost-effectiveness evaluation of treatment choices, particularly the role of neoadjuvant therapy (Abbott *et al.,* 2013). If however value in healthcare is to be defined as value relative to cost, then it must be acknowledged that outcomes in cancer are neither static nor universal and can be highly individual (Russell, 2016). Successful outcomes could be the number of months of survival whilst to others it is the quality and not quantity of survival time that defines successful outcome (Russell, 2016). Similarly length of

disease free survival may represent successful outcome to some whilst others define success as overall survival time regardless of treatment requirements (Russell, 2016). Costs too go far beyond costs of a particular treatment option but include emotional and monetary costs to patients and healthcare systems associated with complications, readmissions, rehabilitation etcetera, as well as wider societal costs in terms of absence from work for both patients and those undertaking informal caring roles (Russell, 2016). Accepting these complexities and challenges this begs the questions: what do we know about the overall cost-effectiveness of the treatment of potentially resectable pancreatic cancer, how have we measured this, and could this help to inform clinical decision making and/or methods of modeling to support better shared clinical decision making?

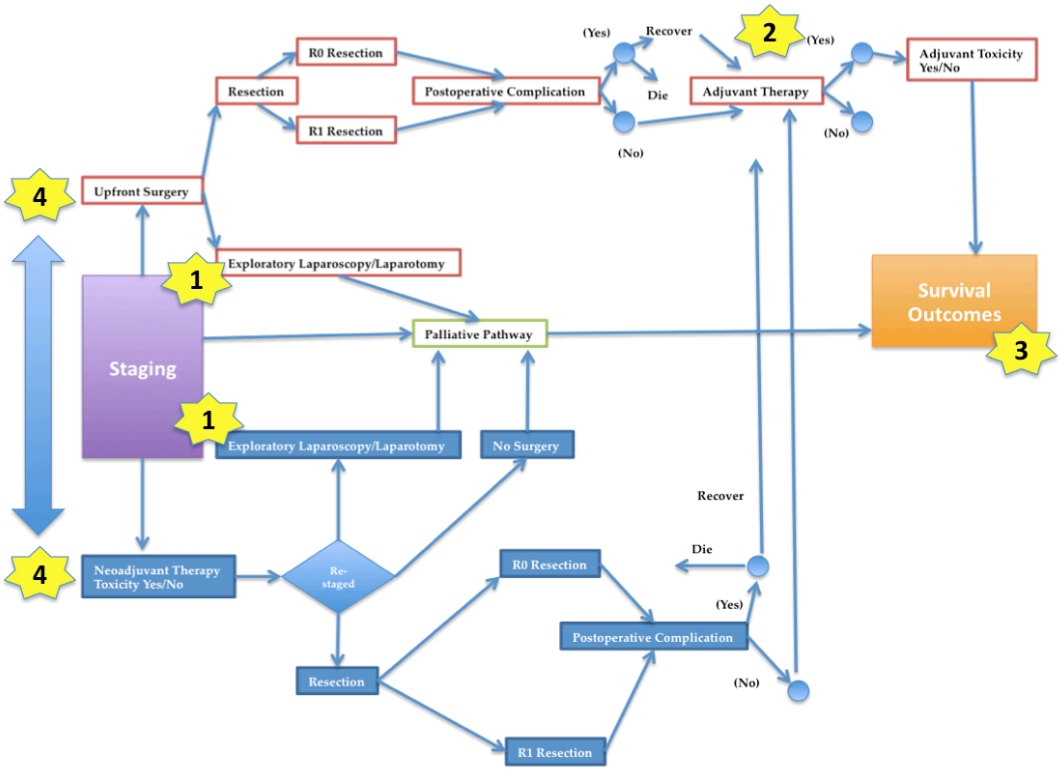
This section is structured as follows. First a critical analysis of cost-effectiveness studies pertaining to the management of potentially resectable PDAC is presented. Secondly, a wider review of cost-effectiveness analysis studies of neoadjuvant *versus* traditional upfront surgery for other solid organ malignancies is presented. The rationale for this is that the evidence base underpinning the management of these malignancies is better established. Therefore the impact of better quality, more certain data on modelling techniques for cost-effectiveness analysis will be critically analysed to ascertain whether this results in better quality of analysis or whether commonalities in flaws prevail. Finally this section concludes with a summation of strengths and limitations of the

current body of literature and how methods of statistical modeling could be improved and applied to the research question.

2.2.1 Cost-Effectiveness Analysis of the Management of Potentially Resectable PDAC

Existing cost-effectiveness analysis studies that have relevance to contemporary clinical practice fall into key areas of the management pathway that include: 1) staging strategies, 2) adjuvant therapy, 3) post resection follow-up strategies and 4) neoadjuvant therapy *versus* upfront surgery approach (Figure 2).

Figure 2: Overview of the Treatment Pathways for the Management of Potentially Resectable Pancreatic Cancer with the Focus of Previous Cost-Effectiveness Analysis Highlighted: 1: cost-effectivness of disease staging strategies, 2: cost-effectivness of adjuvant therapies, 3: cost-effectiveness of follow-up strategies, 4: cost-effectivness of upfront surgery versus neoadjuvant approach



Staging

Current guidelines for the staging of pancreatic cancer recommend pancreatic protocol CT scan including chest, abdomen and pelvis (National Institute of Clinical Excellent (NICE), 2018). Where disease is found to be localised a fluorodeoxyglucose-positron emission tomography/CT (FDG-PET/CT) should be offered to patients who

will be having further treatment in the form of surgery, systemic therapy or radiotherapy (NICE, 2018). If more information is then required to determine an individual's clinical management the following are recommended in specific circumstances: Magnetic Resonance Imaging (MRI) scan where liver metastases are suspect, endoscopic ultrasound scan (EUS) where more information is required for tumour and node staging, and diagnostic laparoscopy with laparoscopic ultrasound where resectional surgery is considered to be a possibility but small volume peritoneal and/or liver metastases are suspected (NICE, 2018). Existing economic analysis studies pertaining to the staging of potentially resectable PDAC focus on FDG-PET/CT and diagnostic laparoscopy.

The Role of FDG-PET/CT

The role of FDG-PET/CT in improving patient selection and being cost-effective has been supported by two key economic analysis (Heinrich *et al*., 2005; Ghaneh *et al*., 2018). Ghaneh *et al*. (2018) went further by not only providing a comprehensive description of the competing alternatives but also in including relevant costs and consequences, measured accurately and in appropriate units, for alternatives identified in their analysis. Heinrich *et al*. (2005) performed a cost-benefit study and failed to perform discounting. Incremental analysis of costs and consequences of alternatives was not performed and, whilst a sensitivity analysis was undertaken this did not account for uncertainty in the estimates of costs and consequences. Ghaneh *et al*. (2018) was a methodologically superior study using Markov modelling to perform both an incremental

analysis of costs and consequences of alternatives as well as an extensive deterministic and probabilistic sensitivity analysis that included measures of the impact of uncertainty in the estimates of costs and consequences. Both studies provided corroborating findings that FDG-PET/CT was cost-effective in improving patient selection for resection surgery.

The Role of Diagnostic Laparoscopy

The role of diagnostic laparoscopy is more ambiguous and controversial. As the NICE guidelines (2018) state, laparoscopy with laparoscopic ultrasound should be performed where resectional surgery is considered to be a possibility but small volume peritoneal and/or liver metastases are suspected. There is some debated evidence that diagnostic laparoscopy could avoid unnecessary exploratory laparotomy. However, ambiguity exists as to the optimal timing of and between diagnostic laparoscopy and, where appropriate, exploratory laparotomy and whether there is any benefit in it becoming routine practice or if and how patient selection for this procedure could be improved. These issues reflect many of the broader issue and ambiguities concerning the management of potentially resectable pancreatic cancer: how to improve patient selection and more effectively target interventions to optimise outcomes in the face of uncertainty.

A detailed critical analysis of cost-effectiveness analysis studies of staging diagnostic laparoscopy for pancreatic cancer is provided in appendix F to assess how statistical modeling can handle such issues.

In summation this analysis highlights that there is a lack of high quality studies exploring the cost-effectiveness of staging diagnostic laparoscopy. Existing studies that report cost-benefit from diagnostic laparoscopy do so only if diagnostic yield is assumed to be high. In an age of advanced imaging modalities and neoadjuvant therapy presumably making occult metastases potentially less likely, the assumption on which these results are made must be questioned. Furthermore based on existing literature there is not clear evidence as to the cost-effectiveness implications of the timing of staging diagnostic laparoscopy (same admission *versus* separate day-case), and its application to patient sub-groups (routine *versus* only patients at high risk of occult metastatic disease).

Adjuvant Therapy

The survival benefits of adjuvant therapy have long been established however cost-effectiveness analysis of competing adjuvant regimes is limited. Two studies were found that provide economic assessment of adjuvant therapy. Abbott *et al*. (2012) compared surgery and adjuvant therapy to: no treatment, surgery only, radiotherapy only, chemotherapy only and chemotherapy combined with radiotherapy for resectable pancreatic head adenocarcinoma using a decision tree approach. This study demonstrated what was already known: surgery and adjuvant therapy is more expensive but yields greater utility. Neoadjuvant therapy was not considered as a competing treatment strategy and the clinical impact of such a study is limited as, for resectable disease, not performing surgery would only be considered as a viable competing treatment option if the patient had

other mitigating circumstances, such as extensive comorbidities. However such factors were not considered within the model. This study was also further limited by the fact that it populated the model with data from multiple databases including a national cancer data registry. Not only does this make the data highly heterogeneous in terms of patients, adjuvant therapy regimes and variations in outcomes between high and low volume centres but, within the USA healthcare system, also carries the potential of bias due to disadvantaged patient groups either not presenting to such renowned institutions or presenting with more advanced disease stages. Costs were taken from Medicare payments yet not all patients would have been covered by public payers, which means such payment data is not generalisable to all patients. Furthermore, although this study is from a societal/payer perspective, indirect costs were excluded. Other significant costs that were excluded included: readmission after surgery, complications of treatment and end-of-life care. Costs and benefits were not discounted and quality adjusted survival outcomes were based on the limited published quality of life data. Such limitations are compounded by the fact that only one-way deterministic sensitivity analysis was performed with no probabilistic sensitivity analyses or other measure of impact of uncertainty surrounding model parameters provided. Willingness-to-pay (WtP) and other such thresholds were also lacking which limits the usefulness of this study.

The second study utilised data from the ESPC-4 RCT within a Markov model to compare gemcitabine monotherapy with gemcitabine combined with capecitabine in patients who had undergone

complete resection of pancreatic cancer (Huang *et al*., 2018). It concluded that although gemcitabine and capecitabine provided 1.23 quality-adjusted-life-years (QALYs) compared to gemcitabine monotherapy, which provided 1.02 QALYs, the incremental cost-effeectivenss ratio (ICER) was $45,191.23, which surpassed the WtP threshold for that country (Huang *et al*., 2018). Both direct and indirect costs were included but palliative care costs were excluded and discounting was not applied. Quality-of-life adjustments for utility outcomes were also based on published literature, which is limited for pancreatic cancer. However this study did perform both deterministic and probabilistic sensitivity analysis including WtP thresholds and measurements of the impact of uncertainty within the model. This study, whilst being the first to compare adjuvant therapies, used data from patients who had undergone resection and were well enough to meet the inclusion criteria for adjuvant therapy within the ESPAC-4 trial. Real-world events such as the discovery of unresectable disease at the time of surgery, or impact of postoperative complications and their impact on overall pathway analysis are therefore not captured. Furthermore, in light of the findings of the PRODIGE 24/CCTG PA.6 trial (Conroy *et al*., 2018), which resulted in mFOLFIRINOX becoming the first line adjuvant therapy over gemcitabine-based alternatives, this study has limited impact. What has yet to be established is whether the increased toxicity profile associated with mFOLFIRINOX has any affect on the cost-effectiveness of this compared to gemcitabine based adjuvant therapy, or whether better patient selection between these competing regimes could maximise cost-effectiveness.

Follow-up Post Resection

Until recently surveillance after potentially curative resection of pancreatic cancer was not considered to be a key issue due to poor survival outcomes and the lack of second line treatments. One study assessed the cost-effectiveness of competing surveillance strategies from a payer's (societal) perspective (Tzeng *et al*., 2013). It used a Markov model to compare: no scheduled surveillance, 6 monthly clinical assessment with Ca 19-9 levels, 6 monthly clinical assessment with Ca 19-9 levels + CT + Chest x-ray, 3 monthly clinical assessment with Ca 19-9 levels, 3 monthly clinical assessment with Ca 19-9 levels + CT + Chest x-ray. This study reported that surveillance beyond 6 monthly clinical assessment with Ca 19-9 levels increased cost but with no clinically significant survival benefit (Tzeng *et al*., 2013). However, the model was populated with retrospective data from a single institution, which carries a risk of bias and limits generalisability of findings particularly as this database only included patients treated within a neoadjuvant pathway. Furthermore costs were taken from Medicare data, but not all patients would have been covered by public payers (Tzeng *et al*., 2013), and discounting was not applied. Outcomes between competing strategies may have been affected by lead time and length time bias related to surveillance intervals and the diagnosis of indolent asymptomatic disease *versus* symptomatic aggressive disease (Tzeng *et al.,* 2012). Only deterministic sensitivity was performed therefore the impact surrounding the degree of uncertainty within model parameters was not fully assessed. Furthermore sensitivity analysis did not account

for the fact that 5-10% of patients do not produce Ca 19-9 (Ballehaninna *et al*., 2012).

Neoadjuvant Therapy versus Upfront Surgery

Two studies compared neoadjuvant and upfront surgery pathways (Abbott *et al*., 2013; Choi *et al*., 2018). Both pose a clearly defined, answerable question: what is the cost-effectiveness of neoadjuvant approach compared to traditional upfront surgery approach. Abbott *et al.* (2013) used a decision analytic model with utility reported as quality-adjusted-life-months (QALMs) whilst Choi *et al.* (2018) used a Markov model with results reported in QALYs. Quality-of-life literature related specifically to pancreatic cancer is limited and therefore in both studies a reliance on the few published quality-of-life indices was not ideal.

In both studies the patient populations were clearly defined. Abbott *et al*. (2013) populated the upfront surgery arm with data drawn from American College of Surgeons National Cancer Database (NCDB 2003-2005) and National Surgical Quality Improvement Programme (NSQIP 2005-2009). However, the neoadjuvant group was drawn from the MD Anderson database 2002-2008. Effectively this meant comparing two different databases. Data from literature was used to populate data points otherwise unavailable and they included phase III RCTs for the upfront surgery arm. Choi *et al.* (2018) synthesised data from published literature but did not provide details of the literature search strategy or quality assessment of included studies. The neoadjuvant arm also included data taken from a single

institution database. In both studies the use of different data sources introduces uncertainty and bias. Furthermore the limitations and ambiguities in existing literature have already been explored in section 2.1 and the supporting appendicies. Institutional data carries bias as patients could have better outcomes due to referral bias, self selection, receiving superior care in larger specialist centres and furthermore, a key subset of patients may never present to these databases due to socioeconomic factors and/or co-morbidities (Abbott *et al.* 2013, Choi *et al.,* 2018). Overall this increases the level of uncertainty as not all patients within these various data sources received uniform care.

In both studies cost data was taken from payer perspective with costs based on Medicare payment estimates and technical and professional services costed from Centers for Medicare and Medicaid Services (CMS). In reality not all costs would have been covered by CMS and this also limits transferability of findings to alternative healthcare systems (Abbott *et al.,* 2013). Hospital payments were estimated based in ICD-9 DRG codes. Costs not included in the Abbott *et al*. (2013) study were: readmission post surgery, complications associated with chemotherapy or radiotherapy, follow-up surveillance and hospice costs. The latter two are potentially a significant omission considering that this study reported significantly prolonged survival time with neoadjuvant therapy. Choi *et al.* (2018) did include the cost of palliative care and treatment complications. Neither study included indirect costs and Choi *et al*. (2018) did not apply discounting of either costs or benefits.

One-way deterministic sensitivity analysis was performed in both studies. In the Abbott *et al.* (2013) study sensitivity analysis was only performed on the upfront surgery group and included adjustment for alternative billing, cost of adjuvant chemotherapy, elimination radiotherapy, survival with node and margin negative resections, rates of perioperative mortality, complications and finding unresectable disease at surgery with justification provided for each form of sensitivity analysis. Choi *et al*. (2018) examined the effect of alterating resection rates, surgical mortality, recurrence rate, cancer mortality and utility values across both treatment strategies. They also performed probabilistic sensitivity analysis to assess the impact of uncertainty across model parameters on model output. This level of analysis was lacking in the Abbott *et al.* (2013) study.

Abbott *et al.* (2013) included three possible neoadjuvant regimes within the neoadjuvant arm of their study (gemcitabine + cisplatin +radiotherapy, or chemoradiotherapy based on either gemcitabine or cisplatin, or capecitabine-based chemoradiotherapy) compared to adjuvant gemcitabine in the upfront surgery arm. Choi *et al.* (2018) compared neoadjuvant FOLFIRINOX to adjuvant gemcitabine monotherapy or gemcitabine/capeciabine in the upfront surgery arm. The conclusions drawn from both studies, that neoadjuvant therapy is more cost-effective than upfront surgery, clearly reflect the results of the studies but, particularly with the Abbott *et al.* (2013) study, the uncertainty surrounding these conclusions were not assessed considering the high degree of heterogeneity within the neoadjuvant arm alone. Furthermore both studies are limited by the

high degree of uncertainty, heterogeneity and quality issues associated with the existing published literature. Neither study examined the role of adjuvant mFOLFIRINOX which, considering the results of the PRODIGE 24/CCTG PA.6 trial (Conroy *et al*., 2018) and subsequent change to practice guidelines making this the first line adjuvant therapy, does now question these conclusions particularly as the survival times reported from both adjuvant therapy cohorts within the PRODIGE 24/CCTG PA.6 trial (Conroy *et al*., 2018) rivals those reported in the neoadjuvant arm of both of these studies.

Conclusion

In the introduction to this section several questions were posed: what do we know about the overall cost-effectiveness of management options for potentially resectable pancreatic cancer, how have we measured this, and could this help to inform clinical decision making and/or methods of modeling to support better shared clinical decision making?

Firstly what we know about the overall cost-effectiveness of the management of potentially resectable pancreatic cancer is limited and permeated with ambiguity. Issues pertaining to methodological quality and the relevance of findings, as new evidence emerges of more effective interventions, limit many existing cost-effectiveness studies. Whilst it is established that pancreatic surgery is expensive, and that improved patient selection for surgery would improve both costs and quality adjusted outcomes, both the most effective and cost-effective way of achieving this is widely debated, arguably with

the exception of FDG-PET/CT for staging. Secondly how we measure cost and effectiveness is also a contested area considering that quality-of-life data for pancreatic cancer is limited. The dominant methods of analysis are decision trees and Markov models. However the quality of data used to populate such models was found to be an issue across many studies with few models performing a full probabilistic sensitivity analysis to gauge the degree and impact of uncertainty across model parameters on outcomes. Included and excluded costs varied widely and discounting across existing studies was sporadic. Ultimately therefore, to answer the third question posed, existing cost-effectiveness analysis cannot yet be said to inform shared clinical decision making. Many of the issue identified in these cost-effectiveness analysis studies emanate from the state of the current evidence base underpinning the management strategies for potentially resectable pancreatic cancer. However whilst it would be convenient to believe that this is the only culprit, the question must be asked as to whether alternative approaches to modeling could better handle the existing data, including its inherent uncertainties. To fully explore this possibility a review of cost-effectiveness analysis studies of neoadjuvant approach for other solid organ malignancies, which have a more established evidence base, must first be undertaken.

2.2.2 Cost-Effectiveness of Neoadjuvant Approach to Cancer Treatment: what have we really learned over the past decade?

The purpose of this section is to: 1) to establish what is currently known about cost-effectiveness analysis of neoadjuvant therapy applied to all solid organ malignancies (SOM) 2) critically appraise research methodology of existing studies and 3) highlight areas and direction for future cost-effectiveness and decision analysis research.

In total 13 studies published since 2000 were identified that have performed cost-effectiveness analysis of neoadjuvant therapy for SOM (pancreatic cancer (previously discussed): n=2, upper gastrointestinal (GI) cancers n= 3, colorectal cancer n=3, cervical n=1, breast: n=1, ovarian: n=2, bladder: n=1). This comprises the cost-effectiveness analysis of a total of 27 neoadjuvant regimes applied to 9 types of SOM (pancreatic: n=4, esophageal: n=1, peritoneal carcinomatosis from gastric cancer: n=4, hepatocellular carcinoma: n=1, colorectal: n=4, cervical: n=1, breast: n=7, ovarian: n=2, bladder: n=3) (table 7).

Table 7: Neoadjuvant regimes included in each study.

Study	SOM	NAT Regime	Alternative
Gordon *et al.*, 2012	Esophageal	NAT regime not specified but separate decision arm included adding FDG-PET to NAT regime	T2-T4 tumours: surgery without NAT or no surgery, treatment with chemoradiation only
Hultman *et al.*, 2012	Peritoneal carcinomatosis from gastric cancer	Irinotecan + 5-FU +LV or Irinotecan + deGramont schedule or decetaxel + 5-FU/LV or EOX all followed by CRS + HIPEC +EPIC	Systemic palliative chemotherapy alone
Vitale *et al.*, 2010	Hepatocellular carcinoma	Sorafenib	No bridging therapy prior to liver transplant
Ercolani *et al.*, 2011	Colorectal with liver metastases	FOLFOX4	Surgery (Hepatectomy) first
Poston *et al.*, 2001	Colorectal with liver metastases	Oxaliplatin + 5-FU/FA or 5-FU/FA as NAT	Comparing 2 NAT regimes
Van der Brink *et al.*, 2004	Rectal	5x5 Gy	Surgery (TME) first
Rocconi *et al.*, 2005	Cervical	Cisplatin, bleomycin and vincristine	Surgery first or primary chemotherapy
Attard *et al.*, 2015	Breast	Neosphere NAT regimes: (trastuzumab+ docetaxel or Pertuzumab+ trastuzumab + docetaxel or Pertuzumab + trastuzumab or Pertuzumab + docetaxel) and TRYPHAENA regimes (FEC+ Docetaxel) (6 cycles or 3 cycles)+ Pertuzumab (6 cycles or 3 cycles) or Docetaxel + Carboplatin + Trastuzumab + Pertuzumab	Comparing NAT regimes from two studies
Poonawalla *et al.*, 2015	Ovarian	NAT (regime not specified)	Primary debulking surgery
Rowland *et al.*, 2015	Ovarian	NAT+ surgery + carboplatin + paclitaxel	Primary debulking surgery + carboplatin + paclitaxel
Stevenson *et al.*, 2014	Bladder	MVAC or gemcitabine + cisplatin, or gemcitabine + carboplatin	Surgery (radical cystectomy) first

Included studies and their methodologies are summarised in table 8. A detailed critical appraisal of each study structured according to the checklist propose by Drummond *et al.* (2015) is presented in appendix G.

Table 8: Summary of cost-effectives analysis studies of neoadjuvant therapy for solid organ malignancies

Study	SOM	Comparators	Design/ Methodology	Economic Perspective	Data Source	Benefit Measures & Time Horizon	Cost Source	Costs Excluded	Sensitivity Analysis
Gordon *et al.* (2012)	Esophageal	T2-T4 non-surgical v SF v NAT	Retrospective. Decision-analytic model	Payer perspective	ACS Adelaide & Brisbane database	QALMs 5 years post surgery	National price schedules, public hospital clinical costings	Indirect costs	Yes: Monte Carlo sensitivity analysis
Hultman *et al.* (2012)	Esophageal	NAT + CRS + HIPEC + EPIC v palliative systemic chemotherapy	Retrospective. Kaplan-Meir + bootstrap sampling	Treatment costs	Patients identified in unit (not randomized) and matched with patients from RCT	QALYs 2005-2009	Uppsala Unit Hospital data, Swedish National Pharmacy pricelist 2008	Indirect costs	Yes
Vitale *et al.* (2010)	Hepatic cell carcinoma	NAT pre LT v no bridging therapy	Retrospective Markov Decision Model.	Payer perspective	Literature review	HR for delay to recurrence, QALDs, LT probability, cost utility ratio, incremental health benefit. 10 years	Italian public healthcare system	Indirect costs	Yes: Monte Carlo sensitivity analysis
Ercolani *et al.* (2011)	Colorectal liver metastases	NAT v SF	Retrospective Markov Decision Model	Societal perspective	Literature review	QALMs, ICER, HR, WtP, RFS. 10 years	Italian public healthcare system	Indirect costs	Yes: Univariant and two-way
Van der Brink *et al.* (2004)	Rectal	NAT v SF	Prospective + RCT. Markov Decision Model	Provider and societal perspective	RCT	LE, QALY, cost per lifetime, ICER. 1996-1999	Price index from Dutch Healthcare Sector	Nil	Yes
Poston *et al.* (2001)	Colorectal liver metastases	Oxaliplatin + 5-FU/FA v 5-FU/FA as NAT	Decision-analytic model.	Provider perspective	Literature review	Incremental cost per life year gained. 6 months follow-up	NHS	Drug administration costs, postsurgical costs including palliative care.	Yes
Rocconi *et al.* (2005)	Cervical	RHYST v CTRT v NAT	Retrospective. Decision-analytic model	Third party payer perspective	Literature review	Cost per cure/ survivor, 5year DFS. 5 years	Local charges	Indirect costs/ reimbursements	Yes: one-way sensitivity analysis
Attard *et al.* (2015)	Breast	NAT: pertuzumab and trastuzu	Retrospective cost-utility	Canadian healthcare payer	NeoSphere and TRYPHAENA trial	LYG, QALYs, ICER, 28 years	NeoSphere and TRYPHAENA	Not stated. Indirect costs	Yes: probabilistic sensitivi

		mab	analysis using Markov decision model.	perspective.	data		trial data, Hoffman-La Roche Unit costs, published sources or Ontario databases.		ty analysis (PSA)
Poonawalla *et al.* (2015)	Ovarian	NAT v SF	Retrospective cohort study	Payer perspective	Surveillance, Epidemiology and End Result (SEER)-Medicare linked database.	Cumulative treatment costs with phase-of-care approach, ICER, OS, NMBs, LYG, 2000-2009	Medicare claims	Indirect costs,	No
Roawland *et al.* (2015)	Ovarian	NAT v SF	Markov decision model	Healthcare system perspective	Literature review	OS, surgical complications, probability of initiation, treatment cost, QoL, 5 yeats	Medicare + hospital costs estimates from Agency for Healthcare Costs and Utilization Project (HCUP) data.	Surveillance, chronic complications and indirect costs.	Yes: Monte Carlo simulation
Stevenson *et al.* (2014)	Bladder	NAT v NAT	Kaplan-Meir, log-rank test, *t* test.	Third party payer perspective	Retrospective review of institutional data	QALY 2004-2011	Local billing, published sources	Indirect costs	None

NAT= neoadjuvant therapy; SF= surgery first

What this critical analysis showed is that the complexities of costs involved in cancer care mean that this is a challenging yet essential area of health economic research. Neoadjuvant therapy represents an emerging and unique area for cost-effectiveness analysis with implications spanning the trajectory of patients' journeys, and impacting far beyond the neoadjuvant phase of treatment. Emerging challenges brought by the advent of neoadjuvant approach to cancer

treatment as well as areas where cost-effectiveness analysis could be improved have been highlighted. Firstly comparability of studies is inhibited by variation in: methods, economic perspective, cost estimates, and sporadic use of discounting. Secondly, neoadjuvant therapy is an emerging treatment option therefore available patient databases and existing published studies regarding its effectiveness carry limitations. How these limitations are addressed within cost-effectiveness analysis studies relies on utilising any of the plethora of techniques of evidence synthesis available which simultaneously raises questions regarding uncertainty and bias in the data undergoing analysis and could further impede comparability of findings for decision makers. Thirdly, although widely used, reporting of QALYs is also shrouded in controversy and ambiguity surrounding quality-of-life measurement. Each of these areas and how future research could be improved will now be addressed.

Comparability of cost-effectiveness analysis studies is essential for decision makers to evaluate trade-offs therefore factors impeding comparability must be analysed. All studies in this review gave details of how costs were arrived at but few studies explicitly detailed costs that were excluded (Appendix G; Table 8). Costs of cancer care are multifactorial involving cost of treatment, individual monetary and emotional costs, costs to healthcare systems and costs to wider society (Russell, 2016) yet only one study did not exclude indirect costs (van der Brink *et al.,* 2004). This corroborates findings that despite earlier recommendations (Gold *et al.,* 1996) only 29% of cost per QALY analysis since 2005 adopted a societal perspective

(Neumann, 2009) and even when they reported to do so important costs were omitted (Sanders *et al.,* 2016). Although some studies converted costs to US dollars, it must be remembered that changing currency does not equate with conversion of actual costs within different healthcare systems or how people in different countries value their health status (Simunovic *et al.,* 2004). Also, cost-effectives ratios and WtP thresholds do not inform decision makers on resources required to implement neoadjuvant therapy and which other interventions should be abandoned to free resources to enable this (Simunovic *et al.,* 2004). Therefore, whilst researchers should compare cost and benefits of treatment approaches, cost-effectiveness cannot be deduced unless opportunity costs of selected treatments (i.e cost-efficient use of redirected resources) are also determined (Simunovic *et al.,* 2004). How then can future studies standardise methodological practices to improve quality and comparability, whilst addressing the theoretical challenge of aggregating costs and effects across different sectors and individual patients and their carers, in a way that reflects consensus position at societal level (Drummond *et al.,* 2015; Brouwer *et al.,* 2008)?

One recommended solution is that all cost-effectiveness studies should include as standard two reference cases: one from health sector perspective and one from societal perspective (Sanders *et al.,* 2016). From health sector perspective effects would be measured in QALYs and results summarised in ICER, net monetary benefit (NMB) and /or net health benefit with a range of cost-effectiveness thresholds considered. Costs would include all health care sector costs reimbursed by third party payers and out-of-pocket costs paid

by patients. The societal reference case would then consider consequences of an intervention including those outside formal healthcare sector therefore would include indirect costs such as transport, patients' and carers' time costs and reduced productivity etc. In practical terms for cost-effectiveness analysis research this would involve a standardized 'impact inventory' that lists all direct and indirect costs making costings more transparent, comprehensive and rigorously assessed (Sanders *et al.,* 2016). The studies included in this review had a wide variation in time horizon and this approach could also go some way to addressing the challenge of estimating future direct and indirect costs during any additional life years gained, although this is an area some would argue requires further research (Sanders *et al.,* 2016).

A further source of variation is that of discounting. Whilst it is recommended that costs and health effects be discounted at the same rate ambiguity surrounds what that rate should be. American recommendations currently stand at 3%. In the UK NICE recommends 3.5% for both costs and benefits (NICE, 2013) but previously recommended 6% for costs and 1.5% for benefits (NICE, 2011). Sensitivity analysis accounting for a range of discounting rates is therefore recommended whilst further research in this area continues (Sanders *et al.,* 2016).

This review found that cost-effectiveness analysis of neoadjuvant therapy for the treatment of SOM utilised a variety of methods although Markov decision models and decision analytic models dominated. Whilst such approaches offer methods for dealing with

uncertainty through evidence synthesis, this review revealed that there was limited explanation of how, or indeed if, such models dealt with heterogeneity and potential bias of the data used to populate these models (Appendix G). Some databases lacked important details related to tumor characteristics or specifics of interventions, which could potentiate bias as could the application of survival analysis, when used in the presence of censored data. Furthermore the majority of studies were retrospective. Future studies should focus on performing economic analysis prospectively, possibly within clinical trial, to ensure high internal validity (Drummond *et al.,* 1997; Simunovic *et al.,* 2004). Where evidence is synthesised a quantitative description and critique of the evidence base must be offered with explicit detail about how bias within and across studies was handled, bias corrected estimates arrived at, and how estimates were adjusted for transferability (Sanders *et al.,* 2016). As was the case with most studies, ambiguities should be tested through sensitivity analysis. In studies of higher quality sensitivity analysis was used to test every assumption or estimate used in the decision model to account for potential impact of such variations on the results and this should be standard practice across future studies.

With the exception of one study in this review (van der Brink *et al.,* 2004) quality-of-life data was not collected as part of the analysis hence introducing limitations in accuracy of QALYs. Whilst some would debate the ability of QALYs to capture all-important factors impacting quality of life, such as short lived but intense experiences, generic preference-based measures would enhance comparability of

findings (Sanders *et al.,* 2016). The proviso is that such instruments exist and are fit for the purpose of measuring differences and changes across interventions being considered (Sanders *et al.,* 2016). Where such instruments do not yet exist, analysts may present quality-of-life estimates based on scores from patients and/or other sources (Sanders *et al.,* 2016).

In summary, current cost-effectiveness analysis of neoadjuvant therapy are limited by factors impeding comparability, limitations of available evidence supporting effectiveness of treatment options and quality-of-life measures on which to base quality adjusted survival outcomes. Reference cases, from both health sector and societal perspectives, introduced as standard reporting of and using standardised methodological practices could promote comparability of future studies, with a set 'impact inventory' improving transparency of included costs (Sanders *et al.,* 2016). The evidence base on which cost-effectiveness analysis is based must also be routinely critiqued with quantitative description of evidence base, accounting for bias within and between studies, offered as the basis on which bias adjusted estimates are calculated (Sanders *et al.,* 2016) and each assumption tested in sensitivity analysis. Attention to collecting quality-of-life data would enhance future studies particularly if included in prospective studies.

2.2.3 Conclusion: lessons from modeling in health economics

To conclude, the flaws highlighted in the reviews of cost-effectiveness analysis pertaining to pancreatic cancer surgery (methodological flaws and heterogeneity limiting comparability and generalisability of findings, uncertainty and quality issues pertaining to data sources and how effectiveness is measured and reported) prevail and permeate through cost-effectiveness analysis of neoadjuvant therapy applied to other malignancies, even where the underlying evidence base for neoadjuvant approach is more matured. This challenges the erroneous assumption that better quality data would automatically equate with better, more useful statistical models of outcome optimisation in pancreatic cancer treatment. However, the lessons to be drawn are more expansive than simply the technicalities of how to improve cost-effectiveness studies and could have much more latitudinous connotations for future research.

Firstly, the limitations of current cost-effectiveness analysis studies must be understood within the context of the limitations of the available evidence-base determining the degree of 'effectiveness' of the intervention, which is after all the 'driver' for cost-effectiveness analysis (Ades *et al*., 2006). Whilst it is widely accepted that RCTs and their meta-analysis provide the highest form of evidence (Garas *et al*., 2012; Centre of Evidence Based Medicine, 2011), the fallibility of this perceived hierarchy must be acknowledged (Ades *et al*., 2006; Garas *et al*., 2012). Firstly, RCTs can report varied outcomes and may not provide all evidence required. Furthermore the inevitable lack of infallible, appropriately designed RCTs reflecting real-life patient

case-mix and complexities of decision making in clinical practice, does not equate with avoiding or improving cost-effectiveness analysis (Ades *et al*., 2006). Increasingly decision analysis is being used as a framework for economic evaluation (Ades *et al*., 2006). How then can clinicians and researchers best make use of the existing imperfect evidence base whilst simultaneously considering the impact of such weaknesses on uncertainty of decisions and for research priorities? (Ades *et al*., 2006; Claxton et al., 2002)

Rather than solely seeking to address the gaps in current literature with further RCTs one additional solution may lie in evidence synthesis; a growing area of interest within cost-effectiveness analysis (Ades *et al*., 2006; Garas *et al*., 2012; Claxton *et al*., 2002). Evidence synthesis is a collective term covering the diversity of methods and mathematical tools utilised for integrating data from a variety of sources into decision and cost-effectiveness analysis models (Ashrafian *et al*., 2010). This approach has been championed as producing evidence with greater accuracy and less uncertainty by utilising data from multiple types of studies (Ashrafian *et al*., 2010). However, as demonstrated by the previous review of existing cost-effectiveness analysis studies, challenges arise when accounting for heterogeneity, degree of bias and uncertainty when combining multiple sources of available evidence in cost-effectiveness analysis (Ades *et al*., 2006).

In response to such challenges Bayesian approach to meta-analytical methodology is gaining precedence (Ades *et al*., 2006; Garas *et al*., 2012; Ashrafian *et al*., 2010; Felli & Hazen et al., 1999). Bayesian

approach is based on the concept of using available evidence to accurately derive the probability of a parameter (Ashrafian *et al*., 2010). It therefore provides the flexibility to incorporate variable randomised and non-randomised data sources in a hierarchical model with sources individually weighted to account for bias and uncertainty (Ades *et al*., 2006; Garas *et al*., 2012; Ashrafian *et al*., 2010; Spiegellhalter *et al*., 2004; Sutton & Abrams, 2001). Consequently this approach is beginning to play a pivotal role in decision modeling (Tapper *et al*., 2011; Garas *et al*., 2012; Spiegellhalter *et al*., 2004; Cooper *et al*., 2002; Cooper *et al*., 2004; Parmigiani, 2002). At present the application of this methodology is stymied by the complex mathematical expertise it demands (Ashrafian *et al*., 2010). Advances in software supporting Bayesian approach juxtaposed with an increasing focus on cost-effectiveness analysis of health interventions, makes this is a rapidly expanding area of research with Bayesian methods being incorporated into the field of machine learning to support decision making. However, the review of cost-effectiveness analysis studies revealed that a significant number of studies used Markov modeling, a statistical model derived from the Bayesian school of statistics, and yet significant flaws prevail. If research is to advance the lessons learned therefore must go beyond those of statistical modeling methodology to consider the theory driving current research and how this relates to its current limitations.

If we consider the basis for Ulrich's seminal work on critical systems thinking (Ulrich, 1983) which was that the definition of a problem, proposals for improvement and outcome are all dependent on the

whole system (Ulrich, 2002). This places great emphasis on how systems boundaries are justified and the implications this has for what modeling a system defined in such a way will, and importantly will not, tell us. The implications of boundary setting across the previously discussed cost-effectiveness analysis studies relate to their limitations and manifest as: exclusion of important alternative treatment strategies, the exclusion of certain costs including indirect costs forfeited by the patient, exclusion of consideration of all relevant potential implications of a treatment strategy including treatment failure and side effects, a lack of quality adjusting survival time or collecting quality-of-life data to more accurately do so, and the setting of time horizons to capture all necessary events. This enables the limitations of the existing body of cost-effectiveness studies not merely to be seen as a series of methodological issues to be corrected, but rather as the system that is the delivery of healthcare being defined in simplistic and reductionist terms dictated by the researcher's agenda which therefore not only defines effectiveness in their terms but also how system boundaries are set which in turn determines how effectiveness is assessed. If a move towards personalised realistic medicine is to be achived then outcomes must instead be defined in more patientcentric ways with models created that can encompass this refocusing.

Problem structuring thinkers, such as Ulrich, argue that systems boundaries must be rationally justified through dialogue with both the involved and affected (Ulrich, 2002; Ulrich, 2012; Ulrich, 1987). Cilliers combined thinking about boundaries with concerns relating to complexity (Kruger *et al*., 2019). Both he and problem structuring

thinkers such as Ulrich agree that both limited knowledge of systems as a result of boundaries and complexity exist and therefore require a critical and ethical imperative in the study and understanding of such systems (Kruger *et al*., 2019). Although this work predated the concept of realistic personalised medicine its relevance is tangible. This view is supported by a growing move within healthcare research to view healthcare systems as complex adaptive systems which have been formally defined as "a collection of individual agents with freedom to act in ways that are not always totally predictable, and whose actions are interconnected such that one agent's actions change the context for other agents" (Plsek & Greenhalgh, 2001 p.625).

To conclude, it is not without coincidence that the field of complex systems developed at a time when statistical theory began to coalesce with machine learning to reliably infer models with large numbers of variables that interact in complex, non-linear ways. It would therefore seem that the potential of, for example, Bayesian statistics has not been fully explored within the area of modeling for cost-effectiveness analysis where the system being modeled has been so reduced and simplified. Therefore the next section will assess how predictive models to support clinical decision making by predicting outcomes have been developed and used, and to what degree of success. After a critical overview of existing models a more detailed examination of the methodological quality of prognostic development studies is offered. After this the specific use of machine learning to support clinical decision making in the management of pancreatic cancer will be critically analysed.

2.3 Predictive Modeling to Support Clinical Decision Making in the Management of Potentially Resectable Pancreatic Cancer.

Publications resulting from this section:

Bradley, A., Van Der Meer, R. and McKay, C.J. (2019) 'A systematic review of methodological quality of model development studies predicting prognostic outcome for resectable pancreatic cancer'. *BMJ Open,* 9**:**e027192. doi: 10.1136/bmjopen-2018-027192

Bradley, A., Van der Meer, R. and McKay, C. (2019) 'Personalized pancreatic cancer management: a systematic review of how machine learning is supporting decision-making'*. Pancreas,*48 (5). pp. 598-604.

Introduction

Traditionally assessment of operative risk has been the domain of surgeons' judgment gained from experience (Lewis & Volmer, 2012). However, there exists a great need to risk-stratify surgical patients pre-operatively in an objective and standardised way (Lewis & Volmer, 2012). This is particularly pertinent in the high precision field of pancreatic cancer surgery where surgical volume is low, with only approximately 10% of cases being resectable at presentation, yet operative mortality and morbidity rates are high (Lewis & Volmer, 2012). Despite advances in surgical technique and adjuvant treatments, the potential benefits of such high-risk surgery are often

nullified by early disease reoccurrence. Effective patient selection for surgery is therefore paramount. Furthermore studies have shown high discrepancies between surgical and survival outcomes in favour of large volume centres (Birkmeyer *et al.,* 2002) which highlights the need for accurate methods of performance adjustment through risk stratification and predictive modeling (Lewis & Volmer, 2012). This makes pancreatic surgery an ideal vector through which to deliver solutions to the complex challenges encountered in prognostic modeling in organ specific surgery.

Options for the management of resectable cases of pancreatic cancer are also becoming more complex with the advent of neoadjuvant therapy. In the absence of large RCTs conclusively proving benefit of either upfront surgery or neoadjuvant approaches to treatment, there exists a need for predictive models to address competing treatment options. These changes have also taken place within a wider socioeconomic context. Prognostic models and risk stratification tools are not only expected to guide treatment approaches but also guide cost-effective use of resources by diverting patients away from unhelpful treatments and investigations. Furthermore there is a move within contemporary healthcare towards personalised predictive medicine whereby probabilistic modeling is used to forecast individual patient outcomes (Velikova *et al.,* 2014; School *et al.,* 2013).

In summary, risk stratification and prognostication are vital in empowering informed consent, supporting clinical decision making, guiding treatment options and patient counseling as well as offering

powerful research tools (Lewis & Volmer, 2012). The following section is structured thus: Section 2.3.1 begins by offering an overview of the role of predictive models in contemporary pancreatic surgery practice through a brief outline of the historical perspective of risks stratification and prognostic modeling to where we are today. From this platform a critical analysis of current methods of predictive modeling will be presented. Section 2.3.2 then takes this discussion further through a systematic critical review of the methodological quality of existing prognostic model development studies. From this basis the application of Bayesian networks as an alternative modeling technique is discussed. Bayesian networks have also been applied within the wider discipline of machine learning therefore section 2.3.3 critically examines how, and to what extent, the application of the emerging discipline of machine learning has been, and could be, applied to the issue of supporting clinical decision making and achieving personalised realistic medicine in the management of potentially resectable pancreatic cancer. Importantly the optimism surrounding this approach is weighted against its current limitations. From this basis the case is made that if research is to progress what is required is not merely an improvement of the application of statistical modeling techniques but rather a revolution in the prevailing Weltanschauung resulting in a fundamental shift in the philosophical paradigm underpinning future research.

2.3.1 Predictive Models For The Surgical Management of Pancreatic Cancer

Predictive models published since 2000 pretainng to the surgical management of pancreatic cancer fall into the following broad categories: predicting mortality and morbidity from pancreatic cancer surgery, complication specific predictions following pancreatic surgery and predicting survival time following pancreatic cancer surgery.

The critical analysis of existing predictive and prognostic models offered in appendix H shows that despite a growing interest in prediction research and its methodologies (Altman & Riley, 2005; Altman, 2007; Altman & Lyman 1998; McShane *et al*., 2005; Rothwell, 2008; Moons *et al.,* 2009; Bouwmeester *et al.,* 2012) there is a lack of rigorous application within surgical centres and wider surgical literature of predictive and prognostic models (Lewis & Volmer, 2012). This is in part due to methodological issues: the inclusion of a wide variety of variables the importance of which clinicians making the decisions may dispute, the use of small single centre data which limits generalisability and the lack of external validation.

Currently the most sophisticated medical predictive models are based on non-liner regression techniques; primarily logistic regression and Cox regression (Lewis & Volmer, 2012). Conversely personalised precision medicine, whereby predictive and prognostic modelling is used to forecast individual patient outcomes, is gaining precedence within contemporary healthcare (Velikova *et al.,* 2014;

School *et* al., 2013) and creates an expectation for models to facilitate decision making and, given the wider socioeconomic context, also guide cost-effective use of resources. A disparity between expectations and the reality of currently available models therefore exists. Juxtapose these growing expectations with the advent of neoadjuvant therapy making treatment options for resectable pancreatic cancer more complex, and it becomes clear that methods of predictive and prognostic modelling must be rigorously assessed if such challenges are to be overcome, as poor methods can result in unreliable and biased results (Bouwmeester *et al.,* 2012).

This research focuses on optimising outcomes for patients with potentially resectable pancreatic cancer. As outcomes are most often measured in terms of survival time, the following section therefore analyses the methodological quality of prognostic model development studies applied to resectable PDAC.

2.3.2 Methodological Quality of Prognostic Development Studies for Resected Pancreatic Cancer

An overview of the current state of prognostic model development studies relating to prognosis following resection of PDAC is presented in appendix I. Areas for improvement and direction for future research have been highlighted by assessing each domain of the Checklist for critical Appraisal and data extraction for systematic Reviews of prediction Modeling Studies (CHARMS) checklist across the 15 included studies (Moons *et al*., 2014). Theses areas for improvement related to general aspects of model development and

reporting, applicability of models and sources of bias (Moons *et al*., 2014).

General Aspects of Models Development and Reporting

General reporting of aspects of model development was found to be clear relating to participant eligibility, recruitment and description as was reporting of follow-up period. Definitions of outcome and number and type of candidate predictors were also generally clearly reported across included studies. Although the number of participants was clearly reported, the number of events at defined time periods of prediction should be more clearly reported to assist assessment of statistical power. Improvement should also be made in the reporting of missing data. The majority of studies used complete case analysis but only 2 of the remaining studies provided details of missing data per variable (Brennan *et al*., 2004; Botsis *et al*., 2009). Across all 15 studies modeling methods were clearly reported. Alternative presentations of models were also offered in all studies to assist application to clinical practice with discussion on strengths, limitations and comparisons also offered.

Applicability

Generalisability of prognostic models is an area for improvement as the majority of models were based in single centre databases. The applicability of these models to patients in neoadjuvant treatment pathways has also not been assessed.

Methods of reporting model performance showed high heterogeneity with only 9 studies providing confidence intervals with results making comment on general applicability difficult (Shen *et al*., 2018; Balzano *et al*., 2017; Dasari *et al*., 2016; Pu *et al*., 2018; Brennan *et al*., 2004; Xu *et al*., 2017; Botsis *et al*., 2009; Smith & Mezhir, 2014; Pu *et al*., 2017). Most models had limited discriminatory performance with area under the curve (AUC) below 0.7 and those reporting an AUC nearing 0.9 being based on small sample sizes therefore raising the possibility of overfitting. The 2 studies employing alternative methods of artificial neural network (ANN) (Walczak & Velonovich, 2012) and Bayesian modeling (Smith & Mezhir, 2014) did not report an improved AUC (0.66 and 0.65 respectively). Furthermore calibration, a crucial aspect of model development, was frequently missing or not performed adequately with the calibration curve based on the derivation dataset (Xu *et al*., 2017). In cases of poor validation whether the model was adjusted or updated was also poorly reported. Only 3 studies performed external validation (Shen *et al*., 2018; Balzano *et al*., 2017; Dasari *et al*., 2016) and none of the studies explored impact analysis of their models making comment on the clinical application of the models difficult. Moving forward this could be addressed through access to datasets from meta-analyses of individual participant data, or registry databases containing electronic health records (Riley *et al*., 2016). Such big datasets would allow researchers to externally validate, and where needed improve through recalibration, model performance across different settings, populations and subgroups (Riley *et al*., 2016).

Source of Bias

Areas for improvement were also found in limiting sources of bias. As previously mentioned overuse of single centre databases is one area but also the reporting of consecutive sampling, number of participants who refused participation, and whether all consecutive participants were included should be more clearly reported. Although handling of candidate predictors, and predictors in modeling, were generally clearly reported including statistical methods for handling categorisation and non-binary variables, their assessment generally did not involve blinding to outcome. Assessment of statistical power of sample size was also not well reported and only 2 studies used the recommended approach of imputation methods to handle missing data with the majority of studies employing complete case analysis which could both potentiate bias and reduce statistical power (Moons *et al*., 2014). None of the included studies gave details on how candidate predictors were identified. In selecting predictors for inclusion in the models the majority of studies employed pre-selection through univariable analysis followed by multivariable analysis. Whilst such an approach is commonplace it does potentiate overfitting of models, an issue poorly discussed across all studies. Only 3 studies included external evaluation (Shen *et al*., 2018; Balzano *et al*., 2017; Pu *et al*., 2018) and classification measures (sensitivity, specificity, predictive value) were poorly reported, as was comparison of distribution of predictors including missing data.

In summary, at a time when an increasing focus and expectation is being placed on personalised predictive medicine, this review highlights fundamental aspects of the methodological quality of models that must be improved if future models are to have a clinical impact by supporting decision making. Whilst many of the models included in this review provided alternative presentations to assist in their clinical application, issues of methodological quality were found that inhibited their clinical impact. These issues included how missing data is handled, the assessment of statistical power, issues of bias associated with candidate predictor selection and a lack of blinding during their assessment. Such issues are augmented by an over reliance on single centre databases which also limits the generalisability of the models. The reporting of model performance is also a key area for improvement. The emerging focus on precision medicine means that the future application of predictive modeling lies in combining each patient's genomic and clinical data in a meaningful way that will support clinical decision making at individual patient level. This can only be achieved if future research focuses on improving the methodological quality of model development, regardless of whether they employ traditional or machine learning methods.

2.3.3 Conclusion: Lessons Learned and Future Direction of Research for Predictive Models in Pancreatic Cancer Surgery

A bamboozling yet flawed array of predictive models and risk stratification tools pertaining to pancreatic surgery exist (Appendix H; Appendix I). Predictive models and risk stratification tools are

widely used in audit and research in many other areas to allow case-mix adjustment when comparing single centre outcomes (Steyerberg *et al.,* 2010) and in defining inclusion and exclusion criteria for RCTs or identifying high risk participants and allowing covariate adjustment (Lewis & Volmer, 2012). However, this review has demonstrated that it is in the area of clinical application that predictive modeling arguably holds most promise yet demonstrates most limitation.

Predictive models can support clinical decision making and assist patient counseling (Braitman & Davidoff, 1996) hence empowering informed consent processes and shared decision making. In the absence of clear contraindications to surgery but where surgeons are faced with difficult decisions about whether to operate or not, predictive models can provide objective predictions about the patient's physiological and immunological responses to surgery (Jarnagin *et al.,* 2011; Christou, 1994). Intra-operative and post-operative application of predictive models can also alter the course of treatment (Lewis & Volmer, 2012). For example a patient identified as being at high risk for developing a pancreatic fistulae may therefore receive more aggressive prophylactic measures (Lewis & Volmer, 2012). Equally lower risk surgical patients could be diverted away from unnecessary referrals or investigations allowing better resource utilisation (Altman & Royston, 2000). The emerging focus on precision medicine means that there will be a demand on future applications of predictive modeling to merge patient's genomic and clinical data to assist decision making on a more individualised basis (Lewis & Volmer, 2012).

The reality however is that such models are not yet in existence and current predictive models are limited in scope and value with most only being descriptive in probabilities of adverse events or survival outcomes (Lewis & Volmer, 2012). Whilst this may help to manage patient expectations, existing models fall short in differentiating patients who would, and more importantly would not, benefit from particular treatment options (Lewis & Volmer, 2012). Furthermore some studies have shown that such models are no better than experience led judgment in predicting morbidity (Markus *et al.,* 2005; Hartley *et al.,* 1994). This corroborates conflicting findings regarding the accuracy of widely used models to predict post-operative morbidity and mortality from pancreatic surgery (Lewis & Volmer, 2012). This is also reflected in the limited application of predictive models within surgical centres and also the lack of rigorous application of predictive modeling in surgical literature (Lewis & Volmer, 2012).

Whilst the plethora of available, disease specific prognostic and risk prediction models may infer a growing interest in the area of predictive modeling, to integrate fully into clinical practice they need to provide predictions beyond length of survival or risk prediction to include fundamentals such as quality of survival time, length of hospital stay, resource utilisation and predicted benefits of competing treatment options available. In short, and echoing the previous conclusions drawn from the review of statistical modeling for cost-effectiveness analysis, predictive models must develop to

engage with the complexity of the system they are attempting to model.

The problem with existing modeling techniques is that they regard prognosis as an isolated event at a pre-determined time, applying attribute selection prior to inducing the model and setting fixed roles of input and output variables to attributes (Verduijn *et al.,* 2007). Variables deemed important by clinicians may therefore be excluded. Furthermore this neglects the dynamic nature of care processes where outcomes today predict those of tomorrow hence expected patient outcomes evolve, as more information becomes available (Verduijn *et al.,* 2007).

Bayesian statistical approach offers an alternative to traditional frequentist paradigm of null hypothesis testing by allowing the integration of prior qualitative and quantitative knowledge (Velikova *et al.,* 2014; School *et al.,* 2013; Verduijn *et al.,* 2007). In this way Bayesian Networks (BN) allow the modeling of relationships between variables at various stages of the healthcare process, with predictions of outcomes evolving throughout the process by utilising all available patient data at that time (School *et al.,* 2013). Predictions can therefore be made for all variables, not just outcome variables (Velikova *et al.,* 2014; School *et al.,* 2013; Lucas *et al.,* 2004). However, despite the potential of BN and the expanse of software supporting their application, their use within healthcare remains under utilised.

BN are based on graphical formalism of a joint or multivariate probability distribution over a random set of variables and are sometimes referred to as acyclic directed graphs (Velikova *et al.,* 2014; School *et al.,* 2013; Stajduhar & Dalbelo-Basic, 2010). In plain language the structure and parameters of BN are nodes, arcs and conditional probabilities. Variables in a BN are modeled as nodes and directed arcs represent causal relationships between nodes. Each node has a conditional probability formula or table that represents the probability of each value contained within that node given the condition of all its parent nodes. Through Bayes theorem the prior distribution and observed data are combined to update knowledge in the form of the posterior distribution (Velikova *et al.,* 2014; School *et al.,* 2013; Stajduhar & Dalbelo-Basic, 2010). Where patient information is limited probabilistic inference can still make predictions based on global averages of the patient population (Verduijn *et al.,* 2007; Lucas *et al.,* 2004). As more information becomes available the predictions become more patient specific (Verduijn *et al.,* 2007).

This has important implications as treatment selection and prognostic reasoning at its very core concerns making predictions of future events despite inherent uncertainties. BN have the capacity to encompasses exploitation of knowledge of evolution of processes over time. Prognostic Bayesian models allow for incorporation of individual patient data, disease progression and impact of different treatment options on the predicted outcome variable, such as life expectancy (Lucas *et al.,* 2004). Therefore, unlike traditional prognostic models that provide predictions of a single outcome

variable, BN are theoretically better equipped to handle complexity, providing information on process variables (conditions that occur during the process) as well as outcome variables (endpoints of that process) (Verduijn *et al.,* 2007; Lucas *et al.,* 2004).

In summary, Bayesian methods underpin BN and allow prognosis to be seen as a dynamic notion through probability updating with new and emerging information (Verduijn *et al.,* 2007). This has several important benefits when considering the clinical application to support decision making. Firstly prognostic updating can capture the reality that as the healthcare process evolves so does a patient's predicted prognosis. In practice this means clinicians involved at the later stages of care can use the same model, adjusted for the events of the preceding care phases (e.g. complex surgical interventions) to make more timely and personalised predictions (Verduijn *et al.,* 2007). This further highlights an aspect of predictive medicine not captured in traditional prognostic models; prognostic scenario analysis. In real life events such as complications and hospital stay do not happen in isolation but rather as scenarios (Verduijn *et al.,* 2007). Algorithms exist within prognostic BN that can perform this type of probabilistic inference to predict a most likely scenario for patients or patient groups (Verduijn *et al.,* 2007). This advantage links beneficially to a further aspect faced by clinicians and patients; the 'what if scenario'. By identifying a specific event the prognostic BN can supply a risk profile of the most likely scenarios leading to the stated event (Verduijn *et al.,* 2007). Such information can be incorporated into decision making regarding treatment options. Similarly BN can be used to perform risk factor analysis as when an

unfavorable event occurs, such as post-operative complication, it is important to identify variables that may have predicted occurrence or nonoccurrence of said event and quantify this in terms of risk ratios (Verduijn *et al.,* 2007; Lucas *et al.,* 2004)

Finally, in addition to the clinical application of BN, an emerging application of BN is in molecular biology (Lucas *et al.,* 2004). As has been explained BN can be understood as representation of uncertain interactions amongst variables. Within bioinformatics BN are being used to explore interactions between genes based on experimentally obtained data in microarrays (Lucas *et al.,* 2004). It is hoped that BN analysis may reveal how the variables interact as a function of time (Lucas *et al.,* 2004). It is possible that through BN the future role of precision medicine within personalised realistic medicine could lie in amalgamating clinically observed patient data with genetic profiling to give patients and clinicians the most accurate predictions of patient outcome when deciding treatment approaches and resource allocation.

The first exciting steps in this path are starting to emerge with the recently published paper by Yamamoto *et al.* (2017) demonstrating that a mathematical model can successfully reproduced clinical outcomes using a predictive signature for lower propensity to metastatic disease based on the finding that these primary tumours contain a small fraction of *KRAS* and *CDKN2A, TP53, or SMAD4* genes. Although this model requires prospective validation it indicates a future direction of research whereby PDAC treatment can be personalised to the most effective therapeutic modality. The next

phase of research will be in integrating breakthroughs in genetic profiling into predictive models for surgical morbidity/ mortality and long-term survival outcomes.

To conclude, the patient with a favorable genetic profile making metastatic disease from their primary pancreatic cancer less likely, but with other pre-existing comorbidities will still want to know how likely they are to survive an operation, their risk of complications from all proposed treatments and their implications including quality adjusted survival predictions across competing treatment strategies such as neoadjuvant or upfront surgery pathways. This is the future of personalised predictive medicine supporting cost-effective healthcare. However, in practical terms this requires the integration of large complex databases.

Bayesian statistics has been offered here as a possible way forward. The potential for the application of this branch of mathematics is only beginning to come into fruition due to advances in the ability of computer software to handle such computational statistics. Bayesianism, in addition to other novel approaches to statistical modeling, have therefore been applied within the wider discipline of machine learning. As previously mentioned the period of time when statistical theory began to coalesce with machine learning was also the period of time when the field of complex systems was developing. This demonstrated the gradual realisation in some fields of the need to develop ways of engaging with complexity including a large number of variables that interact in non-linear, often unpredictable, ways. Yet, based on the existing body of predictive and prognostic

modeling, such an epiphany has yet to dawn within the application of decision support models applied to pancreatic cancer management. Furthermore, where more novel statistical approaches have been employed for prognostic modeling, BN (Smith & Mezhir, 2014) and ANN (Walczak & Velanovich, 2012), their performances did not rival that of models based on traditional regression techniques. However, in both cases the full potential of these techniques were not fully explored as the boundaries of the models failed to attempt to engage with the complexity of the system being modeled. The BN focused predominately on the bearing of lymph node involvement to prognosis (Smith & Mezhir, 2014). The ANN study also used a limited number of variables to predict survival at 7months post resection (Walczak & Velanovich, 2012).

If we consider, as the problem structuring thinkers do, that the definition of a problem, proposals for improvement and outcome are all dependent on the whole system (Ulrich, 2002) and, taking this idea further as Cilliers did in combining thinking about boundaries with concerns relating to complexity (Kruger *et al*., 2019), the current limitations of predictive modeling, whether utilising traditional or newer modeling techniques, reflect the limited knowledge and understanding of the system as a result of boundaries and failure to engage with complexity (Kruger *et al*., 2019).

Considering that methods of machine learning have been championed as having the ability to engage with a large number of variables that interact in a complex, non-linear way, several questions remain. Where, how, and to what degree of success has

machine learning been applied to decision making in the management of pancreatic cancer and do machine learning methods other than Bayesianism offer advantages?

2.3.4 Personalised Pancreatic Cancer Management: how Machine Learning is Supporting Decision Making

Several factors have aligned making decision making in the management of pancreatic cancer more complex. In addition to the pancreatic cancer management pathway issues already discussed, the ageing population and obesity epidemic means patients in general are amassing a greater amount of clinical data to be consider when making clinical decisions (Obermeyer & Lee, 2017). Treatment options are expanding with the emergence of neoadjuvant approach as an alternative to upfront surgery. While some are optimistic about the role of neoadjuvant therapy, others feel the current body of evidence is at best ambiguous with its role in the management of resectable pancreatic cancer being particularly controversial (Tempero *et al*., 2014; Asare *et al*., 2016; Lee *et al*., 2016; Xu *et al*., 2014; Andrulli *et al*., 2012; Versteijne *et al*., 2018). This is compounded by the current lack of RCTs comparing both upfront surgery and neoadjuvant treatment pathways (Versteijne *et al*., 2018). Furthermore with a research move towards precision medicine (gene targeted therapy) databases will expand to reflect our understanding of disease at genomic level, creating a further 'data explosion' (Tonelli & Shirts, 2017). Patients therefore represent a big data challenge not only in the amount of data amassed, but in being extremely complex data systems with multidimensional

problems and interacting parameters with the rules governing behaviours within layers of these systems often unclear or simply unknown (Abbod *et al*., 2014).

Personalised predictive modeling has gained precedence as a means of supporting clinical decision making (Velikova *et al*., 2014). However, as previously discussed existing predictive models, mainly based on non-liner regression techniques are limited in scope and volume regarding prognosis as an isolated event at a pre-determined time (Velikova *et al*., 2014; Verduijn *et al*., 2014). In isolation the factors outlined as contributing to the complexity of decision making may not be unique to pancreatic cancer. However, in the context of being one of the most challenging malignancies (Siegel *et al*., 2015; Ferlay *et al*., 2013), with comparatively lower resection rates compared to other gastrointestinal malignancies (Siegel *et al*., 2015; Ferlay *et al*., 2013; PCUK, 2017), pancreatic cancer is the ideal vehicle to critically examine how successful machine learning is in dealing with complexity and uncertainty to support clinical decision making.

Machine learning methods make predictions within complex systems against a background of competing risks and events (Abbod *et al*., 2014). Machine learning achieves this in one of three ways. Firstly supervised learning, where the computer utilises partial labeling of data (Hashimoto *et al*., 2018; Deo, 2015). Alternatively unsupervised learning allows the computer to make predictions or explain data by utilising structures detected within the data itself (Hashimoto *et al*., 2018; Deo, 2015). Thirdly reinforcement learning whereby, similar to operant conditioning (Skinner, 1938), the computer learns from its

mistakes and successes to achieve a task (Hashimoto *et al*., 2018; Sutton & Barto, 1998).

Commonly employed methods of machine learning include, but are not limited to: Bayesian networks (BN), artificial neural networks (ANN) and Fuzzy-logic (FL) modeling (Abbod *et al*., 2014). The definitions, strengths and limitations of these most commonly employed methods of machine learning are further discussed in appendix J along with a critical analysis of machine learning for decision analysis, prognostic and predictive purposes to support clinical decision making in the management of potentially resectable pancreatic cancer, based on the CHARMS checklist (Moons *et al*., 2014).

The review presented in appendix J found that machine learning, although in its infancy, holds great potential in its application to decision making under complexity (Abbod *et al*., 2014; Bartosch-Härlid *et al*., 2008). However the application of machine learning to predictive modeling pertaining to the management of pancreatic cancer is currently limited in number therefore no conclusion can yet be drawn as to superiority of either machine learning or traditional modeling approaches. Only one study directly compared machine learning methods with traditional approach to modeling (Hayward *et al*., 2010). The accuracy of machine learning predictions, particularly Bayesian modeling, were found to be superior and predictions form log regression approach were improved when combined with machine learning techniques (Hayward *et al*., 2010). However, it is important to note that of the existing predictive studies using

machine learning, limitations in methodological approach were identified using the CHARMS checklist (Moons *et al*., 2014). These issues are similar to issues highlighted in traditional approaches to predictive modeling and include: use of single centre database limiting generalisability, sample size, lack of blinding, lack of transparency in candidate predictor selection, and lack of external validation (Moons *et al*., 2014; Moons *et al*., 2009; Bouwneester *et al*., 2012; Altman, 2001; Altman *et al*., 2009).

Whilst much optimism surrounds the growing use for artificial intelligence (AI) in healthcare delivery, machine learning also carries limitations that must be addressed in future research. Machine learning usually requires large amounts of data (Marcus, 2018), which in the case of potentially resectable pancreatic cancer can be difficult to obtain as the majority of patients present with advanced, unresectable disease (Siegel *et al*., 2015; Ferlay *et al*., 2013; PCUK, 2017). Whilst the creation of national shared databases may be one solution to increase the volume of data, this is not without issue including dimensionality, missing data and control of bias (Lee & Yoon, 2017; Zhang et al., 2017) with minority groups often under represented in such databases (Zhang et al., 2017). Furthermore simply increasing volume of data is not the solution as machine learning is not yet at a stage where it can distinguish correlation and causation (Marcus, 2018). Future research should focus on better integration of machine learning with expert knowledge to overcome this challenge (Marcus, 2018).

This review (Appendix J) found little evidence of machine learning being actively integrated into clinical practice. Whilst this is mainly due to such techniques being in their infancy, it must also be acknowledged that some machine learning techniques are not yet sufficiently transparent which breeds distrust and resistance to their clinical application (Marcus, 2018). Machine learning requires high levels of technical skill and can be difficult to engineer with experts from medicine, computing and data sciences often speaking in different technical language and coming to problems from different perspectives which can inhibit shared understanding and limit achievement of its full potential (Marcus, 2018). The beginning of a possible solution could therefore lie with clinicians expanding their view of the multidisciplinary team to include professionals from computing and data science backgrounds with algorithms developed in conjunction with clinicians and viewed as aids, not replacement, to traditional clinical decision making (Obermeyer & Lee, 2017).

Despite these challenges the study by Hayward *et al.*(2010) does however corroborate other studies where application of machine learning methods to: breast, prostate and bladder cancers have demonstrated superiority in terms of accuracy of predictions over traditional logistic regression (Seker, 2003; Catto *et al*., 2003; Abbod *et al*., 2006; Catto *et al*., 2009). Artificial Neural Networks (ANN) have also been found to perform as well as or better than traditional log regression models and also improve the diagnosis and management of pancreatitis and the diagnosis of pancreatic cancer (Bartosch-Härlid *et al*., 2008). Machine learning methods have also been shown to out perform log regression in: providing individualised prediction

of the need for neonatal resuscitation (Reis *et al*., 2004), predicting early mortality risk in coronary artery bypass graft surgery (Ghavidel *et al*., 2014) and predicting severely depressed left ventricular ejection fraction following admission to intensive care unit (Pereira *et al.*, 2015). However, the studies reporting this advantage are prone to bias. As discussed in appendix J such models share the limitations of more traditional predictive models that where highlighted in appendix I.

Conclusion

To conclude clinical decision making is going to become increasingly complex and orientated twards uncovering causal structures as our understanding of disease and treatment response at genomic level grows, resulting in a further 'data explosion'(Obermeyer & Lee, 2017; Tonelli & Shirts, 2017; Abbod et al., 2014). Utilising this expanse of data to facilitate decision making in a meaningful way for individual patients is beyond the capabilities of the human mind working in isolation (Obermeyer & Lee, 2017; Abbod et al., 2014; Bartosch-Härlid *et al*., 2008). It is in this context that machine learning holds the greatest potential by being able to handle large amounts of data and integrate large, complex and varied databases (Bartosch-Härlid *et al*., 2008). However machine learning also carries limitations and, whilst initial studies are promising, its application has yet to be widely tested (Marcus, 2018). The future direction of research therefore relies on expanding our view of the multidisciplinary team to include professionals from computing and data science backgrounds with algorithms developed in conjunction with

clinicians and viewed as aids, not replacement, to traditional clinical decision making (Obermeyer & Lee, 2017).

2.4 Chapter Summary and Conclusion

The aim of this research is to facilitate the fruition of personalised realistic medicine in the delivery of pancreatic cancer services through statistical modelling that will facilitate better shared decision making with patients and the entire multi-disciplinary team to optimise individual patient outcomes as determined by the individual patient. However, this chapter has demonstrated that the existing body of research pertaining to the management of potentially resectable pancreatic cancer is highly heterogeneous, limited by issues of small sample size and methodological quality potentiating bias, and therefore is permeated by ambiguity, controversy and uncertainty.

Pancreatic cancer is a challenging malignancy associated with poor survival outcomes. In the United Kingdom only 9.8% of cases are resectable at presentation with international estimates ranging from 10-20% (CRUK, 2019). Current guidelines for resectable pancreatic cancer recommend upfront surgical resection followed by adjuvant therapy in the form of mFOLFIRINOX as the first line treatment sequence (Khorana *et al*., 2019). However, up to 50% of patients with resected disease fail to receive adjuvant therapy due a combination of factors including early disease recurrence, post operative complications and decline in physiological function related to pre-

existing comorbidities (Winter *et al*., 2012; Evans *et al*., 2018). Five-year survival rates for resected cases stands at between 7% and 25% (CRUK, 2019). This has resulted in a renewed research interest in neoadjuvant therapy. Postulated benefits of this approach include elimination of micrometastases, conversion to resectability in borderline and locally advanced stages of the disease, increased R0 resection rates, increased likelihood of delivery of multimodal treatment, and allowing time for more aggressive tumours to declare themselves by progressing despite neoadjuvant therapy hence filtering such cases away from costly yet futile surgery with its associated risks of morbidity and mortality impacting on quality-of-life (Evans *et al*., 2018; Asare *et al*., 2016; Lee *et al*., 2016; Abbott *et al*., 2013).

Whilst the role of neoadjuvant therapy has been widely accepted for cases that are borderline resectable or locally advanced at the time of presentation due to the potential for conversion to resectability, particularly R0 resection, its role in the management of resectable pancreatic cancer is controversial. Critics highlight the dangers of loosing the window of resectability and caution against drawing overly optimistic conclusions from small, non-randomised, underpowered studies that display a high degree of heterogeneity (Asare *et al.*, 2016; Lee *et al*., 2016). Currently there is a lack of RCTs comparing upfront surgery and neoadjuvant treatment pathways for resectable pancreatic cancer with many comparison studies including borderline or locally advanced cases in the neoadjuvant arm hence failing to offer a true like-for-like comparison (Versteijne et al., 2018).

The existing body of research on neoadjuvant therapy for resectable pancreatic cancer leaves much room for debate. Preliminary findings from Prep-02/JSAP-05 trial, the first RCT comparing upfront surgery and neoadjuvant therapy in the form of gemcitabine and S1 for resectable pancreatic cancer, has reported improved overall survival outcomes with neoadjuvant therapy (Unno *et al*., 2019). However, another RCT comparing mFOLFIRINOX with gemcitabine in the adjuvant setting within the upfront surgery pathway has reported improved survival outcomes with mFOLFIRINOX that rivals the survival outcomes reported in the neoadjuvant arms of the Prep-02/JSAP-05 trial (Conroy *et al*., 2018). This highlights key challenges. Firstly the superior treatment pathway for resectable pancreatic caner has not been conclusively established. Secondly superior treatment regime combinations within competing pathways have not been conclusively established.

These issues exist within the wider political context of a drive towards the delivery of personalised realistic medicine through more personalised treatment selection strategies that will ensure more cost-effective resource utilisation. This has resulted in the current research focus within pancreatic cancer research being driven in two key areas where trials are underway: 1) the drive for more large multi-centre RCTs comparing neoadjuvant and upfront surgery and 2) precision medicine with the focus on biomarker driven early diagnosis and treatment sequencing and gene targeted therapies.

Precision medicine permeates much of the current medical literature and has been championed by some as hailing a 'brave new world' of future medicine bringing the end of uncertainty in clinical decision-making. This exciting vision of a brave new era of medicine is based on a few impressive studies demonstrating the success of targeted therapies predicted by genetic biomarkers. However, whilst such breakthroughs are both impressive and exciting, in reality the number of patients benefitting from precision medicine currently remains small (MacConaill *et al*., 2015). Whilst such work is highly valuable and important, the ongoing perpetuation of a Newtonian world-view can only ever have a limited impact on moving research forward. To illustrate, a deeper understanding of pancreatic cancer at a molecular level has not influenced clinical decision making to the extent that it has done with other caners (Collisson *et al.*, 2019). Instead this has resulted in pancreatic cancer beginning to be understood as a highly heterogeneous and complex disease at molecular level (Collisson *et al.*, 2019). It follows that breakthroughs in such areas, rather than solving uncertainty and complexity will simply reveal the scale of the challenge particularly when the impact of additional, often ambiguous, clinical information is factored into the decision making process across the trajectory of the patient journey as has been discussed within this chapter that has highlighted the degree of uncertainty pertaining to key aspects of the treatment pathway. Furthermore the enormity of the challenge of delivering precision medicine has become the proverbial elephant in the room. These challenges include:

- Improving infrastructure for data integration: previously unstructured, large scale, detailed datasets must be integrated

into knowledge networks (National Research Council, 2011). This poses questions at policy, financial and technical level to regulate data access and security (Dzau & Ginsberg, 2016).

- Evidence of benefit: a barrier to adoption of precision medicine is the limited evidence that it has improved outcomes at population level or carries benefits in cost-effectiveness. The latter is currently being addressed by collecting data alongside ongoing clinical trials to assess cost-effectiveness (Dzau & Ginsberg, 2016; Joyner & Paneth, 2015). Other options include observational research to identify modifiers of effectiveness, dedicated precision medicine RCTs and disciplined subgroup analysis and interaction testing within standard RCTs of intervention effectiveness (Pletcher & McCulloch, 2017). Here lessons could be learned from robust methods of controlling type I errors and culture of replication developed from exploration of the genome to protect against propagation of spurious findings (Pletcher & McCulloch, 2017).
- Evidence generation: the traditional hierarchy of population based evidence based medicine must be challenged if precision medicine is to address the issue of variance of unknown significance at individual level (Tonelli & Shirts, 2017).
- Incorporating genomic and patient data into clinical care: this includes education, training, decision support and development of techniques and technology to support integration of genomic and patient data into clinical practice otherwise precision medicine will simply be genomic medicine (Tonelli & Shirts, 2017; Dzau & Ginsberg, 2016).

Rather than reducing the complexity of decision making, precision medicine could therefore actually increase the degree complexity and uncertainty inherent in the decision-making process. To put it more simplistically: we are more than our genomes. Furthermore outcomes in cancer care are highly complex and individualised. It would be a sad irony if precision medicine, hailed as the dawning of a 'brave new world' actually became a retrograde step dragging medicine back to a reductionist view of health and disease.

The mapping of the human genome was, at one time, postulated to bring about a 'silver bullet' cure for cancer. What it actually resulted in was a data explosion that eventually has resulted in some significant and impressive breakthroughs but no silver bullet cure. Lessons can be drawn from this. Precision medicine, rather than 'curing' the uncertainty and complexity inherent in clinical decision-making will bring about a further data explosion as our understanding of disease and its treatment deepens. Therefore the actual delivery of precision medicine will entail integrating genomic data with behavioral, clinical, pathological, physiological and epidemiological data. Ultimately clinicians will be expected to make decisions in the face of increased complexity. Practically this means being able to integrate information from large, complex databases drawn from different disciples and sources and apply them to an individual patient who themselves is dynamic with an ever changing clinical picture along the trajectory of their care pathway. Fundamentally the complexity of the challenge of integrating multiple complex databases to achieve personalised predictive medicine is simply too vast for the human mind to handle

unsupported (Obermeyer & Lee, 2017). Unless techniques are developed to use the expanding amount of data to more effectively support decision making, clinicians will simply drown in a data tsunami. We have already seen some early warning signs where electronic data, if handled badly, can overwhelm clinicians rather than assist, leading to the "4000 keystrokes" phenomenon contributing to burnout (Hill *et al*., 2013).

The enormity of this task is compounded when we consider how patients and their expectations are also changing. Firstly an ageing population and obesity epidemic means patients already represent a big data challenge, seeing more specialists and therefore amassing copious amounts of information in their electronic health records (Obermeyer & Lee, 2017). Secondly medicine does not operate in a vacuum. In an increasingly high-tech world personalised predictions from targeted advertisements to credit ratings are commonplace. Therefore it is no surprise that there is a growing expectation for personalised predictive medicine at patient level as well as organizational and political level.

Personalised predictive medicine is captured within the broader term personalised realistic medicine which at its core seeks to deliver the right diagnosis and treatment to the right patient at the right time with the right outcomes determined in collaboration with the individual patient (Alexandrou *et al.,* 2011; The Scottish Government, 2016; The Scottish Government, 2017). Acknowledging the gravitas of the challenge the CMO also called for creative and collaborative working to make this a reality (The Scottish

Government, 2016; The Scottish Government, 2017). Partially in response to this there has been an increasing crossover between healthcare research and operational research as the latter has traditionally focused on the use of mathematical techniques and modeling to support decision making and achieve optimisation of outcomes.

Existing predictive models for pancreatic cancer prognosis, surgical outcomes and cost-effectiveness analysis are limited with most only being descriptive, rather than predictive, of the likelihood of adverse events or survival outcomes (Lewis & Volmer, 2012). Like the majority of predictive models in medicine they focus on "risk" at a population level and then attempt to apply this at individual patient level (Grossi, 2006). Considering the ambiguity permeating the exiting body of studies into the treatment of pancreatic cancer, and the bamboozlement this approach potentially creates in communication between patient and healthcare professional at individual patient level becomes apparent (Grossi, 2015). If the full potential of predictive models are to be realised within personalised realistic medicine, they must integrate fully into clinical practice. To do this they need to provide individualised predictions beyond length of survival or risk prediction to include fundamentals such as quality of survival time, length of hospital stay, resource utilisation and associated costs and predicted benefits of competing treatment options available.

The problem with existing modeling techniques are that they regard prognosis as an isolated event at a pre-determined time, applying attribute selection prior to inducing the model and setting fixed roles of input and output variables to attributes (Verduijn *et al.,* 2007). They neglect the uncertain and dynamic nature of care processes where outcomes today predict those of tomorrow hence expected patient outcomes evolve as more information becomes available (Verduijn *et al.,* 2007). Put simply, traditional decision support models integrate data and knowledge but do not provide reasoning (Muthurmama & Sankaran, 2014).

To achieve personalised predictive medicine statistical models therefore must improve both knowledge representation and reasoning facility, with ontologies employed acting as stepping-stones to achieving this, and ultimately delivering personalised realistic medicine (Muthurmama & Sankaran, 2014). Embracing the call for innovation and creativity the new era of operational research applied to medicine must encompass novel approaches in the world of mathematics, statistics and computer science. Emerging statistical modeling techniques applied within other disciplines, such as engineering, ecology, astrophysics, biomedical sciences and business have made phenomenal advances, moving beyond data explosions within these fields through the application of soft computing techniques such as: Bayesian networks, fuzzy logic and artificial neural networks (Bhatia *et al.,* 2014). Recently several studies have emerged demonstrating that such techniques have improved accuracy of prediction compared to traditional predictive models within medicine (Seker *et al.,* 2003; Catto *et al.,* 2006; Abbod *et al.,*

2006; Catto *et al.,* 2003; Catto *et al.,* 2009). However, the review of cost-effectiveness analysis and prediction modeling studies revealed that studies using newer methods of statistical modeling techniques did not always demonstrate a significant performance advantage in their accuracy of predictions and significant flaws still prevailed.

A key point being made in this chapter is that if research is to advance, and the narrative surrounding the treatment of potentially resectable pancreatic cancer is to evolve towards more personalised medicine, the lessons learned from the existing body of literature must go beyond those of statistical modeling methodology alone to consider the theory driving current research and how this relates to its current limitations. Existing studies utilising statistical modeling techniques still seek to establish a superior treatment pathway at population level rather than engage with the complex adaptive nature of the system being modeled to reveal new insights that could drive future research towards achieving personalised realistic medicine. It follows that where the systems being studied and modeled have been so reduced and simplified, the potential of newer statistical modeling techniques have not yet been fully explored.

There is a growing move within healthcare research to view healthcare systems as complex adaptive systems whereby a collection of individual agents have the freedom to act in ways that are not always predictable, and whose actions are interconnected (Plsek & Greenhalgh, 2001). However, throughout the existing body of research pertaining to the management of potentially resectable pancreatic cancer the definition of the research problem, proposals

for improvement and outcome are not recognised as being dependent on the whole system (Ulrich, 2002). Doing so would place greater emphasis on how systems boundaries are justified and the implications this has for what modeling a system defined in such a way will, and importantly will not, reveal. Recognising this issue also reframes how the limitations of the current body of research are viewed as it highlights the implications of boundary setting across the previously discussed cost-effectiveness and prediction modeling studies which relates to their limitations and manifest as, for example: exclusion of important alternative treatment strategies, the exclusion of certain costs including indirect costs forfeited by the patient, exclusion of consideration of all relevant potential implications of a treatment strategy including treatment failure and side effects, a lack of quality adjusting survival time or collecting quality-of-life data to more accurately do so. Hence the limitations of the existing body of research is not merely to be seen as a series of methodological issues to be corrected, but rather as the system that is the delivery of healthcare being defined in simplistic and reductionist terms which defines how system boundaries are set which in turn determines how outcomes are measured and assessed. It follows that both a limited knowledge of systems as a result of boundaries and a failure to engage with complexity exist and therefore require a critical and ethical imperative in the study and understanding of such systems in order to move research forward (Kruger *et al*., 2019). It is not without coincidence that the field of complex systems developed at a time when statistical theory began to coalesce with methods encompassed within machine learning to reliably infer models with large numbers of variables that interact in

complex, non-linear ways. However, the potential for advancement will remain untapped unless the philosophy driving future research also evolves.

Mirroring the misplaced optimism surrounding precision medicine as a 'cure' to complexity and uncertainty in decision making, such breakthroughs have resulted in some seeing AI as the 'solution' to the challenges of complex decision making. Such developments actually represent an expanse in the capabilities of computational statistics rather than the man-made creation of intelligence. Therefore while many have espoused AI and machine learning as the solution to delivering personlised medicine with the associated cost-effectiveness implications, they are in danger of creating hollow sound-bites by failing to appreciate what lies at both the core of achieving its potential impact and simultaneously also at the core of the barriers to achieving this impact.

At their core these methods make predictions within complex systems against a background of competing risks and events (Abbod *et al.*, 2014). However, a 'black box' approach to machine learning through algorithms alone has led to suspicion regarding its clinical application with some justification. Algorithm based machine learning from databases has failed to consider the impact of clinical judgment on decision making. One notorious example is where such an approach failed to account for the successful clinical protocol of admitting patients with asthma who presented with pneumonia, which resulted in fewer complications (Caruana *et al.*, 2015). As the data consequently did not show an increased rate of complications in

this patient group, the machine learning model erroneously advised no admission for patients with asthma who presented with pneumonia.

The successful application of machine learning in other fields depended on seeking experts in computer science to develop cutting-edge algorithms required for complex problems (Obermeyer & Lee, 2017). However, it was the experts within these fields who set the research agenda and ensured its relevant application to their practice. Rather than see algorithms as a replacement to human decision making processes, the algorithms were viewed as thinking partners, supporting decision making in the face of complexity (Obermeyer & Lee, 2017).

The point being made here is that the simple application of advances in computational statistics to the research problem of how to deliver realistic medicine through personalised predictive medicine is unlikely to provide a solution. This is partly because any such advances themselves would have to be accepted within a wider complex healthcare system (Greenhalgh *et al*., 2017). Despite the expanse of technological innovation now being viewed as a significant contributor to health and wealth, the integration of such technological advances into the healthcare systems and daily practice is plagued by non-adoption and abandonment particularly where change at organisational and the wider systems level is required (Garber *et al.,* 2014; van Limburg *et al.,* 2011; Grin *et al.,* 2010; Greenhalgh *et al*., 2017). Even where initiatives, such as telehealth, were backed by policy-level rhetoric and supported by small scale

proof-of-concepts studies, non-adoption and abandonment by intended users is common place and telehealth services are rarely mainstreamed or maintained (Greenhalgh *et al*., 2017; Greenhalgh *et al.,* 2017; Standing *et al.,* 2016; Bentley *et al*., 2014; Clark & McGee-Lennon, 2011; Wade *et al.,* 2014).

In conclusion, this chapter presents a review of the current body of literature and has revealed that at the core of achieving the aim of evolving research towards personalised realistic medicine in the delivery of pancreatic cancer services through statistical modelling, lies the need to develop ways to engage with the complexity, handle uncertainty and the emergent when examining the complex system of delivering pancreatic cancer care including areas of debate, ambiguity and disagreement (Law & Mol, 2002; Fraser & Greenhalgh, 2001; Star, 2002; Greenhalgh & Papoutsi, 2018). It follows that because the system of delivering pancreatic cancer care and its outcomes are dynamic, the traditional scientific quest for certainty, predictability and linear causality through a focus on RCTs and precision medicine will only answer a fraction of the unanswered questions as the effect of context is controlled for within the artificial setting of such trials (Cohn et al., 2013; Braithwaite et al., 2017; Marchal et al., 2013; Greenhalgh & Papoutsi, 2018). RCTs with their strict inclusion criteria and control of context do not reflect the complexities of a real-life patient case mix and therefore cannot alone provide solutions to the challenge of optimising outcomes on an individual patient level. Therefore what is needed is research that augments such studies by exploring how to deal with uncertainty, unpredictability and general causality through designs and methods

that foreground dynamic interactions and emergence to understand how systems come together as a whole from different perspectives (Cohn et al., 2013; Greenhalgh & Papoutsi, 2018; Flyvbjerg, 2006). This challenge demands more than simply employing different statistical modeling techniques but rather a novel Weltanschauung (Sadegh-Zadeh, 2001) to bring about the necessary scientific change to tackle this problem through what Khun initially termed a 'paradigm-shift' (Kuhn, 1962) and later revised as a shift in the 'disciplinary-matrix' (Kuhn, 1977). To understand the gravitas of the revolution in scientific thinking required, the following Methods chapter will critically examine the prevailing dominant philosophy driving medical and operational research before an alternative paradigm, ontology, epistemology and theoretical framework for this research is offered.

Chapter 3

Methods

3.1 Research Philosophy

Introduction

This chapter opens with a critical analysis of the prevailing dominant philosophy driving medical and operational research before an alternative paradigm, ontology, epistemology and theoretical framework for this research is offered. From this basis the strengths and weakness of decision models in handling uncertainty and complexity will be examined before Bayesian methods are discussed as a vehicle for taking statistical modeling and personalised realistic medicine to a new level of insight through complexity theory.

3.1.1 The Current Philosophical Direction of Research: the case for a new roadmap

Positivism has reigned as the dominant philosophy across much of scientific research including operational research and medicine. Classic reasoning, for over two millennium, has been dominated by the Aristotelian disciplinary matrix, which gravitates around 'truth' and 'falsehood' hence arguably being viewed as the progenitor of Tarski semantics of classical two-valued logic and Cantor's two-valued set theory (Sadegh-Zadeh, 2001). Aristotelian ontology

postulates that classes must have defined sharp boundaries and rejects any intermediate between such states, or a 'doctrine of crisp existence' (Sadegh-Zadeh, 2001; Grossi, 2015). Consequently much of the existing and emerging research in the fields of operational research and medicine take a mechanistic world view inspired by a Newtonian framework that postulates an understanding of the universe through a process of reductionism of systems and an analysis of their parts to understand the whole with the methodologies and practices employed in this research further propagating such assumptions (Kruger *et al.*, 2019).

Comte viewed the original aim of positivism as providing an unambiguous and accurate knowledge of the world through application of methods from natural science to social science (Bridges, 2009; Bisman, 2010). The Vienna Circle later applied mathematical exactitudes to philosophy introducing 'logical positivism', embracing empiricism and rejecting all else (Sahotra, 1996; Houghton, 2011). Epistemologically an objective view of reality is held. Through quantitative methods of statistical modeling, this view is juxtaposed with the ontological view of reality comprising determined, observable, measurable events that interact in an observable, measurable manner (Houghton, 2011; Smith, 2005; Bisman, 2010). Statistical models are populated with data from patient databases and clinical trials selected in a hierarchical order whereby RCTs reign as "gold-standard" evidence. Thus natural and social sciences amalgamate in a shared logic of enquiry to explain and predict treatment outcomes and cost-effectiveness based on

factual, value-free judgment with reliability evidenced in replicability (Houghton, 2011; Bisman, 2010).

However in the previous chapter a deeper analysis of existing statistical models pertaining to the management of potentially resectable pancreatic cancer revealed limitations mirroring the philosophical criticisms of positivism. Comte warned against the danger of confusing signs for ideas when blindly introducing mathematics to investigation of social science (Houghton, 2011). Later quantum theory usurped the perceived infallibility of positivism by questioning both the human ability to determine true accuracy of information and maintain complete objectivity (Houghton, 2011).

Complete objectivity within statistical modeling is highly questionable (Mingers, 2004; Mingers, 2005; Zachariadis *et al.,* 2010). Results cannot be described or classified without an element of interpretation (Mingers, 2004; Mingers, 2005; Zachariadis *et al.,* 2010). Under positivism philosophy and the associated Aristotelian disciplinary matrix of 'truth' and 'falsehood' it is assumed that factors not included within statistical models (because they are unknown or difficult to measure) have random or insignificant effects on outcome (Mingers, 2004). Through positivism's 'naïve realism', results incompatible with theory are dismissed as an anomaly (Mingers, 2004; Mingers, 2005; Zachariadis *et al.,* 2010). Furthermore Kuhn postulates that use of a paradigm can limit the questions the researcher asks and their interpretation of results therefore the researcher is not objective (Mingers, 2004; Mingers, 2005;

Zachariadis *et al*., 2010; Steele, 2005). This is evidenced within existing statistical models through: excluded costs including indirect costs, lack of quality-of-life data, exclusion of surveillance data and associated costs, and utilisation of narrow models. Also in selecting between competing models for the same data, despite creation of elaborate methods, statistical models are often selected on subjective grounds including 'best-fit' with the researcher's view (Mingers, 2004; Mingers, 2005). Hendry *et al*. (1990) posit that by maintaining a solely positivism stance, model selection can become adhoc and atheoretical (Mingers, 2005; Zachariadis *et al.,* 2010). Acknowledging these limitations, contemporary positivism maintains emphasis on empiricism but deals in partial objectivity and probability rather than unquestionable facts (Bisman, 2010; Houghton, 2011; Smith, 2005). However this does not address the issue of distinguishing natural and social sciences, which has further implications for statistical modeling.

The most vehement criticism of positivism comes from interpretive view, seeing the Humean notion underlying empiricism of causality as a constant conjunction of events as impoverished (Mingers, 2004; Mingers, 2005; Zachariadis *et al.,* 2010). Statistical modeling is criticised on the basis that material and social worlds are different, and human social construction cannot be captured or understood within statistical models (Mingers, 2005). This view is contested by the fact that computational and mathematical models are necessary for the development and progress of operational research and can and will continue to produce important results in many areas (Kruger *et al.*, 2019). However the interpretivism view is important

as it highlights two key issues. Firstly social structures are a product of, and shape, activities of society, but are not independent of them (Houghton, 2011; Marsh & Stoker, 2002). To illustrate, statistical models can be populated with databases from large specialist centres but in private health-care systems patients may not present to such institutions due to socio-economic reasons (Abbott *et al*., 2013). Secondly, social structures are not independent of the agent's view but are shaped by their actions and may change (Houghton, 2011; Marsh & Stoker, 2002). Statistical models must be able to adapt to unforeseen circumstances such as unanticipated complications of treatment, or changes in, for example funding and costs, political prioritisation of health-care resources and society's willingness-to-pay.

Russel Ackoff, in his 1979 paper "The future of operational research is past" addressed some of these issues by highlighting problems with the pursuit of objectivity and instead argued in favour of expansionism over reductionism through systems thinking (Ackoff, 1979a). Systems thinking is central to the methodological pluralism view in operational research and marked a move away from the positivist stance within operational research towards what Midgley termed the 'second wave systems thinking' (Midgley, 2000). Rather than dismiss existing mathematical and computational models Ackoff sought to develop, enrich and complement these models and their underlying theories through methodological pluralism (Ackoff, 1979a; Kruger *et al.*, 2019). For him the main critique of existing methods was that these deterministic models assumed the problem

context of a closed system, which raised six deficiencies in the prevailing epistemology:

1. the need for decision making systems to learn to adapt
2. the need for decision making systems to consider quality-of-life values
3. model abstraction of systems of problems as problems cannot be treated effectively by deconstructing them analytically into separate problems
4. the need for a synthesising planning paradigm rather than a problem-solving paradigm
5. interdisciplinary interaction is required to deal with complex issues
6. the pursuit of complete objectivity when in reality the view of all those affected by the outcome of a decision making process must be considered (Ackoff, 1979a; Kruger *et al.*, 2019).

The common denominator in all six issues is a human characteristic that requires an integrated, holistic approach to address them (Kruger *et al.*, 2019). Ackoff was proposing a move away from deterministic statistical models towards a systems thinking approach whereby "purposeful systems that contain purposeful parts with different roles or functions and that are themselves parts of larger purposeful systems" (Ackoff, 1979a, p.96) are created and can serve its own purpose (self-control), the purpose of its parts (humanisation) and the purpose of the larger system of which they are a part (environmentalisation) (Ackoff, 1979a; Kruger *et al.*, 2019).

In summary, existing research conceived and critiqued within the philosophy of positivism, although valuable, are not infallible. The argument being made is that by aligning operational research epistemologies with the acknowledgement of the complexity inherent in the real-world, new methods of modeling decision making can be developed (Kruger *et al.*, 2019).

3.2 A New Philosophical Direction

A refocusing on alternatives to positivism has aligned with a greater use and acceptance of Bayesianism, which is based on probability theory, rather than solely relying on classical Frequentists statistics. In the post-positivism era alternative paradigms have emerged. One of the most prominent, critical realism, has been closely aligned with the shift towards Bayesian statistics.

Critical realism has been championed as a half-way-house between empiricism and positivism on one hand and anti-naturalism and interpretivism on the other (Mingers, 2004; Mingers, 2005; Zachariadis *et al*., 2010; Steele, 2005). This is a disservice. Critical realism introduces a more sophisticated paradigm simultaneously addressing the concerns of natural science (through technological characteristics) and social science (by applying human contexts) (Mingers, 2004; Mingers, 2005; Zachariadis *et al*., 2010; Steele, 2005; Smith, 2005). Critical realism proports better understanding of causal forces, underpinned by deep social structures that are not always identifiable by material properties or outward behaviours, through retroduction: analogy, metaphor, intuition and rhetoric

(Steele, 2005; Bhaskar, 1975; Bhaskar, 1979). Tendency to extreme apriorism therefore traditionally led critical realism to dismiss closed statistical models as merely observed "event regularities", unable to predict how isolated variables behave when exposed to 'real-world' exogenous factors (Steele, 2005). Russell (1929) however accounted for exogenous influence whereby causal sequence arises, but with probability of expected outcome less than $_{I}$. Yet within open models, whereby "no constant conjunction of events prevail", critical realism acknowledges, but dismisses causality of, sequential event regularities (Steele, 2005; Bhaskar, 1975; Bhaskar, 1979). Some conclude that critical realism therefore dismisses the value of statistical analysis completely, replacing mathematical formulae and statistical inference with retroduction (Steele, 2005). Conversely retroduction can be applied to any science scrutinising complex phenomenon (medicine describes the heart as a pump after all) hence both are not mutually exclusive (Steele, 2005).

Ontologically critical realism posits that the interplay of causal powers or tendencies of domains of 'the real' (structures, mechanisms, events and experiences) leads to particular events, 'the actual' (Mingers, 2005). These domains may be physical, social or conceptual (Mingers, 2005, Zachariadis *et al*., 2010). Events may be observable or experienced by people and therefore become empirical (Mingers, 2005). Epistemologically in recognising that all knowledge, whilst provisional, is historically and culturally relative, critical realism also accepts both epistemic reality (observer-independent access as a fallacy) and judgmental relativity (rational grounds for theory preference) (Mingers, 2005; Zachariadis *et al.,* 2010). Both

quantitative and qualitative methods are therefore acceptable within the essence of science: explanation, understanding and interpretation (Mingers, 2005; Zachariadis *et al.,* 2010).

Critical realism suggests that statistical models can be developed from a plethora of resources: experiments, theoretical work, expert opinion (Mingers, 2005). However, as with positivism, the challenge of inferring unknown mechanisms from limited observations and experiences remains (Mingers, 2005; Zachariadis *et al.,* 2010). The argument being made is that future research should move from solely quantifiable data and Humean causality to incorporating complex, multi-dimensional, underlying mechanisms within the empirical domain (Mingers, 2005). In practice this means employing quantitative statistical methods that concern themselves with discovering causal mechanisms (Mingers, 2005).

3.2.1 Critical Realism and Bayesian Models

Whilst accepting that models as a representation of reality are never truly exact; to quote Box:
"all models are wrong, but some are useful" (Box, 1979),
a continued adherence to classical Frequentists mathematics, underpinned by Aristotelian classical reasoning, in a complex environment of imprecision and uncertainty has resulted in decision support models that fall sadly short, leading to the longstanding belief that decision making is part of the 'art of medicine' as opposed to a science (Sadegh-Zadeh, 2001). This soporific view is why, despite advances in biomedicine and technology, clinical judgment has

largely remained archaic in the face of uncertainty (Sadegh-Zadeh, 2001; Sadegh-Zadeh, 1981; Sadegh-Zadeh, 1994; Sadegh-Zadeh, 1998; Sadegh-Zadeh, 1999; Sadegh-Zadeh, 2000). This also permeates research into cost-effectiveness analysis in healthcare. Despite recommendation that such reports include both a payer's perspective case report and a societal case report (Sanders *et al.,* 2016) there is a distinct lack in the current literature of research that includes attempts to analyse indirect costs to patients, their carers and wider society.

Bayesian statistics traditionally takes an inductive approach, learning about the general from particulars through inverse probability, starting with prior distributions, getting data and moving to posterior distribution (Gelman & Shalizi, 2013; Bernardo & Smith, 1994; Earman, 1992; Savage, 1954). Frequentists believe probability must reflect repetitive, objectively measured occurrences and the central goal is computing the posterior probabilities of hypothesis (Gelman & Shalizi, 2013; Bernardo & Smith, 1994; Earman, 1992; Savage, 1954). Through Bayes' Theorem probabilities are updated as new data emerges. Heavily steeped in the positivism philosophy of 'natural science' it holds that anything not contained in the posterior distribution is irrelevant (Gelman & Shalizi, 2013; Bernardo & Smith, 1994; Earman, 1992; Savage, 1954).

However, there is a move towards viewing Bayesian models in a deductive light with greater acceptance of subjective probabilities (Gelman & Shalizi, 2013). Bayesian models are characterised by subjective and objective knowledge, modeling information from a

variety of sources enabling changes in held beliefs on causal structures in light of occurrence, or absence of, events and emergence of new data (Gelman & Shalizi, 2013). This has attracted the attention of critical realists, seeing Bayesian modeling and its wide-ranging applications as a means of operationalising critical realism's retroductive methods (Mingers, 2005). But is critical realism masking the truth of contemporary Bayesians' argument?

Bayesians accepting subjective probabilities and postulating a deductive approach to modeling, do so from a positivism stance, using Popper's ideas of falsification to argue that Bayesian modeling is better understood from a hypothetico-deductive perspective (Gelman & Shalizim, 2013). Gelman & Shalizi (2013), although cited by those championing critical realism (Mingers, 2005), are actually following positivism philosophy of traditional statisticians emphasizing the importance of model checking and frequency evaluation to guide Bayesian inference and obtain statistical methods with good frequency properties (Gelman & Shalizi, 2013; Rubin, 1984; Wasserman, 2006). Accepting that all scientific statements must remain eternally tentative (Popper, 1959), Gelaman & Shalizi (2013) argue that Bayesian model checking must go beyond inductivist view of comparing posterior odds to support model selection. Instead models should be compared to data and, if falsified, rather than being rejected, aim to understand cause of failure to expand and evolve the model (Gelman & Shalizi, 2013). Therefore when severe testing cannot falsify the model, the inferences drawn become more credible (Gelman & Shalizi, 2013).

This is not an argument against, but can be accommodated within, critical realism. Importantly this highlights the danger of a subjective view of Bayesian statistics leading to complacency in selecting or averaging over existing models (Gelman & Shalizi, 2013). Within critical realism, complex models can and should be rigorously checked and falsified if they, and the credibility of their findings, are to be improved (Mingers, 2005; Gelman & Shalizi, 2013).

However, the complexities involved in the research question must also be fully appreciated. Within the arena of ambiguity regarding best treatment approach for potentially resectable pancreatic cancer patients, clinicians and policy makers are expected to make difficult treatment choices with wide ranging implications for many stakeholders. Juxtaposed with contemporary economic restraints on healthcare resources, ambiguity surrounding treatment benefits mandates cost-effectiveness analysis of treatment selection (Greenberg *et al*., 2010; Luengo-Fernandez *et al.,* 2013; Department of Health, 2015; Russell, 2016; Abott *et al.,* 2013). Challenges include simultaneously handling ever-emerging quantitative data from drug trials and the concept of value outcomes in cancer, which are neither static nor universal (Russell, 2016). Successful outcomes could be defined by the quantity of disease-free and overall survival time, regardless of treatment requirements, or they could be defined by the quality of survival time (Russell, 2016). Costs also go beyond costs of a particular treatment and include indirect costs that could be emotional as well as monetary costs to patients and healthcare systems (associated with complications, readmissions, et cetera) and

wider societal costs through work and leisure activities absences of patients and informal carers (Russell, 2016).

How then can such complex real-life concepts be best captured within a statistical model? Whilst it has been argued that such an approach can be accommodated within critical realism, Zadeh in 1969 proposed that a "radically different kind of mathematics" in the form of Fuzzy Logic was required to address the issues of uncertainty in 'real-world' problems. Fuzzy logic can be seen as both a mathematical tool and an overall theory that could encompass Bayesianism. Fuzzy logic as a theory will now be discussed to ascertain what, if anything, it can add to the philosophy driving this research.

3.2.2 Fuzzy Logic: a new map or a fellow traveller?

Building on Bertrand Russell and Max Blacks' analysis of the problems of uncertainty and vagueness in 'real-life' problems and the challenge this posed to classical logic (Black, 1937; Black, 1963; Russell, 1923), Zadeh in conceiving fuzzy theory offered a method for dealing with uncertainty (Sadegh-Zadeh, 2001). Zadeh hit upon what lies at the core of limitations of current approaches to statistical modeling in medicine (Zadeh, 1965a; Zadeh, 1965b). Patients are animate systems, therefore orders of magnitude much more complex than man-made systems (Zadeh, 1962; Zadeh, 1969). Traditional mathematical techniques dealing with probability, precisely defined points and sets et cetera are therefore simply inadequate (Zadeh, 1962). It follows that medical professionals, in research practice

terms, are animate systems analysts (Zadeh, 1969). Decision making in pancreatic cancer surgery is complex and rife with uncertainty. Methods equipped to deal with uncertainty surrounding complex animate systems (patients) therefore need, as Zadeh put it a "radically different kind of mathematics" (Zadeh, 1969).

Although fuzzy logic based methods for decision making are in their infancy, its application to: breast, prostate and bladder cancers have demonstrated superiority in terms of accuracy of predictions over traditional log regression and artificial neural networks (Seker *et al.,* 2003; Catto *et al.,* 2006; Abbod *et al.,* 2005; Catto *et al.,* 2003; Catto *et al.,* 2009). The models for breast and prostate cancer prediction were based on histopathology and molecular data only and, despite small sample sizes (breast: n= 100; prostate n= 41) reported greater than 80% accuracy of prognostic prediction (Seker *et al.,* 2003). Fuzzy models dealing with bladder cancer combined clinical, histopathological and molecular data (n=109 to 609) and reported model prognostic predictive value of greater than 88% (Catto *et al.,* 2006; Abbod *et al.,* 2006; Catto *et al.,* 2003; Catto *et al.,* 2009). Individualised risk prediction using fuzzy logic has been demonstrated elsewhere. Brand *et al.* (2006) used fuzzy modeling to show that the influence of smoking on development of colorectal cancer in hereditary non-polyposis was dependent on gene mutation, gender and age. This will enable the development of clinical risk scoring and individualised prevention strategies (Brand *et al.,* 2006). A fuzzy expert system has also been shown to provide individualised prediction of the need for neonatal resuscitation with 74% sensitivity and 94.8% specificity enabling streamlining of patients and planning

for resource availability in high risk cases (Resi *et al.,* 2004). Ghavidel *et al.* (2014) found that a fuzzy decision tree model was slightly more superior to a crisp decision tree model (AUC 0.9 versus 0.86; accuracy 0.98 versus 0.95; sensitivity 75% versus 58.3% and specificity 98.6% versus 97.4%) in predicting early mortality risk in coronary artery bypass graft surgery. Both decision trees outperformed log regression models. Furthermore, fuzzy modeling has been shown to out perform logistic regression in predicting severely depressed left ventricular ejection fraction following admission to intensive care unit based on variables acquired within 6 hours of admission (Pereira *et al.,* 2015).

Impressive as some of these results might be perceived, the doctrine driving the logic behind fuzzy method is that of approximate reasoning based on inference, therefore validity of these methods will only ever be approximate (Haack, 1979; Haack, 1980). How then can such a method ever lead the researcher on a path to personalised precision medicine? Such an approach stands accused of actually replacing scientific precision with scientific permissiveness resulting in imprecise thinking (Zadeh, 1996a; Zadeh, 1996b). These studies also have the six epistemological deficiencies as outlined by Ackoff in his critique of classic (positivist) operational research. None of the models displayed a learn and adapt ability, quality-of-life values were not factored in, models presented an abstraction of systems problems with a predict and prepare rather than synthesising planning paradigm, the interdisciplinary nature of healthcare system was neglected and finally these studies seemed to value objectivity of their predictions rather than considering all stakeholder affected by

the decision making process. These studies did not explore the "what if" scenarios of differing treatment options or individualised risk predictions associated with these options such as postoperative complications or treatment side effects. If fuzzy logic is to move research beyond Kantian's "What can I know?" to "What shall I do?" as promised (Sadegh-Zadeh, 1983), it can be seen as a fellow traveller on the journey to dealing with uncertainty and complexity rather than a new map to guide the way.

Ackoff, in his own critique of the second wave of systems thinking in operational research introduced one of the first examples of problem structuring method when he talked about replacing the problem-solving paradigm with an interactive planning method (Ackoff, 1979b2). A key epistemological aspect of problem structuring methods is the aspect of multiple perspectives and navigating human relationships (Midgley, 2000; Checkland, 1981; Checkland, 1987; Eden, 1987; Friend, 2001; Rosenhead, 1996; Rosenhead, 2006; Mingers & Rosenhead, 2004; Rosenhead & Mingers, 2001). This led to the third wave of systems thinking within operational research known as critical systems thinking which Ulrich formulated in his critical systems heuristics framework (Ulrich, 1983). I will now present critical systems thinking as a complementary and overlapping field with critical realism that can have a symbiotic relationship resulting in the enhancement of both. Furthermore, critical systems thinking provides a framework for enhancing modeling techniques. I will then outline how the work of Cilliers on complexity theory provides a framework for further enhancing and

enriching these fields before detailing how complexity theory will provide a theoretical map for my research.

3.2.3 Critical Systems Thinking

Ulrich, in his seminal work 'Critical heuristics of social planning' (Ulrich, 1983) introduced both a philosophical foundation and a practical framework, termed critical systems heuristics, for critical systems thinking (Kruger *et al.*, 2019). The basis of this framework was that the definition of a problem, proposals for improvement, and outcome are all dependent on the whole system (Ulrich, 2002) therefore systems boundaries must be rationally justified through dialog with both the involved and affected (Ulrich, 2012; Ulrich, 1987; Kruger *et al.*, 2019). Therefore Ulrich argues that boundary judgments cannot be separated from value judgments hence embedded in Ulrich's work are the guiding principles of rationality and universalisation (moral judgments are applicable to everyone equally) (Kruger *et al.*, 2019; Midgley, 2000).

In a practical sense his work provides a framework for the ethical process of debating systems boundaries. Set around four categories of: motivation, control, expertise and legitimacy this provides 12 boundary questions that have been used in a heuristic manner to debate what the system in question is and what is ought to be (Table 25). Midgley extended this work by considering situations where conflict arises between different values and boundary judgments (Midgley, 2000). They postulated that stabilisation of a situation where a conflict between two ethical boundary judgments arise can

be achieved by imposing a 'sacred/valued' or 'profane/devalued' judgment on marginal elements (Midgley, 2000; Midgley, 2007). Hence where the marginal element is deemed profane the primary ethical boundary becomes the main reference for decision making. Where the marginal element is deemed sacred the secondary ethical boundary becomes the main reference for decision making. Ulrich and Reynolds (2010) later built on this earlier work by focusing on working constructively with tensions between opposing perspectives. Accordingly, table 9 represents boundary critique of this research not as an expert-driven process of boundary setting but rather a participatory process of unfolding and questioning boundary judgements as set out by Ulrich and Reynolds (2010) by addressing conflicts including; 'situation' *versus* 'system', 'is' *versus* 'ought', concerns of 'those involved' *versus* 'those affected', stakeholders' 'stakes' *versus* 'stakeholding issues'.

Table 9: Boundary Judgement Applied to the Management of Potentially Resectable Pancreatic Cancer (Ulrich & Reynolds, 2010)

Source of Influence	Boundary Judgement Informing Pancreatic Cancer Management as the System of Interest			
	Social Roles (Stakeholders)	Specific Concerns (Stakes)	Key Problems (Stakeholding issues)	
Motivation	*1.Beneficiary: who ought to be/is the intended beneficiary of the system?*	*2.Purpose: what ought to be/is the purpose of the system?*	*3.Measure of Improvement: what ought to be/is the system's measure of success?*	The Involved
	Patients, clinicians and policy planners	To optimise outcomes for patients with potentially resectable pancreatic cancer by delivering the right treatment to the right patient at the right time with the right outcomes determined in collaboration with the individual patient. To maximize cost effectiveness of service delivery.	<u>Short term</u>: Accuracy of models' predictions of individualised outcomes across competing treatment strategies. Revealing new insights that will direct future research. <u>Longer term</u>: Acceptance and utilisation of predictive model into clinical practice. Prospective cost effectiveness and cost benefit analysis of the impact of model implementation.	
Control	*4.Decision maker: who is/ ought to be in control of the conditions of success of the system?*	*5.Resources: what conditions of success are/ought to be under the control of the system?*	*6.Decision Environment: what conditions are/ought to be out of the control of the decision maker?*	
	Initially health professionals involved in delivering the service with organisational backing and support.	Research project, financial and human resources, wider professional and social network to raise awareness of the project.	i) Interested groups affected by the outcomes (patients) ii) Expertise un beholden to the decision maker	
Knowledge	*7.Expert: who ought*	*8.Expertise: what*	*9.Guarantor: who*	

	to be/is providing relevant knowledge and skills for the system?	*ought to be relevant new knowledge and skills for the system?*	*ought to be/are regarded as assurances for successful implementation?*	
	i)The multidisciplinary team of the West of Scotland Pancreatic Unit. ii)Experts in decision-making, health technology assessment and cost-effectiveness analysis. The above informed by natural and social sciences.	Interdisciplinary and intersectional facilitation skills. Technical skills in computational statistics.	Competent and validated professional and non-professional knowledge. Avoidance of scientism (sole reliance on objectivity and statistical facts). Avoidance of managerialism (sole reliance on facilitating communication).	
Legitimacy	*10.Witness: who ought to be/ is representing the interests of those negatively affected by but not involved with the system?*	*11.Emancipation: what ought to be/are the opportunities for the interests of those negatively affected to have expression and freedom from the worldview of the system?*	*12.Worldview: what space ought to be/is available for reconciling differing worldviews regarding the system among those involved and affected?*	The affected
	Collective representation of professionals and patient bodies through liaisons with Pancreatic Cancer United Kingdom to gain qualitative assessment of views of all affected.	Open to challenge from all those potentially affected including patients, patient advocacy groups, professionals and funding bodies	Manage conflicts of interest between a political drive to effectively manage resources and the needs of individual patients affected by changes to the system	

Much of Ulrich's work in problem structuring is concerned with the socially constructed power struggles and the different frameworks of people (Kruger *et al*., 2019). However the research questions being addressed in this thesis involves a system that includes not only frameworks of people but many other elements as well.

Cilliers combined thinking about boundaries with concerns relating to complexity (Kruger *et al*., 2019). Both he and problem structuring thinkers such as Ulrich agree that both limited knowledge of systems

as a result of boundaries and complexity exist and therefore require a critical and ethical imperative in the study and understanding of such systems (Kruger *et al*., 2019). However, for Cilliers complexity has to do with the interactions and relationships amongst elements (Kruger *et al*., 2019). By combining thinking about boundaries with concerns pertaining to complexity and uncertainty Cilliers' work came to represent a new critical complexity paradigm giving a philosophical perspective on complex systems by taking cognisance of the insights from the field of post-structural philosophy (Midgley, 2007; Kruger *et al*., 2019). This work also provides an opportunity to challenge the role of operational research in how it relates to bigger societal questions (Preiser & Woermann, 2016).

This mirrors a move within healthcare research to view healthcare systems as complex adaptive systems which have been formally defined as "a collection of individual agents with freedom to act in ways that are not always totally predictable, and whose actions are interconnected such that one agent's actions change the context for other agents" (Plsek & Greenhalgh, p.625). Although there had been some debate over the precise terminology, complex adaptive systems are widely accepted to include: embeddedness, fuzzy boundaries, nested systems, self-organization, distributed control, emergence, non-linearity, unpredictability, historicism, change phases, sensitivity to initial conditions, non-equilibrium, adaptation and co-evolution (Kernick, 2006; Litaker *et al*., 2006; Plsek, 2003; Holland, 2014; Byrne, 1998; Manson, 2001; Long *et al.,* 2018). Importantly these key features of healthcare as a complex adaptive system mirror the work of Cilliers who provides ten propositions that represent the

characteristics of a complex system that are necessary for a 'lean ontology' of complexity.

Over recent decades there has been a growing appreciation within healthcare research of complexity theory (Long *et al.,* 2018). The argument supporting this move amongst healthcare researchers is that healthcare systems, due to their social nature, are qualitatively different from other systems and therefore require a different set of methods (Klein & Young, 2015; Eldabi, 2009; Tako & Robinson, 2015; Kernick, 2006). The argument continues that by continuing to hold the traditional and dominating Newtonian mechanistic conception of healthcare (Plsek & Greenhalgh, 2001; Plsek & Wilson, 2001) implementation of evidence-based medicine and healthcare innovations is being stymied (Kernick, 2006; Plsek & Wilson, 2001; Sanderson, 2009; Litaker *et al*., 2006; Plsek, 2003; Anderson *et al*., 2005). However, classical approaches to complex theory that have included agent-based modeling, simulation, and network analysis have made limited impact on healthcare (Long & Meadows, 2018; Fone *et al*., 2003; Bailsford *et al*., 2009; Long *et al.,* 2018). Such a classic approach to complexity theory involves researchers identifying rules that govern behaviours attributing them to the agent (local rules) or the environmental pattern (agents) (Long *et al.,* 2018). A theory of local rules are then built into a statistical model and tested against reality (Holland, 2014; Manson, 2001; Byrne *et al.*, 2013; McKelvey, 1999). Low implementation rates of such models have been attributed to: lack of good quality data (Brailsford *et al*., 2013; Brailsford, 2005; Robinson & Pidd, 1998; van Lent *et al*., 2012), complex nature of healthcare systems with multiple intersecting

stakeholders (Klein & Young, 2015; Eldabi, 2009; Brailsford *et al*., 2013; Robinson & Pidd, 1998; Brailsford *et al*., 2009; Kirchhof & Meseth, 2012) and the high time and expertise cost required to build sufficiently complex models that are ecologically valid (Brailsford *et al*., 2013; Brailsford, 2005; Robinson & Pidd, 1998; van Lent *et al*., 2012; Brailsford *et al*., 2009; Tunnicliffe-Wilson, 1981; Lane *et al*., 2003; Barnes *et al*., 1997).

The challenge of delivering personalised realistic medicine by optimising outcomes for potentially resectable pancreatic cancer through personalised predictive medicine is framed as both an operational and healthcare research problem. The next step will be to use the work of Cilliers on complexity theory as a lens through which to view this problem in the hope that this will provide new insights and broaden perspectives for informing contemporary practice (Kruger *et al*., 2019).

3.3 Complexity Theory: the new road map

A prerequisite for this section is defining the research problem as complex. After all it could be argued that superficially many components of the pancreatic cancer management pathway appear to be well defined and protocol driven inferring a complicated, rather than complex system. Jackson and Keys proposed a framework, known as a system of systems methodologies, to classify a problem context as either simple or complex reflecting Ackoff's terminology of 'mechanical' and 'systemic' respectively (Jackson & Keys, 1984). They also defined the relationship between stakeholder and

participants as unitary (when all agree on a common goal), pluralistic (where views and objectives differ) and later added coercive (irreconcilable differences in views and objectives). However this framework creates difficulties when defining the problem context of this research. Whilst aspects of the problem can are unitary, everyone wants to optimise treatment outcomes and ensure the most effective use of resources when delivering a service, other aspects are pluralistic, the most obvious example being the debate as to whether to use neoadjuvant therapy in cases where pancreatic cancer is resectable at presentation. There are also differences of opinion on the definition of resectability when interpreting imaging scans and the definition of R0 resection. Arguably aspects of the debate tend towards the coercive with some commentators adamant that for those who present with resectable disease and are treated with neoadjuvant therapy but do not proceed to surgery, the window of resectability was lost. Conversely others believe that such patients were filtered away from futile surgery with the associated impact on cost and quality-of-life. This reflects wider criticisms of the framework related to problems that do not fit unambiguously into one category, and where participants might well disagree on the unitary, pluralistic or coercive context (Midgley, 2000; Mingers, 1992).

Snowden and Boone developed the Cynefin framework to offer insight into how problem contexts can be classified in a way that assists decision makers in understanding the context in which they are operating (Snowden & Boone, 2007; Kruger *et al*., 2019). This framework centres around cause and effect (Kruger *et al*., 2019) and

consists of five difference contexts: simple (stable cause-and-effect relationships), complicated (known unknowns), complex (no apparent cause-and-effect relationship established), chaotic (the unknowable) and disorder (simultaneous multiple opposing perspectives that compete for prominence) (Snowden & Boone, 2007). As table 10 shows aspects of the research problem belong to the first two contexts, an ordered world in which fact based decisions can be made (Kruger *et al*., 2019). However, on deeper analysis many of the research questions move into the complex and chaotic contexts, an unordered world where patterns are used to make decisions (Kruger *et al*., 2019). Overall it could be argued that we are actually dealing with the context of disorder. As Kruger *et al*. (2019) pointed out it can be particularly difficult to recognise when one is operating in the context of disorder. Snowden & Boone (2007) offered a way out of this realm through breaking the situation down into constituent parts (table 10).

Table 10: Classification of Problem Context

Simple (cause-an-effect)	Complicated (known unknowns)	Complex (no established cause-and-effect)	Chaotic (the unknowable)	Disorder (simultaneous multiple perspectives)
Outcomes from pancreatic cancer are poor. Surgical resection is the only potentially curative treatment and adjuvant therapy has been proven to prolong survival time. Neoadjuvant therapy can convert borderline resectable and locally advanced cases to resectability. Obtaining multimodal treatment in either upfront surgery or neoadjuvant pathway prolongs survival time.	Up to 50% of patients in the upfront surgery pathway fail to receive adjuvant therapy due to a combination of early disease recurrence, post operative complications and decline in function. Which patients will and will not receive adjuvant therapy is unknown at the time of treatment pathway selection. Multimodal treatment in either pathway prolongs survival but, as with the above, the pathway in which multimodal treatment is most likely is unknown at the time of treatment pathway selection. Which pathway is more cost effective?	Patients with the same tumour type, location, stage, resection margins and postoperative course have differing survival outcomes in both pathways. Outcomes from trials of neoadjuvant therapy for borderline resectable and locally advanced disease cannot be assumed to apply to cases of disease that is resectable at presentation. Outcomes from a recent RCT comparing adjuvant therapies and establishing the use of mFOLFIRINOX as the first line adjuvant therapy agent cannot be assumed to equate with upfront surgery pathway being the superior pathway choice for resectable pancreatic cancer. Equally preliminary results from a RCT reporting superior survival outcomes with neoadjuvant therapy for resectable	Would patients with resectable disease treated in the neoadjuvant pathway and who do not proceed to resection have been better served in the upfront surgery pathway? Would patients in the upfront surgery pathway who did not receive adjuvant therapy have been better served in the neoadjuvant pathway? Can gene sequencing lead to better patient selection for available treatment pathways (at present unknown) As the percentage of patients presenting with resectable disease is small: 1) will studies underway to establish earlier diagnosis through identification of biomarkers be successful 2) if such breakthroughs are made how will the resulting	Upfront surgery for resectable disease is proven to prolong survival therefore the research focus should be on developing more effective adjuvant therapies and the focus at service delivery level should be on fast-tracking patients with resectable disease to early surgery followed by measures to increase the percentage of patients receiving adjuvant therapy. The use of neoadjuvant therapy carries the risk of loosing the window of resectability and optimistic assumptions about its use are based on small, underpowered studies with a high degree of heterogeneity. Neoadjuvant therapy has an increasing body of evidence demonstrating a survival advantage over upfront surgery. It allows time to

		pancreatic cancer compared to upfront surgery cannot be assumed to equate with neoadjuvant pathway being the superior pathway choice for resectable pancreatic cancer in light of the fact that the reported survival times in the neoadjuvant arm of this trial are lower than those reported in the former trial. The impact on quality-of-life of different treatments and interventions both in the short and long term, and how this might affect patients' decision-making.	data and its analysis affect pathway decision making, i.e. will an increased pool of patients with earlier disease show that early resection and adjuvant therapy is better or will the marginal survival advantage with neoadjuvant therapy reported in some studies still stand? Will gene targeted therapy come to fruition and if so will it be cost-effective? Will it be delivered pre or postoperatively? Will patients with the same or similar genetic profiles still have variation in their treatment outcomes and what other factors will determine this and to what extent? Will biomarker driven treatment sequencing be established and how will this affect outcomes in, and comparison of, treatment pathways? Will the afore mentioned potential breakthroughs be cost-effective and how will they affect the	filter more aggressive disease, which progresses despite neoadjuvant therapy, away from costly, high risk yet futile surgery therefore has a cost-effectiveness advantage. Any cost-effectiveness advantage reported with neoadjuvant therapy must be offset against moving costs and resource utilisation away from a surgical service budget to oncology and palliative care service budget. Such reports must also be reconsidered in light of emerging improvements with new adjuvant therapies. The key to optimisation of outcomes is better patient selection. The focus should therefore be on precision medicine through gene targeted therapies. The key to optimisation of outcomes is earlier diagnosis followed by early surgery. The focus should

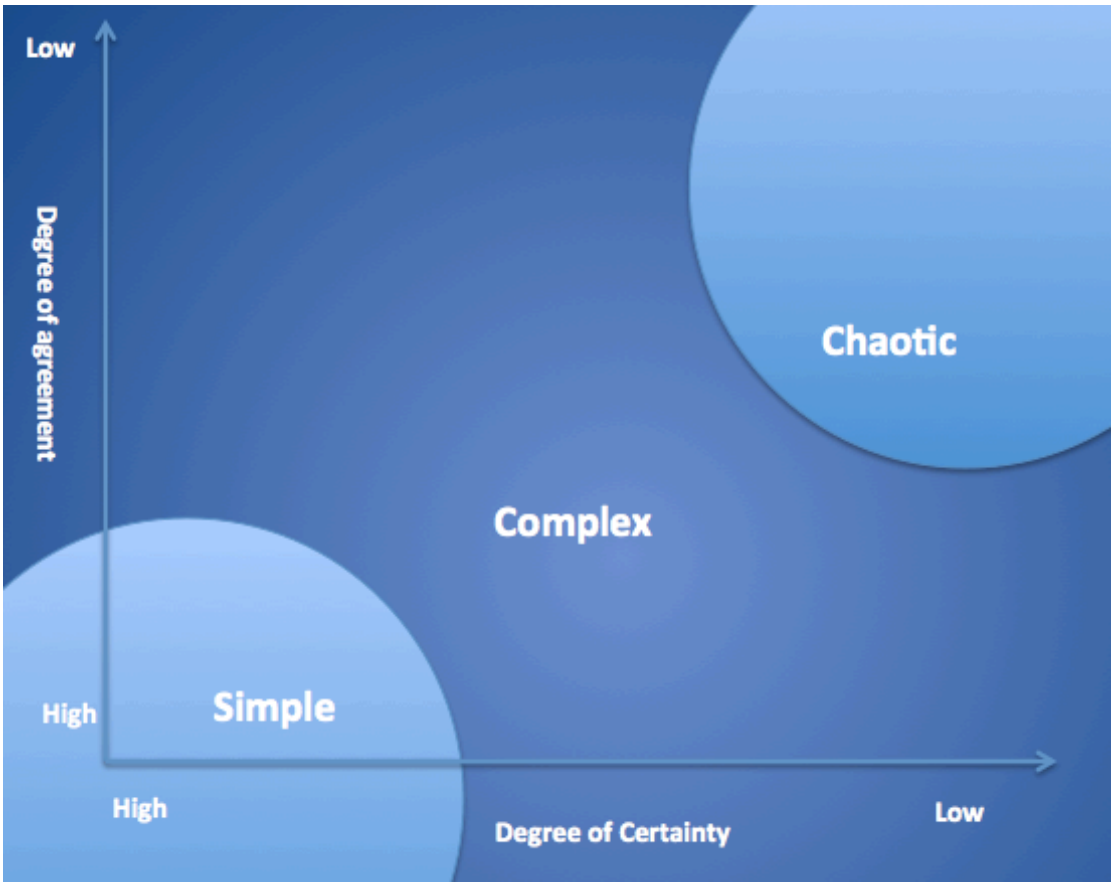
			cost-effectiveness analysis of competing treatment pathways? What impacts will this have on the wider NHS budgeting and the structuring of service delivery?	therefore be on screening and early detection combined with improvements to the upfront surgery pathway.

Such unfolding theorisation within operational research has particular implications for research that seeks to use complexity theory as a lens through which to view healthcare systems. On a daily basis decisions regarding the management of pancreatic cancer, and indeed throughout all areas of healthcare, are being made based on contested, limited and incomplete data (Greenhalgh & Papoutsi, 2018). The Newtonian world-view that permeates much of medical practice (Waldrop, 1992; Plsek & Greenhalgh, 2001) assumes that the scientific quest for certainty, predictability and linear causality will be achieved through ongoing RCTs (Greenhalgh & Papoutsi, 2018). Yet there is growing and insurmountable evidence that RCTs can address only a fraction of the unanswered questions within a healthcare system (Cohn *et al*., 2013; Braithwaite *et al*., 2017; Marchal *et al*., 2013; Greenhalgh & Papoutsi, 2018). To illustrate, preliminary results from the first RCT reporting superior survival outcomes with neoadjuvant therapy for resectable pancreatic cancer compared to upfront surgery cannot be assumed to equate with neoadjuvant pathway being the superior pathway choice for resectable pancreatic cancer in light of the fact that the reported survival times in the neoadjuvant arm of this trial are lower than

those reported in a RCT comparing adjuvant regimes in the upfront surgery pathway (Unno *et al*., 2019; Conroy *et al*., 2018). But how do the findings of the latter trial affect decision making when up to 50% of patient in the upfront surgery pathway do not receive adjuvant therapy? How do the findings from both of these RCTs apply to patients who would not have met the trail inclusion criteria? RCTs rather than providing clear guidance for adaptive behaviours have actually resulted in raising more questions.

Langton has termed circumstances that require adaptive behaviours, such as adopting the neoadjuvant pathway instead of the traditional upfront surgery pathway for resectable pancreatic, "the edge of chaos" (Langton, 1989). The centre area in Figure 3 therefore represents the reality of daily clinical practice where adaptive solutions, workarounds and general muddling-through (Greenhalgh & Papoutsi, 2018) are deployed in a reality where there exists insufficient agreement and centrality to make the correct decision or next step obvious but not so much uncertainty or disagreement to send the system into complete chaos (Stacey, 1996). The implications for this research is that such complexity must be placed at the centre of the unfolding story regarding how to optimise outcomes for pancreatic cancer with emerging and ongoing RCTs being augmented by the study of how we can best deal with unpredictability, uncertainty and generative causality (Greenhalgh & Papoutsi, 2018). This will now be taken forward by using Cillier's ten point 'lean' ontology of complexity.

Figure 3: Diagrammatic depiction of the relationship between complexity in daily clinical practice and degree of agreement and certainty



3.3.1 The Application of Cilliers' 'Lean' Ontology of Complexity

Complex Systems Consist of a Large Number of Elements

In viewing personalised realistic medicine as an operational research application, it must be acknowledged that the discipline of operational research constitutes several types of analytical models, which have been distinguished as classical and enhanced operational research (Jackson, 1988). Enhanced operational research encompassed newer methodological approaches in different contexts

involving a large number of conceptual models but both enhanced and classical disciples encompass a large number of concepts, elements and ideas applied to an even larger number of combinations (Kruger *et al*., 2019). This implies a finite number of elements that will impose an artificial boundary in the operational research system which some have argued is necessary from the observer's perspective to study a complex system (Cilliers, 2008; Mowat & Davis, 2018). However, Merali (2006) goes further in conceptualising the world as a networked world. Hence it stands that a collection of concepts and techniques does not constitute an operational research system and equally an operational research application does not exist in isolation (Merali, 2006; Kruger *et al*., 2019). Therefore whilst elements, such as society and the economy for example, may not interact with the application in a deterministic way they will interact and merge with the application (Kruger *et al*., 2019).

The Level of Interaction Among Elements is Fairly Rich and there are Loops in the Interactions Amongst Elements

The implementation of an operational research model will result in a multi-level, cross-scale, cross-sector interaction both within the boundaries of the model and the wider economic and societal environment in which it operates. The relationships between these elements are both mathematical and application concepts, and it is precisely these relationships that have resulted in many advances in the field of operational research (Kruger *et al*., 2019).

To illustrate these first three points, at a biological level pancreatic cancer begins with the complex process of carcinogenesis which itself depends on a large number of variables. Indeed it has been hypothesized that every tumour is unique and that the spectrum of biological changes that determine human tumour formation and behaviours is infinitely variable and regulated at multiple spatial and temporal scales (Grizzi & Chiriva-Internati, 2006). It follows that different treatment sequences will have infinite variability in their impact across these multiple scales on a biological level. Through multiple processes and controls that involves feedback loops within molecular carcinogenic pathways, these micro-scale processes have macro scale manifestations (Grizzi & Chiriva-Internati, 2006). This will culminate in differences in how the tumour behaves at organ, system and entire organism level, which will determine overall and disease-free survival times. Environmental, societal and political elements will also interact either in a deterministic way or through mergence. Such factors influence diet and lifestyle which can alter both an individual's risk of developing cancer as well as their general health and predisposition to other comorbidities which will affect their physiological reserve to withstand chemotherapy and/or major surgery with any associated complications that might occur. Socio-economic factors can also affect access to health care services with poorer socioeconomic conditions having been proven to result in poorer outcomes across all cancers. At politico-economic level government funding and the setting of society's willingness-to-pay threshold by the treasury, which is directly integrated into statistical models to perform cost-effectiveness analysis, will determine which treatments are made available within the NHS. Loops in the

interaction amongst elements at this level could manifest as strong public pressure to make certain cancer treatments available on the NHS.

The Elements Interact Dynamically

An operational research model is only meaningful within the real-life context within which it is applied therefore any such model will be meaningless unless it can interact with the environment within which it operates (Cilliers, 1998). This not only means that elements within the model can interact mathematically but also that they represent the dynamic interaction with the wider environment within which the model is operating. This interaction between the operational research application and reality could be either physical or transference of information to facilitate shared decision making (Cilliers, 1998).

Interactions are Non-Linear

The concept of non-linearity within an operational research model means that small causes can have large results, as previously illustrated in the example of micro-scale processes having macro-scale manifestations in the process of carcinogenesis, with the converse being also true (Cilliers, 1998). For Cilliers the concepts of non-linearity is closely aligned with the principle of asymmetry. For example personalised realistic medicine within the context of this research is seeking the most effective delivery of treatment for individual patients. Even without the addition of considering

treatment delivery at the lowest cost this already introduces an element of competition that also introduces asymmetry because if the model was perfectly symmetrical there would be no need for an operational research model. For Cilliers this meant that non-linearity, asymmetry and competition are all inevitable components of complex systems (Kruger *et al*., 2019).

Conditions are Far From Equilibrium

Within a complex system non-linearity between components, the environment and whole systems results in a state of non-equilibrium (Capra, 2007; Prigogine, 1987). Interactions with the world are dynamic therefore systems in non-equilibrium have multiple states of states and become more robust through a process of adaptation than statical systems operating in a state of equilibrium (Kruger *et al*., 2019; Capra, 2007; De Villiers-Botha & Cilliers, 2010).

Each Element Is Ignorant to the Behaviour of the Whole System

Although there are mathematical relationships between elements within a model, each element only reacts to information available to itself. Cilliers emphasises that this characteristic should be considered carefully as the information applied to the individual element may be rich but that individual element cannot contain the complexity of the whole system (Cilliers, 1998).

The complex system as one that is dynamic, non-linear, not in equilibrium and with elements ignorant of the whole can also be

illustrated through the example of pancreatic cancer. Mathematically a dynamic system depends on either a set of different states or configuration patterns (z), a number of transition steps (ż) from one state to another caused by a generating factor (u), over a period of time (t) thus:

$\dot{Z} = f(z,u,t)$

where f is a non-linear function that is continuous and the dot denotes differentiation in time (t) (Grizzi *et al*., 2004; Abraham, 1991; Abraham & Shaw, 1992; Grizzi & Chiriva-Internati, 2006). Therefore, for example, time (t) to disease recurrence, depends on a large number of dynamic elements that are themselves connected and interact in non-linear ways (Grizzi & Chiriva-Internati, 2006). Therefore the time to disease recurrence depends not only on treatment received but highly heterogeneous tumour factors at molecular level, the patient's physiological reserve to cope with the insult of intervention and / or an associated complication of this (which is itself affected by wider societal, environmental, genetic and lifestyle factors), and whether such an occurrence results in incompletion of intended course of therapy or delay in, for example, commencing adjuvant therapy due to postoperative complication. What this is illustrating is that treatment pathways for pancreatic management are actually systems comprising parts that show systematic heterogeneity and have non-linear relationships with the variables influencing the system also being connected in a complex manner, therefore small alterations in variables can lead to very different outcomes (Grizzi & Chiriva-Internati, 2006). Clinicians will recognise concepts in irregular modes of carcinogenesis, erratic tumour growth, poorly understood patters of metastatic spread, and

variations in the response of similar tumour types to the same chemotherapy agents (Grizzi & Chiriva-Internati, 2006). The implications for this in moving research forward is that cancer does not conform to simple mathematical principles but instead classical notions of cause and effect must be replaced by concepts of control, bifurcation and turbulence (Grizzi & Chiriva-Internati, 2006) which will mean expanding the repertoire of modeling methods and further exploring the development of simulation modeling to better handle complexity and uncertainty (Long *et al*., 2018).

Interactions Have a Fairly Short Range

Operational research models have the aim of bringing about improvement with changes occurring locally, close to the application (Kruger *et al*., 2019) although interactions can also be wide ranging (Cillers, 1998). Wide-ranging interactions can mean that changes at local level and have regional and national impact (Kruger *et al*., 2019).

An Open System

When an operational research model is applied to a specific problem the model becomes exposed to the real world as an open system with a large number of elements having an influence on its formulation (Kruger *et al*., 2019). Therefore to properly formulate an operational research model it must be implemented and applied to a specific problem otherwise the model itself becomes a closed system with

results confined to a set of variables, which would be a gross oversimplification (Kruger *et al*., 2019).

Complex Systems Have a History

Cilliers viewed the history of a complex system as a collection of traces left distributed over the system open to multiple interpretations (Cilliers, 1998). Therefore good models can be reused but their use and history should not be determined by the provision of an optimal answer in one application as a plethora of other factors must be taken into account such as the influence of different stakeholders and unforeseen developments during implementation (Kruger *et al*., 2019). This also means that the memory of the model will vary between different applications, even when the same type of model is applied, hence the model memory appears to be contingent and dynamic (Cilliers, 2010).

In summary the alignment of the research question with the ten characteristic of complexity thinking as a 'lean' ontology means that the epistemological questions raised by Ackoff can be addressed (table 11). While it is acknowledged that a complex world cannot be simplified into a list of characteristics this allows one to indicate the complexity in the context of the reality in which this research will operate (Kruger *et al*., 2019). No claim is being made that existing mathematical and computational models are wrong or of little use, but rather that by adopting methodological pluralism and using complexity theory as a lens through which to view the system of pancreatic cancer management new methods of modelling for

decision making could be developed by gaining new perspectives in terms of emergence, boundary setting, lack of complete knowledge and responsibility or ethics for the consequences regarding definitions and choices of boundaries (Kruger *et al*., 2019).

Table 11: Summary of how complexity theory can address Ackoff's epistemological concerns

Ackoff's Epistemological Concern	Complexity Characteristics Addressing This Concern
The need for a system to learn and adapt	Non-equilibrium History
Lack of quality of life values	Rich interaction between elements Open systems
Systems of problems	Large number of elements Elements interact dynamically
Synthesizing Planning Paradigm	Large number of elements Elements interact dynamically Open system
Interdisciplinary interaction	Loops in interactions Non-equilibrium Open systems
Pursuit of objectivity: who can be affected	Dynamic interaction Short Range Open system

Emergence

Checkland (1999) defined emergence as the principle that entities exhibit properties, which are meaningful only when attributed to the whole, and not its parts. This corroborated Cilliers claim that one of the defining characteristics of a complex system is its emergent properties which cannot be reduced to the system component properties (Cilliers, 2010). Therefore complexity emerges as a result of the dynamic and non-linear interactions between elements within the system (Cilliers, 1998). Juxtapose this with the context of an open system and the magnitude of emergence can be difficult to quantify (Paul *et al.*, 2014) particularly as emergence can take many forms

including deeper understanding or raised consciousness about issues (Kruger *et al*., 2019).

Boundary Setting

It is impossible to solve a real-world problem by including all of reality therefore the problem must be framed in a specific way and the system must be modeled in a recognisable way which requires that it be bounded (Cilliers, 2005a; Kruger *et al*., 2019). The setting of boundaries can also generate knowledge through dialogue with stakeholder (Audouin *et al.,* 2013; Kruger *et al*., 2019). Importantly however the setting of a boundary constitutes that which is bound rather than intended to separate things (Cilliers, 2008). Furthermore boundary setting is not objective, and is both artificial and temporary (Kruger *et al*., 2019). As emphasised by the work of Cilliers, operational research's epistemology, through complexity theory, can both accommodate and be broadened by this (Kruger *et al*., 2019).

Lack of Complete Knowledge

Complete knowledge of a complex system is not possible but rather knowledge in terms of a certain framework is (Kruger *et al*., 2019). As Midgley *et al*. stated " if we accept the systems idea that everything is ultimately connected, then no theoretical knowledge, however well elaborated, can accurately reflect reality" (Midgley *et al*., 1998, p.160). Hence the generation of knowledge within a complex system is exploratory and temporary (Cilliers, 2005b). Cilliers (2005b) and Woermann (2010) argue that rather than this

being an excuse for relativism, this represents a challenge to develop a new kind of scientific understanding (Cilliers, 2007).

Responsibility (Ethics)

Both the artificial nature of boundary setting and the provisional nature of knowledge means that a level of uncertainty will prevail in model outputs which means that responsibility must be taken for intended and unintended consequences when a system does not reflect reality (Cilliers, 2008; Woemann & Cilliers, 2012; Ackoff, 1974; Gallo, 2004; Ormerod & Ulrich, 2013) particularly as boundary definitions involve a value based judgment (Audouin *et al*., 2013; Ulrich, 1983; Midgley, 2000).

In summary as Law and Mol (2002) have suggested other ways of relating to, accepting, producing or performing complexity must be developed. Therefore this research seeks to engage with what Tsoukas called conjunctive theorising by avoiding simplification and abstraction (or disjunctive theorising) and instead drawing on different kinds of data from multiple sources to move research towards a theory that can build a rich picture of pancreatic management pathways as a complex phenomenon (Tsoukas, 2017). Combining operational and healthcare research and drawing on influences from complementary paradigms of critical realism and systems theory and enhancing their impact by using Cilliers' complexity theory 'lean ontology' an open-world ontology is held. This posits that the interplay of causal powers or tendencies of domains of 'the real' leads to particular events, 'the actual' (Mingers

2005). These domains may be physical, social or conceptual (Mingers 2005, Zachariadis *et al.,* 2010) and these events may be observable or experienced by people and therefore become empirical (Mingers 2005) but that as a whole the world is open to multiple interacting influences and to ignore such layers of influence serves no analytical benefit (Tsoukas, 2017). Epistemologically in recognising that all knowledge, whilst provisional, is historically and culturally relative both epistemic reality (observer-independent access as a fallacy) and judgmental relativity (rational grounds for theory preference) are accepted (Mingers 2005, Zachariadis *et al.,* 2010). By amalgamating operational and healthcare research disciples in this way this research seeks to be theory driven and empirically focused from a complexity perspective. Through a 'systems mindset' methodological pluralism is embraced to expand the methodological repertoire. Specifically how imperfect data can be better utilised within statistical simulation models will be explored so that, as Long *et al*. (2018) have suggested, the potential for simulation modelling in the study of complexity in healthcare can be explored to attempt to expand capabilities for handling uncertainty, the emergent and engage in disagreements (Star, 2002; Fraser & Greenhalgh, 2001; Greenhalgh & Papoutsi, 2018). Methods of modeling and their ability to cope with uncertainty and capture system complexity will now be explored.

3.4. Methodology

3.4.1 Methods: Modeling

The term 'model' has been overused recently (National Research Council, 1991) which mandates clarification of the use of the term within this research, first in relation to decision analysis then in relation to predictive modeling. Although several definitions of a 'model' exist (Milton *et al.,* 2001) Box *et al.* (1978) defined two types of models: empirical and theoretical. Empirical models are used when the mechanism is either not understood or too complex to allow an exact model postulated from theory (Milton *et al.,* 2001; Box *et al.,* 1978). An example would be drug trials as here data speaks for itself in connecting cause (inputs) and effect (outputs). Theoretical models are based on physical or mechanistic theory governing the system. Pertaining to predictive models, Box *et al.* (1978) would therefore site logistic regression models as examples of empirical models. Considering the ambiguity in the existing body of research relating to the best treatment pathway for potentially resectable pancreatic cancer, and the complexity of factors that influence clinical decisions not captured in drug trials, the case could be made for moving away from empirical towards theoretical modeling on the basis that a "basic understanding of the system is essential to progress" as this provides a better basis for extrapolation than empiricism (Box *et al.,* 1978). It does however follow that with any model extrapolation beyond the range of data is never a safe option (Milton *et al.,* 2001). This leads to the second point that must be

addressed in regards to modeling; the controversy that surrounds its application within healthcare.

Models for decision analysis, cost-effectives and prediction have been met with criticism. Empiricists have emphasised the potential inaccuracies of input data whilst epidemiologists have raised subsequent concerns that logical assumption regarding cause and effect are wrong (Milton *et al.,* 2001; Henschke & Flehinger, 1967; Schwartz, 1979). This has led some to conclude that clinical judgment is an 'art' that cannot be quantified whilst others express distrust in the hidden nature of the 'black box' of modeling software that can easily be manipulated by proponents of a particular treatment option (Milton *et al.,* 2001; Schwartz, 1979). However, the basis of explainable AI in healthcare is being able to sit down with a patient and describe the basis for a particular course of action. Furthermore all bias that exists in data cannot be adequately addressed if decisions are not interpretable.

At the core of this cynicism lies disagreement about the degree of experimental or empirical evidence required prior to modeling (Milton *et al.,* 2001; Henschke & Flehinger, 1967; Schwartz, 1979). Clearly the evidence base surrounding the best treatment approach to pancreatic cancer is inconclusive. However, to simply wait until 'perfect' evidence exists paralyses progress in medicine, stymies the realisation of realistic medicine, and is negligent in light of the fact that decisions must be made on a daily basis despite limited data and implicit value on qualitative outcomes (Milton *et al.,* 2001; Kuntz *et al,.* 2013). Furthermore when debating the quality of available data to

model it must be remembered that data from trials themselves are not infallible nor generalisable as treatment is protocol driven, context is controlled for, study participants are a selected subset of the population, and late outcomes may not be recorded due to trial time horizon despite this potentially being a key factor in influencing clinical decisions (Kuntz *et al.*, 2013; Milton *et al.,* 2001). Therefore blinkered reliance on 'perfect' trial data can actually result in cognitive bias when making these daily clinical decisions (Kuntz *et al.,* 2013). This leads to the second misconception underpinning cynicism towards modeling in medicine, that models are to establish 'truth' when they are actually meant to guide clinical decisions. This also counters the concern by some that models may produce results that conflict with the decision maker's view (Kuntz *et al.,* 2013). It follows that models can reduce the cognitive bias inherent in decision making and help understanding of the decision process and the inherent trade-offs in complex decisions with sensitivity analysis showing effects of varying model parameters (Kuntz *et al.,* 2013; Milton *et al.,* 2001).

What the criticisms of modeling within health care have done is focus attention on adhering to principles of good modeling practice: transparency, verification (outputs being consistent with observed data), corroboration (results produced are similar other models), validation (internal, calibration, face, convergent and where appropriate, predictive validation) and accreditation (peer review of models) (Kuntz *et al.,* 2013; Milton *et al.,* 2001). This highlights the imperative need to make both methods and goals of modeling transparent (Kuntz *et al.,* 2013).

Wolpert & Rutter (2018) offered a potential solution to these issue when they developed a framework for using flawed, uncertain, proximate and sparse (FUPS) data in the context of complexity. This will now be explored followed by an examination of models for decision analysis of neoadjuvant versus upfront surgery pathways and then methods of individualised predictive modeling in the context of complexity.

Flawed Uncertain Proximate and Sparse (FUPS) Data

Data surrounding the treatment of pancreatic cancer is currently flawed, uncertain, proximate and sparse and is likely to remain so even with the emergence of further RCTs (Table 12). A healthcare system delivering pancreatic cancer management is therefore faced with the challenge of dealing with the gap between the ideal of comprehensive, clear data used in complicated contexts, and the reality of FUPS data used within the context of complexity (Wolpert & Rutter, 2018). Clinical decision making therefore moves from what the urban planner and philosopher Donald Schon called the "high ground" where manageable problems lend themselves to solutions through the use of research-based theory to the "swampy lowlands" where problems are more confusing and messy (Schon, 1984).

Table 12: Summary of how existing data on the management of potentially resectable pancreatic cancer can be viewed as FUPS data

FUPS Category	State of Existing Data
Flawed	Missing data from institutional data base Erroneously recorded or coded data on institutional data base Deviation from the planned treatment protocol, the impact of which is not fully explored through solely intention-to-treat analysis. High degree of heterogeneity across study populations, treatment protocols and clinical practices. Many existing studies are small, underpowered and have a high degree of bias.
Uncertain	How data is rated or conceptualised for example: how treatment toxicities are categorised and reported (all occurrences recorded, worst event recorded, total number of toxicities in population versus percentage of population experiencing a particular grade of toxicity), how post operative complications are rated, variations in the definitions of resectable disease and R0 resection across studies and variation in follow-up practices creating uncertainty in the accuracy of disease free survival time.
Proximate	A proxy for the focus of interest: overall and disease free survival commonly reported but quality adjusted survival time poorly understood.
Sparse	Only an estimated 10% of cases are resectable at presentation. Drug trails comparing upfront surgery and neoadjuvant approach often include borderline resectable and locally advanced cases in the neoadjuvant arm.

Lessons drawn from cognitive psychology and sociology suggest that the use of findings from data are influenced by key factors including the tendency to reject that which challenges the prevailing

assumptions and the tendency for power elites to protect their interests (Currie et al., 2012; Raghupathi & Raghupathi, 2014). Large multicenter RCTs are costly and carry a high level of prestige for the institutions involved. The data from such trials are therefore held in high esteem and often readily accepted within medicine with limited criticism. However, the reality is that the majority of daily clinical decisions are made based on FUPS data out-with controlled trial conditions and concerning patients who are not selected based on a strict inclusion criteria. Yet agents in the system are more likely to apply higher standards of evidence than to traditional practice regardless of the apparent flaws in the evidence supporting such traditional approaches (Muir, 2001). This highlights the need for decision support models not only to pay sophisticated attention to the merits and detriments of using FUPS data but to also answer the call for a greater consideration of the implications of the complexity of the healthcare system in both research and practice (Plsek & Greennalgh, 2001; Rutter et al., 2017). This means paying close attention to the properties of the complex system in which the data will be used (Wolpert & Rutter, 2018) which can be characterised as a collection of individual but interconnected agents with the freedom to act in ways that are not completely predictable, and whose actions changes the context for other agents (Plsek & Greennalgh, 2001).

Taking a lead from Wolpert and Rutter (2018) to move beyond the biomedical model as the only model of evidence to simultaneously acknowledge the dangers of both over-interpretation of FUPS data as well as non-use, the aim of modeling FUPS data pertaining to the management of pancreatic cancer will be to open up conversations

on findings rather than treating them as definitive facts (Wolpert & Rutter, 2018). In this way findings will be used to consider the complex reality they relate to but cannot fully capture so that narrative arguments and hypothesis can be contested and debated within the system rather than being dismissed due to FUPS-ness of the data or presented as definitive facts in order to aid decision making in the "swampy lowlands" of clinical practice (Wolpert & Rutter, 2018). In practical terms this means instantiating the key principles of analysing FUPS data within statistical models. Firstly all data is treated as a partial remnant with findings presented to convey associated limitations to interpretation. Secondly 'black box' statistical modeling will be avoided in favour of transparency and clarity. Thirdly triangulation will be used to contextualise findings from models based on FUPS data to explore how other information and modeling techniques refute or support these findings (Wolpert & Rutter, 2018). Specific decision analysis modeling methods and their ability to handle uncertainty will now be explored.

3.4.2 Decision-Analysis Modeling

With its roots in mathematics, ethics, game theory and economics (Albert, 1978), decision-analytical models are designed to perform decision analysis in a systematic, transparent and quantitative way under uncertainty (Kuntz *et al.* 2013). Von Neumann & Morgensterm (1953) first provided a mathematical framework, based on the axioms of utility theory, and synthesising concepts of probability and value, for 'rational' decision making under uncertainty. Ledley & Lusted (1959) then attempted to apply decision-analysis to medical

diagnostics, focusing on probability, logic and value and emphasizing the importance of Bayesian formulae in achieving this. However, they then drew on game theory when deciding upon best treatment selection, discussing what a physician could "win" and what nature could "lose" with various treatment options (Ledley & Lusted, 1959). Later Henschke & Hehingers (1967) in 'Decision Theory in Cancer Therapy' attempted to tackle decision making at a more complex level when attempting to analyse risk and benefits of prophylactic neck dissection for head and cancer. Neither of these seminal papers however formalised an analytical tool to analyse decisions (Kuntz *et al.,* 2013). It was not until the 1980s that the first textbook on decision analysis introduced the decision tree (Weinstein *et al.* 1980) followed by the first application of a Markov model, utilizing markov chains, in medical decision making (Beck & Pauker, 1983). Despite postulations in 1975 that the time had come for modeling in decision-analysis to become more widely used and accepted, despite not being a new idea then (Inglefinger, 1975), it is only in recent years that it has gained precedence. The reason for this is two-fold; firstly the perception of decision-analysis modeling within medicine (available evidence is too limited to model, qualitative factors are important in decision making but not suitable for quantitative analysis, and doctors could not be expected to use complex and sophisticated models in their daily practice) and secondly the development of software to support its application has accelerated the use of decision-analysis modeling (Ledley & Lusted, 1959; Henschke & Flehinger, 1967). Models for decision-analysis will now be explored and include: decision-trees, Markov, micro and discrete event simulation.

Decision trees use utility theory, or multiattribute until theory is developed to attach value to an outcome, even qualitative outcomes (Kuntz *et al.* 2013; von Neumann & Morgensterm, 1953). This type of model would have the benefit of modeling the complex and numerous branches involved in treatment pathways when deciding between treatments for pancreatic cancer in a logistical and linear manner that is easy to follow with all pathways and their transiton probabilities being transparent (Kuntz *et al.,* 2013). However, this model does not capture reoccurring events and is only applicable in decisions with short time horizons. For this reason I contest that Markov modeling would be a more appropriate choice of model as many features of the clinical process are captured taking into account the timing of these events, such as changing health states over time. Furthermore a Markov model can be computed analytically to give expected values such as life expectancy, or stochastically to measure predicted outcomes in addition to expected values (Kuntz *et al.* 2013). It allows both decision analysis in terms of health outcomes as well as cost-effectiveness analysis.

Markov modeling has been employed for decision analysis of upfront surgery versus neoadjuvant therapy for resectable pancreatic cancer (deGeus *et al*., 2016; Sharma *et al*., 2015). These studies were based on synthesised data from published trials and their output was not validated against patient level data. Whilst this approach adds flexibility in sensitivity analysis by incorporating explicit links between end points, it also carries methodological limitations that could inhibit its future application (Caro *et al*., 2010; Miettinen &

Caro *et al*., 2010). The challenge in applying the Markov model is in addressing its main disadvantage; the Markov property. This property means that transition probabilities within the model are treated as independent of the past history. However, in clinical practice a significant past medical history may make a post operative complication more likely which in turn could reduce the probability of a patient receiving adjuvant therapy. One strategy to deal with this would be to define health states according to past events (Kuntz *et al.* 2013) although there is a danger that the model could grow too large to manage as the number of possible health states could increase exponentially (Kuntz *et al.* 2013). Given the anticipated move towards future personalised targeted treatments the memory-less property of the Markov cohort model makes it less well equipped to handle individual patient data, which can result in reduced accuracy due to depletion of susceptibles and an over simplification of assumptions (Caro *et al*., 2010; Miettinen & Caro *et al*., 2010). Furthermore in light of the afore mentioned current challenges in pathway assessment for resectable pancreatic cancer, implementation of time-dependent transition probabilities when multiple health states and treatment sequences are considered would make programming and utilising such a model difficult (Caro *et al*., 2010; Miettinen & Caro *et al*., 2010).

A better framework for modeling treatment pathways for resectable pancreatic cancer could be offered through discrete-event-simulation (DES) approach as it captures a patient's experience in terms of events and also has the ability to track changes in patient characteristics, health status and treatment history in relation to

their impact on outcomes (Pan et al., 2018; Caro *et al*., 2010). By tracking the individual patient's simulated history, including multiple comorbidities allowing them to interact and effect outcomes, the number of health states could be reduced within the model (Kuntz *et al.* 2013). This could be achieved on a cycle-by-cycle basis or by using data distributions to simulate time-to-event hence allowing flexibility in how data is modeled (Kuntz *et al.* 2013). This potentially makes this approach a more accurate and efficient framework with the flexibility to incorporate future anticipated breakthroughs in personalised targeted treatments. This would however require a high number of simulations to reach a stable expected value which amounts to high costs in terms of time and computing power. Debugging such a model would be difficult compared to the Markov model which has a Markov trace which means that the proportion of the cohort in each health state can be given per cycle time which results in face validity of the model and good accuracy testing (Kuntz *et al.* 2013). Furthermore, the proportion of disease that is resectable at presentation is small considering the large data requirements for such a modeling framework. DES approach has not yet been applied to treatment pathway analysis for resectable pancreatic cancer to assess its level of accuracy.

Both these modeling approaches could be complementary. The curiosity is whether the Bayesian approach could be taken forward to achieve precision medicine with a model that can give individualised predictions of prognosis as well as failure events such as the risk of treatment complications.

3.5 Bayesian Theorem

Pearl (1990) defined Bayesianism by the attributes of: willingness to accept subjective belief as a substitute for raw data, reliance on coherent probabilistic models of beliefs and updating belief in light of new information through adherence to Bayes' conditionalisation.

A probability distribution, P is defined on a proposition space S, which contains all propositions that a system can represent and process (Kolmogorov, 1950). The probability evaluation $P(x)$, comes from P being defined on S for every proposition $x \in S$. Both x and y are contained within S, so the probability of x under the condition of y is a conditional probability evaluation from which we get Bayes' Theorem (Kolmogorov, 1950; Wang 2004):

$$P(y|x) = P(y|x)\, P(x) \,/\, P(y)$$

Accordingly the probability of the proposition h is the systems belief in h according to background knowledge K that could be data, experience evidence et cetera. The system starts with determining prior probability P_0 from knowledge K_0 at time t_0. When new knowledge, e, becomes available P_0 becomes a posterior distribution, P_1 so:

$$P_1(h) = P_0(h|e) = P_0(e|h)\, P_0(h) \,/\, P_0(e)$$

P_1 is based on K_1 which is a combination of previous and new knowledge, K_0 and e. Therefore Bayes' Theorem when repeatedly applied in this fashion is known as conditioning process and means that the system can learn and adjust beliefs according to this new

knowledge (Heckerman, 1999; Pearl, 2000). This also means that a probability evaluation *P(h)* is always conditional due to the implicit condition that *P(h)* is conditional on the relationship between *h* and *K* and not the objective property of *h* (Wang, 2004):

$$P_{K1}(h) = P_{K0}(h|e) = P_{K0}(e|h)\, P_{K0}(h) \,/\, P_{K0}(e)$$

Often however this implicit condition on which the dependency of a probability is dependent, is represented as a conditional probability or 'explicit condition' (Cheeseman, 1985; Heckerman, 1999; ,Pearl 1988; Pearl, 2000):

$$P(h|K_1) = P(h|e \wedge K_0) = P(e|h \wedge K_0)P(h|K_0)/\, P(e|K_0)$$

which some contest is improper and central to understanding the often under reported limitations of Bayesianism (Wang, 2004) as will be discussed later. First I will discuss the application of Bayes' theorem to the research question through Bayesian Networks (BN) before outlining these limitations and how they might be overcome.

Precision Medicine and the Role of Bayesian Networks

There is a move within contemporary healthcare towards precision medicine whereby probabilistic modeling is used to predict likely disease progression and/ or treatment outcomes for individual patients based on interpretation of patient data (Velikova *et al.* 2014; School *et al.* 2013). However, decision making in medicine can be fraught with difficulty due to underling uncertainties.

Treatment selection and prognostic reasoning at its very core concerns making predictions of future events despite inherent uncertainties. Traditionally prognostic models utilise supervised data analysis methods based on frequentist statistical paradigm, such as multivariate logistic regression analysis (School *et al.* 2013; Verduijn *et al.,* 2007). Limitations of this approach highlight the gap between theory and practical application of models. Such models regard prognosis as an isolated event at a pre-determined time, applying attribute selection prior to inducing the model and setting fixed roles of input and output variables to attributes (Verduijn *et al.,* 2007). Variables deemed important by clinicians may therefore be excluded. Furthermore this neglects the dynamic nature of care processes where outcomes today predict those of tomorrow hence expected patient outcomes evolve as more information becomes available (Verduijn *et al.,* 2007).

Prognostic Bayesian models, although in their infancy, allow for incorporation of individual patient data, disease progression and impact of different treatment options on the predicted outcome variable, such as life expectancy. This can be defined very simply as a probability distribution:

$$\Pr(\text{outcome}/\delta,\Im)$$

where δ denotes available patient data and $\Im$ denotes sequence of treatment events impacting on the outcome variable (Lucas *et al.,* 2004).

Through Bayes theorem the prior distribution and observed data are combined to update knowledge in the form of the posterior

distribution. Posterior probability intervals (PPI), or credibility intervals, represent the 95% probability that the predicted outcome lies between two values. This is often erroneously confused with frequentist-based 95% confidence interval, which means that 95% of the confidence intervals capture the true outcome under the null hypothesis (School *et al.,* 2013). Regarding personalised predictive outcomes for patients, PPI is therefore more accurate and easier to communicate to patients the predicted probability of their outcome lying between two values (School *et al.,* 2013).

BN are based on graphical formalism of a joint or multivariate probability distribution over a random set of variables and are sometimes referred to as acyclic directed graphs (Velikova *et al.,* 2014; School *et al.,* 2013; Stajduhar & Dalbelo-Basic, 2010). BN are based on the following set of formulisation.

BN are defined as a pair:

$$BN = (G, Pr)$$

where G is a graphical structure and *Pr* is the probability distribution. G = (V(G), A(G), where V(G) is a random variable taking on a set of values. Variables are represented as nodes within BN and any number of nodes can be included, therefore:

$$V(G) = \{V_1, V_2 \ldots V_n\}$$

where $n>1$. A(G) represent arcs which indicate probabilistic influence between two nodes: $V_i \text{è} V_j$ where V_i is termed the parent node and V_j the child node. The joint probability distribution (*Pr*) respects the dependence and independence between nodes and is defined as:

$$Pr(V_1, V_2 \ldots V_n) = \prod^{n}_{i\ 1} Pr(V_i / \boldsymbol{\pi}(V_i))$$

where $\boldsymbol{\pi}(V_i)$ represents the covariates of parent nodes to V_i. (Velikova *et al.,* 2014, Lucas *et al.,* 2004, Verduijn *et al.,* 2007, Stajduhar & Dalbelo-Basic, 2010).

Unlike traditional prognostic models that provide predictions of a single outcome variable, BN can be more complex, providing information on process variables (conditions that occur during the process) as well as outcome variables (endpoints of that process) (Verduijn *et al,.* 2007; Lucas *et al.,* 2004). Therefore in practice BN can predict outcomes pertaining to quality and not just amount of survival time (Lucas *et al.,* 2004). Furthermore predictions from prognostic BN can be used to support decision making in resource allocation as well as individual cases or case-mix adjustment or benchmarking in groups or populations (Verduijn *et al.,* 2007).

Where patient information is limited probabilistic inference can still make predictions based on global averages of the patient population (Verduijn *et al.,* 2007; Lucas *et al,.* 2004). As more information becomes available the predictions become more patient specific (Verduijn *et al.,* 2007). This highlights a further key benefit of prognostic BN; prognosis updating (Verduijn *et al.,* 2007). As the healthcare process evolves so does a patient's predicted prognosis. Bayesian methods underpin BN, which allows prognosis to be seen as a dynamic notion through probability updating with new and emerging information (Verduijn *et al.,* 2007). In practice this means clinicians involved at the later stages of care can use the same model,

adjusted for the events of the preceding care phases (e.g. complex surgical interventions) to make more timely and personalised predictions (Verduijn *et al.,* 2007). This further highlights an aspect of predictive medicine not captured in traditional prognostic models; prognostic scenario analysis. In real life events such as complications and hospital stay do not happen in isolation but rather as scenarios (Verduijn *et al.,* 2007). Algorithms exist within prognostic BN that can perform this type of probabilistic inference to predict a most likely scenario for patents or patient groups (Verduijn *et al.,* 2007).

This advantage links beneficially to a further aspect practice faced by clinicians and patient; the 'what if scenario'. By identifying a specific event the prognostic BN can supply a risk profile of the most likely scenarios leading to the stated event (Verduijn *et al.,* 2007). Such information can be incorporated into decision making regarding treatment options. Similarly BN can be used to perform risk factor analysis. When an unfavorable event, such as a post-operative complication, occurs it is important to identify variables that may have predicted occurrence or non-occurrence of said event and quantify this in terms of risk ratios (Verduijn *et al.,* 2007; Lucas *et al.* 2004):

$$RR(X^1) = \frac{P(X^1=x^1 / X=x,\Im)}{P(X^1=x^1 / X\neq x,\Im)}$$

X^1 is a variable that precedes the adverse event. $\Im$ is the background knowledge of the patient, or patient group. A high value for risk

ration means that X^1 is an important predictor for the event, X, occurring in that patient or patient group (Verduijn *et al.,* 2007).

In summary, BN are emerging as a promising, but as yet under utilised, solution with potentially extensive application to medicine owing to their ability to model uncertainty and causal relationships between variables. Bayesian statistical approach offers an alternative to the traditional frequentist paradigm of null hypothesis testing by allowing the integration of prior qualitative and quantitative knowledge (Velikova *et al.,* 2014; School *et al.,* 2013; Verduijn *et al.,* 2007). In this way BN allow the modeling of relationships between variables at various stages of the healthcare process, with predictions of outcomes evolving throughout the process by utilising all available patient data at that time (School *et al.,* 2013). Predictions can therefore be made for all variables, not just outcome variables (Velikova *et al.,* 2014; School *et al.,* 2013; Lucas *et al.,* 2004). How then can BN be applied to guide decision making through the extensive, but inconclusive and arguably ambiguous body of evidence underpinning the treatment of resectable pancreatic cancer?

Modeling Under Uncertainty: the unique challenge of potentially resectable pancreatic cancer and the application of Bayesian Networks

Arguably one if the most controversial aspects of Bayesian statistics is the elicitation of priors (Johnson *et al.,* 2010). Where considerable prior knowledge of a high quality exists, prior distribution can be objectively derived through meta-analysis (School *et al.,* 2013; Hampson *et al.,* 2014; Johnson *et al.,* 2010). Challenges however arise

in cases of rare disease or where existing prior knowledge is limited or ambiguous.

The evidence-base underpinning treatment options for resectable PC are multifaceted in the challenge it poses to predictive modeling. As outlined in previous sections trials of both adjuvant and neoadjuvant therapy are inconclusive in proving superiority of one treatment approach over another. Furthermore large, multi-centred RCTs comparing one treatment approach over another do not yet exist. The extensive body of research thus far accumulated in this field however means that pancreatic cancer cannot be viewed as a rare disease where no prior knowledge exists. In addition there also exists a separate body of research identifying predictive variables of survival outcome pertaining to pancreatic cancer (tumour size, lymphovascular invasion, albumin: CRP ratio et cetera) as well as extensive work from other disciplines within medicine looking at predictive modeling for outcomes of major surgery based on pre-existing patient factors that cannot be ignored. To summarise, the challenge of uncertainty with regards potentially resectable pancreatic cancer is not a lack of existing knowledge, as with rare diseases, but rather uncertainty permeates the extensive existing body of research in addition to separate but highly relevant body of prior knowledge accumulated out-with drug trials. How then can such information be modeled to meaningfully make individualised predictions of outcome?

In cases of uncertainty one option is to set objective priors, or uninformed prior distributions, which assumes ignorance of any

prior knowledge and accepts that observed value can lie between minus and plus infinity (Lin & Haug, 2008). Whilst those of a more objective, or even frequentist, stance would champion this approach as letting the observed data speak for itself free from bias, subjective Bayesians would argue that this neglects existing empirical knowledge hence stymying the progression of knowledge (School *et al.,* 2013; Lin & Haug, 2008). Instead they argue in favour of informative prior distributions where knowledge can be drawn from quantitative and qualitative sources, including expert opinion, where existing knowledge is limited (Hampson *et al.,* 2014). If this approach is adopted it is essential that prior precision, the degree of certainty in the prior knowledge, is specified with low-informative prior distribution being generally accepted as having limited impact on results (School *et al.,* 2013; Lucas *et al.,* 2004; Lin & Haug, 2008). Furthermore when uncertainty exists regarding prior distribution, sensitivity analysis exploring the impact of different prior distributions on the results is required (School *et al.,* 2013; Lucas *et al.,* 2004).

Building a Bayesian Network

There are various methods for implementing the afore mentioned components in a BN. Firstly in a naïve or uninformed BN parameters are either learned from data or expert estimations, with all independent variables acting as child nodes of dependent variables (Velikova *et al.,* 2014; School *et al.,* 2013; Verduijn *et al.,* 2007; Lucas *et al.,* 2004, Lin & Haug, 2008).

Alternatively Bayesian network structure algorithms can be employed to derive machine-learned network structure and parameters from data (Lin & Haug, 2008). Where parameters are learnt from data, the dataset must be complete, comprehensive, and comprise enough data to reliably identify probabilistic relationships between variables (Lin & Haug, 2008). Biases introduced during data collection will also reflect on the BN (School *et al.,* 2013). As outlined in the previous section BN, through its Markov condition, models a collection of dependence and independence statements (Verduijn *et al.,* 2007; Lucas *et al.,* 2004). An alternative approach to data learning may be to incorporate dependence analysis such as information-theoretical algorithms (Cheng *et al.,* 1997). Here mutual information for each linked variable is established from the data using an algorithm. Arcs are then added between variables, which are not conditionally independent given a conditioned set of variables (Lucas *et al.,* 2004; Cheng *et al.,* 1997). Each arc is then tested using conditional independence test whereby if independence is proven the arc is removed (Lucas *et al.,* 2004; Cheng *et al.,* 1997). In larger conditioning sets this approach however can become infeasible and less reliable (Lucas *et al.,* 2004). A hybrid approach of constructing the graph from data using lower-order dependence test then using this graph to restrict the search space of graphical structures in the second stage which is to use an algorithm to find a diagraph that best explains the data (Lucas *et al.,* 2004).

A combination of both approaches can also be utilised with human experts defining nodes and directed arcs to create network structure, and parameters then machine-learned from data. This approach is

particularly applicable where logical constraints, derived from functional relationships between variables, and qualitative probabilistic constraints, for example derived from stochastic dominance of distribution, can assess and verify the number of probabilities required for network construction (Lucas *et al.,* 2004).

Judgmental probabilities can be obtained through rigorously tested expert opinion, or from data. To begin the local conditional probability distributions are filled in: $Pr(V_i/\boldsymbol{\pi}(V_i)$ (Lucas *et al.,* 2004). Network conditional probability distributions are then often computed as a weighted average of a probability estimate based on available data and a prior multinonimal distribution defined as:

$$Pr(V_i/\boldsymbol{\pi}(V_i), D) = [n/n+n0]\ Pr_D(V_i/\boldsymbol{\pi}(V_i))+ [n_0/n+n_0]\ \Theta\ V_i/\boldsymbol{\pi}(V_i))$$

D is the dataset from which the probability distribution, $Pr_{D,}$ is estimated. n is the size of the dataset, D. Θ is the mulinomial prior over all possible values of V_i and n_0 is the number of past cases that contribute to Θ (Lucas *et al.,* 2004).

Finally the quality and clinical application of the BN must be tested before its use in practice. There are a number of methods available to do this including using the patient data to assess robustness BN output to inaccuracies in the probability distribution (School *et al.,* 2013; Verduijn *et al.,* 2007; Lucas *et al.,* 2004).

Addressing the Limitations of Bayesianism

In summary, BN can be understood as representation of uncertain interactions amongst variables. Prior probabilities are conditional upon the relationship between proposition *h* and prior knowledge K and are therefore more accurately termed 'implicit condition':

$$P_1(h) = P_0(h|e) = P_0(e|h)\ P_0(h)\ /\ P_0(e)$$

Bayesian learning is carried out by the above equation therefore the knowledge the system can learn must be represented as an 'explicit condition'. This however carries some restrictions. *h* must be a binary proposition that must be in *S* so that its probability, $P_0(e)$, can be defined and as greater than 0 otherwise it cannot be a denominator (Diaconis & Zabel, 1983; Pearl, 1990). Furthermore these restrictions are not applied to the implicit conditions, which need only be related to *S* and can include non-binary propositions such as subjective probabilistic estimates. Also a proposition assigned a prior probability of 0 could be assigned a non-zero prior probability from another source (Wang, 2004). What this means is that not all implicit conditions can be represented as explicit conditions and that knowledge not available when deciding priors cannot be learned or acquired in the system through Bayesian conditioning. In practical terms this means that prior knowledge can be probabilistic-valued but all new knowledge must be binary valued, propositions given a value of 0 or 1 cannot have this belief altered in light of new knowledge, and no novel concept can appear in new knowledge (Wang, 2004). This counters the claim by some that Bayes' Theorem is a generally applicable learning rule sufficient for reasoning in uncertainty. However without distinguishing implicit and explicit conditioning an illusion arises that knowledge supporting a probability distribution function can be expressed as an

explicit condition therefore learned by the system (Wang, 2004). How then can the model learn from new evidence that is not binary?

If there is a prior probability distribution, P_0, assigned to a proposition space, S, and new evidence shows that the probability of proposition e should be changed to p, ($P_1(e) = p$), assuming that the explicit condition is unchanged ($P_1(h|e) = P_0(h|e)$), then using Jeffrey's rule (Diaconis & Zabel, 1983, Jeffry, 1965; Kyburg, 1987; Pearl, 1988) every proposition of h in S can be updated resulting in a new distribution function:

$$P_1(h) = P_0(h|e) \times p + P_0(h|\neg e) \times (1-p)$$

In this way, if new evidence shows that e happens, or e's probability changes to 1, then Bayes' Theorem becomes a special function of Jeffrey's rule: $p = 1$.[2]

The second challenge is then how to process uncertain evidence e. If a similar approach is taken and a virtual proposition v is taken to represent new knowledge and:

$$P_0(e \mid v) = p \text{ (Cheeseman, 1986; Pearl, 1988)}$$

Then in consideration of this new knowledge a new probability calculation can be offered whereby the prior probability is conditonalised to v rather than updated:

$$P_1(h) = P_0(h \mid v) = P_0(h \mid e \wedge v) \times P_0(e \mid v) + P_0(h \mid \neg e \wedge v) \times P_0(\neg e \mid v)$$

In this way Jeffrey's law can overcome the restriction that new evidence must be binary (Wang, 2004). Furthermore if conditional

probability is defined by de Finetti's coherent ($P(x \mid y) = P(x \wedge y) / P(y)$) events with a prior probability of 0 can also be conditioned (Coletti *et al.,* 1993). Juxtapose this with Pearl's Neo-Bayesianism which adds topological structure in the form of a BN to traditional Bayesianism, and the afore mentioned limitations of conditioning can be overcome as conditional probability can be introduced independent of absolute probability values (Pearl, 1990; Wang, 2004). However, this does not mean that the system has a general way to revise implicit conditions, or put another way the background knowledge behind probability distribution. This means that if BN is to be applied to the research problem a choice must be made between: 1) accepting that the implicit condition, or domain knowledge determining probability distribution, is immune from modification or 2) all modifications of implicit condition are treated as updating therefore when new knowledge conflicts with old knowledge the old knowledge is abandoned (Wang, 2004).

Even though the distinction between implicit and explicit conditions are rarely made, this serves to prove that Bayesianism has limitations in handling uncertainty. This is because probability distribution function alone fails to show the degree of uncertainty about the function itself (Wang, 2004; Diaconis & Zabel, 1983; Demster, 1967). Although some contest that a point value and a density function produce the same results in decision making (Cheeseman, 1985) the counterargument is that standard deviation cannot capture the change in expectations (Wang, 2004). To illustrate, if a proposition is tested n times and produces the same results standard deviation remains independent of n at 0. However our confidence that the

results will remain the same in light of new information will obviously increase, which would not be captured through use of standard deviation (Wang, 2004). This could be addressed by replacing precise probability values with either a probability interval, with width of the interval indicating degree of certainty of the system (Grosof, 1986; Kyburg, 1988), or imprecise probability, where upper and lower probability values are used (Walley, 1991; Walley, 1996). Alternatively high-order probability where a second probability value is introduced to specify accuracy of first-order probability (Kyburg, 1988; Paab, 1991) could be considered. A belief function and a plausibility function could be introduced using Dempster-Shafer theory so that an evidence combination rule could reduce ignorance amount uncertainty (Dempster, 1967; Shafer, 1976). A frequency value and a confidence value could also be employed to represent uncertainty with confidence value being used to measure degree of ignorance (Wang, 1993; Wang, 2001). Pearl stated that ignorance was the lack of confidence and that confidence was the measurement of the extent to which a degree of belief could be modified by future evidence (Pearl, 1988; Wang, 2001). In other words an assessment of $P_0(e)$ measured by narrowness of distribution of $P_0(\mathrm{e} \mid c)$, as c ranges over all combinations of contingencies and is weighted by its belief in $P_0(c)$ (Wang, 2004). However this still does not capture ignorance about the implicit conditions leading some, like Wang (2004), to contest that whilst these approaches may handle representation of ignorance or uncertainty, Bayesianism alone cannot truly handle uncertainty.

To summarise BN are a powerful tool in modeling in uncertainty and have shown grate promise in their application to personalised realistic medicine. Bayesian approach enables calculation of other values from values in the same probability distribution and can even update previous probability distributions given some values in a new probability distribution, which has several advantages centering on capturing the dynamic nature of the healthcare process. However caution must be taken when appreciating how Bayesianism handles uncertainty in the knowledge base on which prior probabilities are calculated and updated in light of new evidence (Wang, 2004). Importantly however it must be remembered that within this research the aim of statistical modeling, and the exploration of the potential benefits of Bayesianism, is to facilitate personalised realistic medicine through better shared clinical decision making by finding new ways of engaging with complexity, including uncertainty, not to attempt to "solve" these issues. This includes using FUPS data pertaining to the management of pancreatic cancer to open up conversations on findings rather than treating them as definitive facts (Wolpert & Rutter, 2018). In this way findings will be used to consider the complex reality they relate to but cannot fully capture so that narrative arguments and hypothesis can be contested and debated within the complex system (Wolpert & Rutter, 2018). In practical terms for this research this means instantiating the key principles of analysing FUPS data within the statistical models developed. Firstly all data is treated as a partial remnant with findings presented to convey associated limitations to interpretation. Secondly 'black box' statistical modeling will be avoided in favour of transparency and clarity. Thirdly triangulation will be used to

contextualise findings from models based on FUPS data to explore how other information and modeling techniques refute or support these findings (Wolpert & Rutter, 2018).

Chapter 4

Results

Introduction

The purpose of this research is to view the management pathways for pancreatic cancer through the lens of complexity theory and in so doing develop and expand the application of statistical modeling techniques to engage with the complexity in order to uncover new insight and move the research narrative towards the goal of more personalised realistic medicine.

As outlined in previous chapters the existing data fulfills the FUPS data criteria. Therefore the results are presented following the key principles for analysing FUPS data as proposed by Wolpert & Rutter (2018). This means that all reported results are presented as partial remnants with the limitations of interpretations stemming from FUPS characteristics clearly conveyed (Wolpert & Rutter, 2018). This is done not only through subjective assessment of the quality and risk-of-bias assessment of any included data, but also through an exploration of statistical techniques in quantifying such an assessment. Secondly all statistical analysis follows the principles of transparency and clarity. Thirdly all results are considered within the context of other existing information to explore what supports and undermines emerging findings (Wolpert & Rutter, 2018). This means that through the principle of triangulation findings from models

populated with the West of Scotland Pancreatic Unit database are triangulated with outcomes when the model is populated with internationally published data and vice versa. Each modeling approach is then discussed in terms of emergence, boundary setting, lack of complete knowledge and responsibility before moving on to assess how statistical modeling could be taken forward to gain further insights. This allows for a further layer of triangulation between outcomes form different statistical modeling approaches.

The rest of the chapter is structured as follows. Section 4.1 focuses on meta-analysis of existing studies comparing neoadjuvant and upfront surgery approaches for the treatment of potentially resectable pancreatic cancer. This section begins with a Bayesian network meta-analysis to assess overall resection, R0 resection and survival outcomes between neoadjuvant and upfront surgery pathways. This approach allows a synthesisation of phase II trials and observational studies comparing neoadjuvant and upfront surgery, as well as RCTs comparing upfront surgery and surgery alone. This means that an indirect comparison can be offered between neoadjuvant outcomes and surgery only outcomes. This also offers triangulation of findings between only including neoadjuvant phase II trials and then additionally including observational comparison studies. Furthermore triangulation of outcomes between the inclusion of studies exploring neoadjuvant therapy in all potentially resectable pancreatic cancer versus studies including resectable only cases is also offered. Bayesian network meta-analysis also offeres a more detailed quantification of the limitations of the analysis due to FUPS characteristics, which improves the transparency of the analyses.

Overall the results of the Bayesian network meta-analysis again suggested that although a marginal benefit was found with neoadjuvant approach neither pathway could be considered to be conclusively superior. The possibility still remained that optimal pathway selection could still depend on individualised factors.

Section 4.2 focused on Markov decision-analysis comparing upfront surgery and neoadjuvant pathways. Firstly the pathways are modeled to include the real-world scenario whereby borderline and locally advanced cases are treated within the neoadjuvant pathway. A like-for-like comparison for resectable only cases treated within upfront surgery and neoadjuvant pathways is then performed. Through deterministic and probabilistic sensitivity analysis of these Markov models not only is a transparent assessment of the degree of model uncertainty offered, but this allows the models to engage with the complexity of the system being examined. The result is that new insights into optimal treatment pathway selection begin to emerge. The findings corroborate those of the Bayesian network meta-analysis in section 4.1 that suggest a marginal overall survival advantage with neoadjuvant therapy, but this analysis goes further. Specifically probability thresholds for obtaining multimodal treatment in either pathway emerge as determining the superior treatment pathway. This further challenges the current narrative focusing on trying to prove whether upfront surgery or neoadjuvant pathway is superior for all patients and moves the research narrative towards a more personalised approach. The results of the Markov model were then triangulated by populating the model with patient

data from the West of Scotland Pancreatic Unit database, which corroborated the findings of the model using synthesised data.

Markov modeling provided some important insights but methodological issues surrounding its memory-less property and attrition of susceptible on a cycle-to-cycle basis mandated the further triangulation of these findings with another modeling technique that focused on micro simulations and modeling data at individualised patient level as opposed to at cohort level. Section 4.3 and 4.4 therefore centered on the use of Discrete Event Simulation (DES) modeling for decision analysis. Again synthesised data was triangulated against patient level data and discrete and probabilistic sensitivity analysis transparently quantified the degree of uncertainty raised by the FUPS characteristics of the data populating the model. This form of modeling corroborated many of the findings from the section on Markov modeling but was able to uncover further new insights by assigning more individualised data distributions to patients within the model depending on their disease stage at presentation. Specifically DES modeling had the flexibility to simulate the results of emerging RCTs into 'real-world' scenarios where not all context was controlled for when delivering pancreatic management pathways. This produced new insights into individual thresholds that determined how and to what extent reported findings from RCTs could be expected to apply to individual patients in a system where the complexity was not controlled. A further emerging insight was that for patients who did not progress to surgery within the neoadjuvant pathway, their corresponding

maximum benefit from being treated within the upfront surgery pathway was less than 5months prior to quality adjustment of survival time. This adds a further dimension to the debate regarding the criticism that neoadjuvant approach could result in losing the window of resectability and further challenges the narrative that resection is the only potential cure when for many the reality is that resection may be of limited benefit. Overall Markov and DES modeling approaches corroborate the emerging need to focus on personalised treatment pathway selection to optimise individual patient outcomes. Section 4.4 offers a further assessment by triangulating the outcomes of both Markov and DES modeling approaches by comparing their accuracy against the actual patient outcomes contained within the West of Scotland Pancreatic Unit database. This additional analysis raises the possibility that whilst Markov modeling is a more established technique for cost-effectiveness analysis, DES modeling could actually increase the accuracy of such models.

The Markov and DES models using both synthesised and actual patient data were then used separately to perform cost-effectiveness analysis of the competing treatment pathways in Section 4.5. Once again by using complexity theory as the lens through which to focus this research new insights began to emerge regarding the cost-effectiveness analysis of upfront surgery and neoadjuvant pathways. Uncertainty surrounding discounting rates of costs and benefits as well as debate over WtP thresholds were transparently incorporated into the analysis and both deterministic and probabilistic sensitivity analysis not only assessed the degree of model uncertainties

including the impact of altering costs, but also enabled an exploration of the impact of boundary setting on model outcomes. Specifically the inclusion of costs of palliative care and follow-up were important additions to the analysis. Once again, rather than simplistically concluding that one pathway was more cost-effective this analysis suggested that cost-effective delivery of treatment actually lay in better patient selection at individual patient level.

Although the insights gained through Markov and DES modeling have provided new insights and attempted to engage with complexity to change the research narrative from a "which pathway is superior for all" narrative towards a more personalised approach to patient selection, the question still remained as to how this could be achieved on a practical level. In section 4.5 lessons are drawn form the application of Bayesian statistics in other disciplines that use FUPS data within high-risk, complex adaptive systems that contain multiple potential points of risk of failure. These lessons are combined with the Bayesian analysis of the West of Scotland Pancreatic Unit database presented in appendix O that focuses on identifying individual patient factors that could determine whether patients with potentially resectable disease are likely to have a good or poor post resection prognosis. A prognostic Bayesian network is created and validated against the West of Scotland Pancreatic Unit database that makes individualised predictions of outcomes pre-operatively, across competing treatment pathways, and also performs prognostic updating at the post-operative phase of the patient journey.

4.1 Bayesian Network Meta-Analysis

Publications resulting from this analysis:

Bradley, A. and Van Der Meer, R. (2019). 'Upfront surgery versus neoadjuvant therapy for resectable pancreatic cancer: systematic review and Bayesian network meta-analysis'. *Nature Scientific Reports*, 9(1):4354. doi:10.1038/s41598-019-40951-6

Bradley, A., Van Der Meer, R., McKay, C.J. (2020) 'Bayesian network meta-analysis of upfront surgery versus neoadjuvant therapy for potentially resectable pancreatic ductal adenocarcinoma'. *British Journal of Surgery*: accepted

Abstract

Background: Current treatment recommendations for resectable pancreatic cancer support upfront surgical resection and adjuvant therapy. RCTs offering comparison with the emerging neoadjuvant (NAT) approach are lacking. This review aims to compare both treatment strategies first for potentially resectable pancreatic cancer and then separately for only disease that is resectable at presentation.

Methods: PubMed, MEDLINE, Embase, Cochrane Database and Cochrane Databases were searched for studies comparing neoadjuvant therapy and upfront surgery with adjuvant therapy pathways for potentially resectable pancreatic cancer. A Bayesian

network meta-analysis was conducted using the Markov chain Monte Carlo method. Cochrane Collaboration's risk-of-bias, ROBINS-I and GRADE tools were used to assess quality and risk-of-bias of included trials. Convergence was assessed using the Brooks-Gelman-Rubin method. In accordance with the NICE decision-support recommendations inconsistency was measured by comparing deviance residuals and deviance information criteria statistic in fitted consistency and inconsistency models.

Results: 25 studies comparing neoadjuvant and upfront surgery approaches (n=32,921), and 5 studies comparing upfront surgery plus adjuvant therapy and surgery only (n=899) were included. Aggregate rate (AR) of R0 resection was marginally higher, but not statistically significant according to 95% Credible Intrvals (CI), with neoadjuvant therapy (0.7389 versus 0.7306, Odds Ratio (O.R) 1.12, 95% CI 0.60-2.08). AR of 1,2,3,4 and 5-year survival were higher with neoadjuvant therapy (1-year survival: 0.8109 versus 0.6403, O.R: 2.12, 95% CI: 1.59-2.93; 2-year survival: 0.5135 versus 0.3002, O.R: 1.65 95%, CI: 1.16-2.34; 3-year survival: 0.3151 versus 0.2147, O.R: 1.50, 95% CI: 1.10-2.04; 4-year survival: 0.2114 versus 0.1647 O.R: 1.57, 95% CI: 0.80-2.99; 5-year survival: 0.2118 versus 0.1736, O.R: 1.65, 95% CI: 0.68-3.73).

For cases of pancreatic cancer that were resectable at presentation 9 studies compared neoadjuvant therapy and upfront surgery with adjuvant therapy (n=22,285). AR of R0 resection for neoadjuvant therapy was 0.8008 (0.3636-0.9144) *versus* 0.7515 (0.2026-0.8611), O.R. 1.27(95% CI 0.60-1.96). 1-year survival AR for neoadjuvant

therapy was 0.7969 (0.6061-0.9500) *versus* 0.7481 (0.4848-0.8500) O.R. 1.38 (95% CI 0.69-2.96). 2-year survival AR for neoadjuvant therapy was 0.5178 (0.3000-0.5970) *versus* 0.5131 (0.2727-0.5346) O.R. 1.26 (95% CI 0.94-1.74). 5-year AR survival for neoadjuvant therapy was 0.2069 (0.0323-0.3300) *versus* 0.1783 (0.0606-0.2300) O.R. 1.19 (95% CI 0.65-1.73).

Conclusion: Neoadjuvant therapy may offer benefit over surgery-first and adjuvant therapy for some patients. However, further RCTs are needed in collaboration with research developing methods of engaging with system complexity as multimodal treatment in either pathway is not obtained by all patients yet is a pivotal factor in achieving optimal patient outcomes.

Introduction

Pancreatic cancer is the fourth and fifth most common cause of cancer deaths in the USA and Europe respectively (Ferlay *et al*., 2013; Siegel *et al.*, 2015). Despite advances in surgical technique and adjuvant treatment, survival rates remain poor (Ferlay *et al*., 2013; Siegel *et al.*, 2015). Early complete surgical resection is the only potentially curative treatment and adjuvant therapy has been proven to prolong survival leading to surgery first with adjuvant therapy becoming the standard of care for resectable pancreatic cancer (Neoptolemos *et al*., 2001). However in reality most patients develop early recurrence, nullifying the potential benefits of high-risk surgery (Winter *et al*., 2012) with up to 50% of patients failing to receive

adjuvant therapy due to: post-operative complications, early metastases, reduced performance status and comorbidities (Bilimoria *et al*., 2007a). This has resulted in the advent of neoadjuvant therapy with the postulated benefits of: identifying more aggressive tumours hence avoiding futile surgery, elimination of micrometastesis, increased feasibility of R0 resection and completion of multimodal treatment (Asare *et al*., 2016; Lee *et al*., 2016).

Neoadjuvant therapy for resectable pancreatic cancer is an area of prime controversy and ongoing debate with a lack of large prospective RCTs offering direct comparison with upfront surgery and adjuvant therapy pathway (Tempero *et al*., 2014). Existing comparison studies often include borderline resectable and locally advanced cases in the neoadjuvant arm hence they do not offer a true like-for-like comparison. Ambiguity surrounding the existing body of research has led critics to highlight the limitations of drawing optimistic conclusions from small studies that are underpowered and caution against losing the window of resectability (Asare *et al*., 2016; Lee *et al*., 2016). Previous Markov decision analysis studies have reported slight survival benefit with neoadjuvant therapy but they only focused on a base-case intention-to-treat comparative analysis (Sharma *et al*., 2015; de Geus *et al*., 2016; Van Houten *et al*., 2012).

In the clinical setting the role of neoadjuvant therapy has widely been accepted for the management of locally advanced and borderline resectable cases of pancreatic cancer to increase the likelihood of achieving resection, particularly R0 resection (Asare *et al*., 2016; Lee

et al., 2016; Tempero *et al*., 2014; de Geus *et al*., 2016). However, ambiguities in the existing body of research concerning the management of resectable pancreatic cancer create a dilemma in clinical decision making. It has been established that optimal survival outcomes are not obtained by resection alone, but require the delivery of additional treatment whether delivered as neoadjuvant or adjuvant therapy (Neoptolemos *et al*., 2001; Xu *et al.*, 2014; Andriulli *et al*., 2012; Sharma *et al*., 2015; de Felice *et al*., 2014; de Geus *et al*., 2016; Versteijne *et al*., 2018; Van Houten *et al*., 2012). Both treatment pathways carry the risk of failing to achieve multimodal treatment delivery. Upfront surgery pathway carries the risk of failing to receive adjuvant therapy despite having undergone surgery with its associated risks of morbidity and mortality (Winter *et al*., 2012; Bilimoria *et al*., 2007a). Neoadjuvant approach also carries the risk of disease that was initially resectable at presentation progressing to become unresectable which makes its role in the management of resectable pancreatic cancer controversial (Asare et al., 2016; Lee et al., 2016). The question therefore arises as to whether neoadjuvant pathway represents a less superior treatment approach, or if it has the advantage of identifying aggressive tumour types that would have resulted in early disease recurrence precluding adjuvant therapy, being identified prior to patients undergoing high-risk, costly yet futile surgery (Asare et al., 2016; Lee et al., 2016). The aim of this meta-analysis is to compare upfront surgery and neoadjuvant approach for the management of potentially resectable pancreatic cancer and then separately for resectable pancreatic cancer on an intention-to-treat basis.

Treatment outcomes include: R0 resection rates and 1,2,3,4 and 5-year survival.

Methods

This review followed the PRISMA checklist (Moher *et al*., 2009). The protocols for this review and analysis are published on the PROSPERO online database of systematic reviews (CRD42018108676 and CRD42018108673). A search was undertaken using MEDLINE, Embase, PubMed and Cochrane database. For each of the searches the entire database was included since 2000 up to and including 31st August 2018, with no further date restrictions or limits applied.

Search Strategy

After removal of duplicates, manual screening was carried out based on the title and abstract of articles identified in the database searches. Articles of probable or possible relevance to this review based on the title and abstract were reviewed in full. Following screening, reference lists and citations of all included papers were manually searched to identify any additional articles. This process was repeated until no new articles were identified.

Inclusion Criteria and Outcomes

RCTs and prospective phase II/III studies offering comparison of neoadjuvant therapy *versus* upfront surgery plus adjuvant therapy for pancreatic cancer, published in English language since 2000,

involving chemo/radiotherapy-naive human subjects over 18 years of age with pancreatic cancer preoperatively staged as being potentially resectable (i.e. resectable, borderline resectable and locally advanced) were included. RCTs comparing upfront surgery plus adjuvant therapy and surgery only and cohort studies comparing neoadjuvant therapy and upfront surgery plus adjuvant therapy, with the same participant inclusion criteria, were included for separate sensitivity-analysis. Included trials had to report: protocol design, treatment regimes, number per arm, median age and co-morbidities of subjects, pre-treatment disease staging, outcome of post neoadjuvant therapy re-staging, surgical outcomes including resection rates, R0 resection rates and survival time. Case series and case reports, studies from identical patient cohorts, trials involving intra-operative radiotherapy and trials including disease other than pancreatic cancer were excluded. For the analysis of resectable pancreatic cancer the same inclusion and exclusion criteria was applied but studies had to include only preoperatively staged resectable pancreatic cancer, or report outcomes for resectable pancreatic cancer separately.

Data Collection

The following data was extracted from each study: study details (country, year, design, number of participants, mean age, sex, co-morbidity profile and presenting disease stage of participants in each arm), details of treatment protocols, treatment outcomes (rates of tumour resection, R0 resection rates, overall survival and disease

free survival and 1,2,3,4 and 5 year survival rates) and risk-of-bias data.

Statistical Analysis

Transparency of Analysis

This study conducted base-case analysis on an intention-to-treat basis. Patients who dropped out, or who failed to receive multimodal treatment within either pathway in the included studies were included in the overall and disease free survival analysis. The number of patients in the neoadjuvant pathway who failed to undergo resection, and the number of patients who underwent surgery but failed to receive adjuvant therapy, were analysed using weighted pooled estimates of proportions calculated using Freeman-Tukey arcsine square root transformation under random effects model to account for heterogeneity (Freeman & Tukey, 1950).

For each outcome of interest, NetMetXl was used to draw a weighted network for all treatments assessed for the specific outcomes that accounted for the study population size of each included study (Brown *et al*., 2014; Brown *et al*., 2018; Chaimani *et al*., 2013). This ensured that larger studies carried a greater weight within the network. A Bayesian network meta-analysis was conducted using the Markov chain Monte Carlo method in WinBUGS 1.4.3 (MRC Biostatistics Unit, Cambridge, and Imperial College School of Medicine, London, UK). To account for the inherent heterogeneity as a result of the different chemotherapy regimes, variations in multimodal treatment completion rates and differences in reported

survival outcomes, analysis was run using a random effects model, in addition to a fixed effects model, using vague priors as outlined in National Institute of Clinical Excellence (NICE) Evidence Synthesis Series (Brown *et al*., 2014; Dias *et al*., 2013a). Pairwise comparisons between interventions were also summarised to provide ranking of impact of intervention on each outcome based on the surface under the cumulative ranking (SUCRA) and were summarised in rankograms (Brown *et al*., 2014).

Assessment of Limitations of Interpretation Stemming from FUPS Characteristics of Data

To further minimise the impact of heterogeneity of different chemotherapy combinations, treatment completion rates and reported survival analysis on the overall analysis, convergence was assessed using the Brooks-Gelman-Rubin method and by checking whether the Monte Carlo error is less than 5% of the standard deviation of the effect estimates and between-study variance (Brown *et al*., 2014). The Markov chain Monte Carlo (MCMC) Bayesian network meta-analysis was fitted with three chains as a means of checking MCMC convergence (Brown *et al*., 2014). The Brooks-Gelman-Rubin method compares within-chain and between-chain variances to calculate the potential scale reduction factor with a value close to one indicating when approximate convergence is reached (Brown *et al*., 2014; Brooks & Gelman, 1998).

Inconsistency assessment, the conflict between direct and indirect evidence, is crucial to any network meta-analysis (Dias *et al.*, 2013b).

In accordance with the NICE decision-support documents (Spiegelhalter *et al.*, 2002) inconsistency was measured by comparing deviance residuals and deviance information criteria (DIC) statistic in fitted consistency and inconsistency models (Brown *et al*., 2014; Dias *et al*., 2013b). Posterior mean deviance of the individual data points in the inconsistency model were plotted against their posterior mean deviance in the consistency model to identify any loops in the treatment network where inconsistency is present (Brown *et al*., 2014).

The Cochrane Collaboration's risk-of-bias tool (Higgins *et al*., 2011) and Risk Of Bias In Non-randomized Studies - of Interventions (ROBINS-I tool) (Sterne *et al*., 2016) were also used to assess the quality of included studies. Grading of Recommendations Assessment Development and Evaluation (GRADE) tool was used to provide additional assessment of quality of evidence and rate certainty in estimates from the network meta-analysis (Shunemann *et al*., 2018; Brignardellu-Petersen *et al*., 2018).

Triangulation

Sensitivity network meta-analyses that included cohort studies for neoadjuvant therapy *versus* upfront surgery plus adjuvant therapy and RCTs of upfront surgery plus adjuvant therapy *versus* surgery only were also performed with the latter offering an indirect comparison between neoadjuvant therapy and surgery only.

Results: Bayesian network meta-analysis of treatment options for potentially resectable pancreatic cancer

Eligible Studies

A total of 14375 studies were identified through a search of the electronic databases. 452 studies underwent full text review. 25 studies were identified that offered comparison between neoadjuvant therapy and upfront surgery plus adjuvant therapy (Figure 4). Nine of these studies were phase II/III trials, 3 of which were randomised. 16 cohort studies comparing neoadjuvant therapy and upfront surgery plus adjuvant therapy were also included in a separate network for sensitivity analysis. 6 studies were prospective and 10 studies were retrospective (Table 13).

Figure 4: PRISMA Flow Chart for Neoadjuvant Therapy versus Upfront Surgery plus Adjuvant Therapy studies

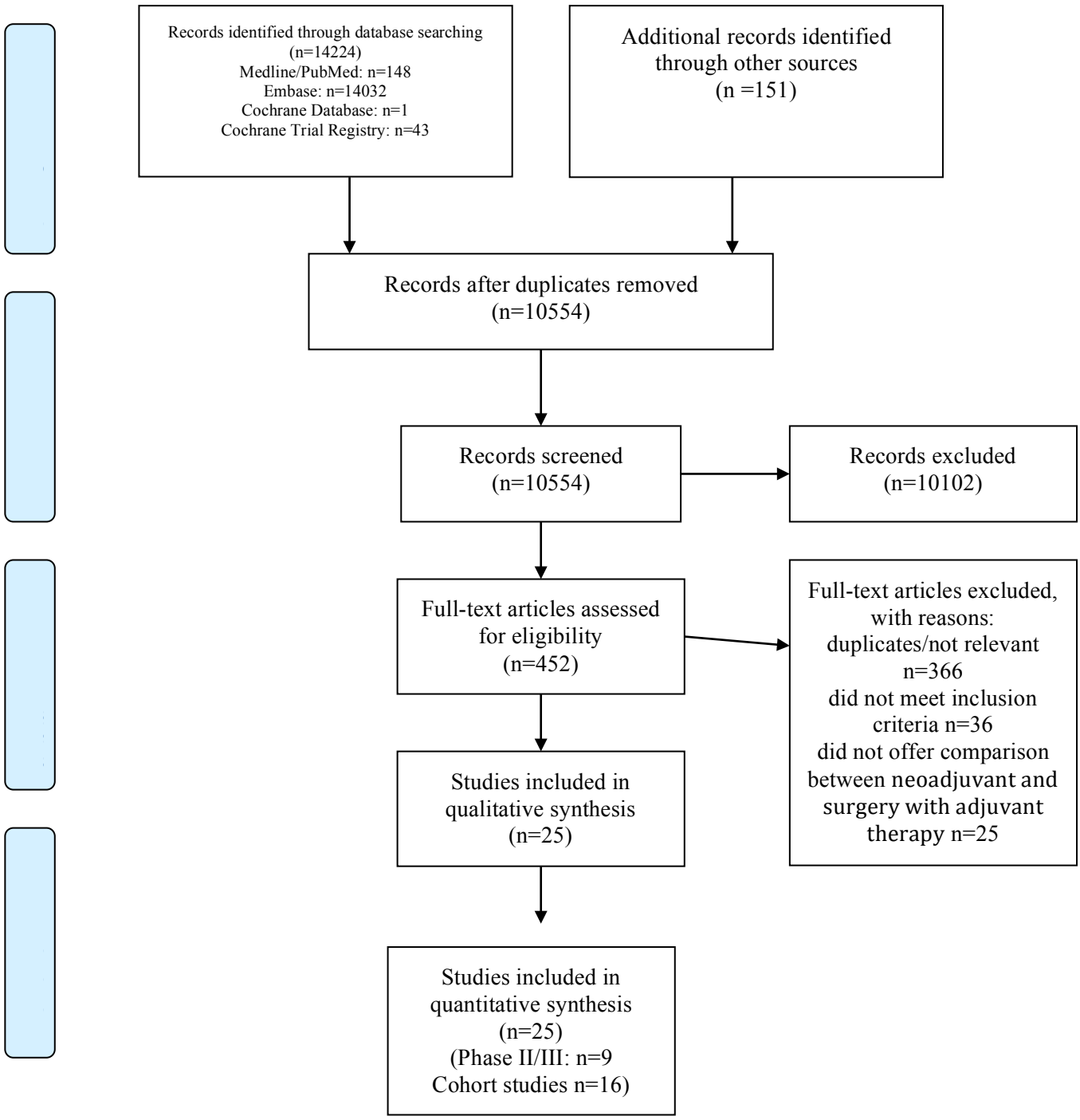


Table 13: Summary of Included Trials of Neoadjuvant Therapy versus Upfront Surgery Plus Adjuvant Therapy in Bayesian Network Meta-analysis for Potentially Resectable Pancreatic Cancer.

Study	Centre	Randomisation	Type of Trial	No. Neoadjuvant arm	Neoadjuvant regime: CRT= chemoradiotherapy CT=chemotherapy NAT= no further details given about regime	Neoadjuvant arm overall survival in months	No. surgery first arm	Surgery first arm overall survival in months
Al-Sukhum *et al.*, 2003	Single	No	Phase II	20	CRT	13.4	21	18.1
Casadei *et al.,* 2015	Single	Yes	Phase II	18	CRT	28.3	20	27.5
Golcher *et al.*, 2008	Single	No	Phase II	121	CRT		58	21
Golcher *et al.*, 2015	Single	Yes	Phase II	33	CRT	17.4	33	14.4
Lind *et al.*, 2008	Multiple	No	Phase II	17	CRT	19	35	11
Massucco *et al.*, 2006	Single	No	Phase II	28	CRT	15.4	44	14
Satoi *et al.*, 2009	Single	No	Phase II	35	CRT	24.5	41	18.5
Vento *et al.*, 2007	Single	No	Phase II	22	CRT	30.2	25	35.9
Jang *et al.,* 2018	Single	Yes	Phase II/III	27	CRT	21	23	12
deGus *et al.*, 2017a	Multiple	No	Retrospective	1077	NAT	25.9	6840	24.2
Mellon *et al.*, 2016	Multiple	No	Retrospective	159	CRT	17	241	22.1
Nurmi *et al.,* 2018	Single	No	Retrospective	75	CRT/CT	35	150	26
Shubert *et al.*, 2016	Multiple	No	Retrospective	377	NAT	20.7	216	13
Artinya *et al.*, 2011	Multiple	No	Retrospective	39	NAT	33.8	419	19
Ielop *et al.*, 2016	Multiple	No	Prospective	45	CRT	21.65	36	22.1
Roland *et al.*, 2015	Single	No	Prospective	222	CRT		85	
deGus *et al.*, 2017b	Single	No	Retrospective	1541	NAT	Resectable: 26.2 Borderline: 23.5 Locally Advanced: 23.5	11316	Resectable: 24.5 Borderline: 20.0 Locally advanced: 15.5
Mokdad *et al.*, 2017	Multiple	No	Retrospective	2005	NAT	26	6015	21
Chen *et al.*, 2017	Multiple	No	Retrospective	98	NAT	25	98	17
Tzeng *et al.*, 2014	Single	No	Prospective	115	NAT	28	52	25.3
Fujii *et al.*, 2015	Single	No	Prospective	21	CRT	29.1	71	13.1
Fujii *et al.*, 2017	Single	No	Prospective	88	CRT	Resectable: 24.9 Borderline: 28.4	416	Resectable: 23.5 Borderline: 20.1
Papalezova *et al.*, 2012	Single	No	Retrospective	144	CRT	15	92	13
Hirono *et al.*, 2016	Single	No	Prospective	46	CRT	19.3	124	13.7
Murkakami *et al.*, 2017	Single	No	Retrospective	52	CT	27.1	25	11.6

For further sensitivity analysis, RCTs offering comparison between upfront surgery plus adjuvant therapy and surgery only were also

included in a separate network meta-analysis (Figure 5). Five RCTs offered comparison between upfront surgery plus adjuvant therapy and surgery only and were included in the sensitivity analysis (table 14).

Figure 5: PRISMA Flow Chart for Randomised Controlled Trials of Upfront Surgery plus Adjuvant Therapy versus Surgery Only

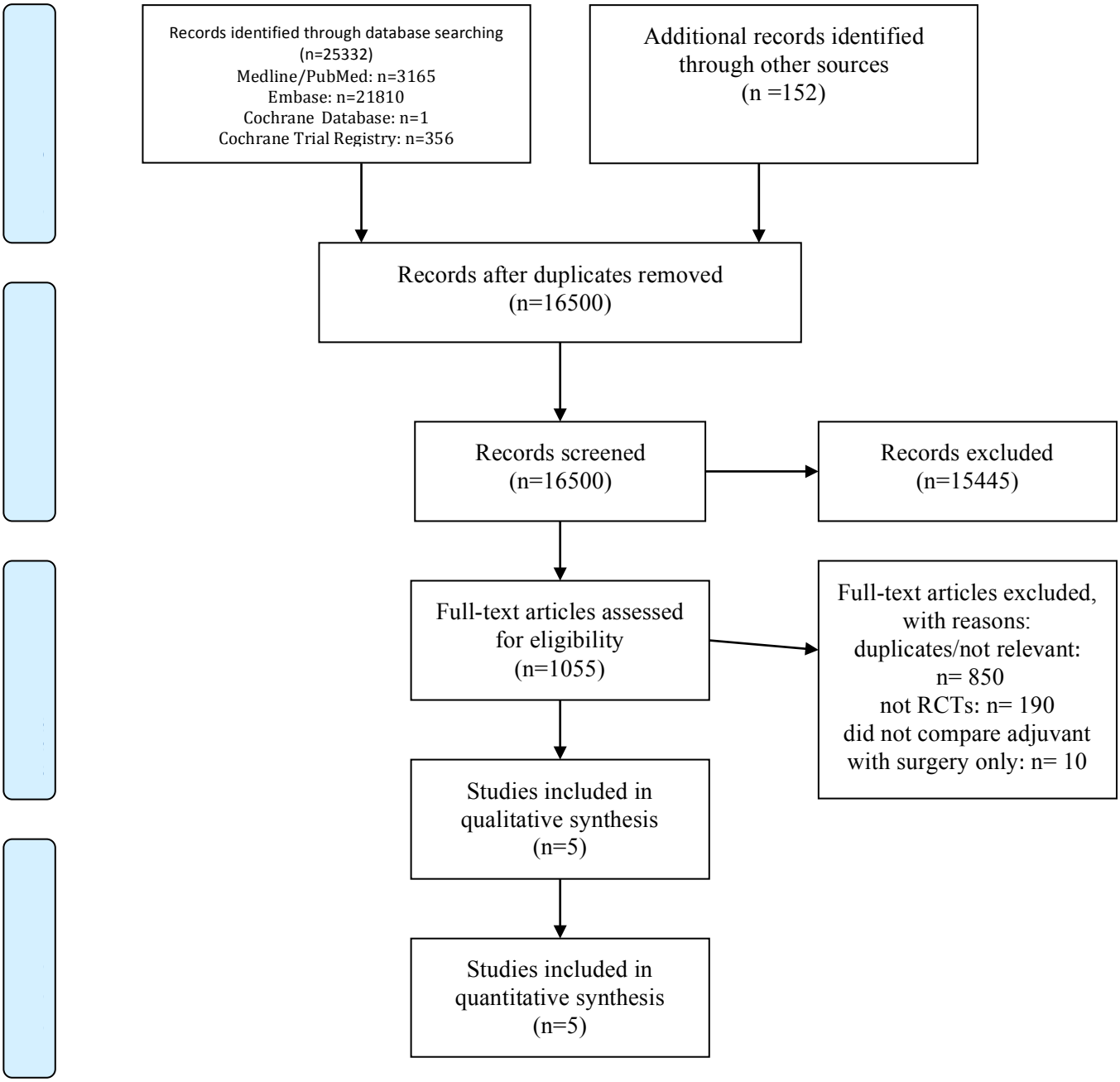


Table 14: Summary of Randomised Controlled Trials comparing Upfront Surgery plus Adjuvant Therapy *versus* Surgery Only

Study	No. in Upfront surgery and adjuvant therapy arm	Adjuvant Regime	Upfront surgery and adjuvant therapy Overall Survival in months	No. Surgery Only arm	Surgery Only Overall Survival in months
Ueno *et al.*, 2009	58	Gemcitabine	22.3	50	18.4
Oettle *et al.*, 2013	179	Gemcitabine	22.8	175	20.2
Kosuge *et al.*, 2006	45	Cisplatin + 5-FU	12.5	44	15.8
Smeenk *et al.*, 2007	110	5-FU +radiotherapy	21.6	108	19.2
Morak *et al.*, 2008	59	5-FU+folic acid+ mitocantrone + cisplatinur + radiotherapy	19	61	18

A summary of overall findings for each outcome measure is provided in Figures 6 and 7.

Figure 6: Summary of Findings of Network Meta-analysis of Phase II/III trials comparing Neoadjuvant Therapy *versus* Upfront Surgery and Adjuvant Therapy

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	№ of participants (studies)
	Risk with Surgery First + adjuvant therapy	Risk with Neoadjuvant		
Resection Rate	877 per 1,000	**332 per 1,000** (221 to 439)	**OR 0.07** (0.04 to 0.11)	621 (9 observational studies)
R0 resection	509 per 1,000	**293 per 1,000** (225 to 371)	**OR 0.40** (0.28 to 0.57)	580 (8 observational studies)
1-year survival	614 per 1,000	**771 per 1,000** (674 to 847)	**OR 2.12** (1.30 to 3.49)	332 (6 observational studies)
2-year survival	347 per 1,000	**413 per 1,000** (317 to 512)	**OR 1.32** (0.87 to 1.97)	460 (6 observational studies)
3-year survival	253 per 1,000	**253 per 1,000** (155 to 386)	**OR 1.00** (0.54 to 1.86)	241 (4 observational studies)
4-year survival	86 per 1,000	**128 per 1,000** (35 to 361)	**OR 1.56** (0.39 to 6.01)	118 (2 observational studies)
5-year survival	116 per 1,000	**246 per 1,000** (126 to 437)	**OR 2.50** (1.10 to 5.95)	189 (3 observational studies)

*The risk in the intervention group (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI). 'Risk' is the risk of the event occurring i.e 'risk' of being alive at the set time interval.

Figure 7: Summary of Findings of Network Meta-analysis of Neoadjuvant Therapy *versus* Upfront Surgery and Adjuvant Therapy with inclusion of cohort studies

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	№ of participants (studies)
	Risk with SFadj Phas II/III+pro/retro	Risk with Neoadjuvant		
Resection Rate	884 per 1,000	**550 per 1,000** (498 to 605)	**OR 0.16** (0.13 to 0.20)	2941 (18 observational studies)
R0 reaction rate	731 per 1,000	**783 per 1,000** (769 to 797)	**OR 1.33** (1.23 to 1.45)	19369 (21 observational studies)
1-year survival	640 per 1,000	**808 per 1,000** (796 to 819)	**OR 2.37** (2.19 to 2.55)	32094 (21 observational studies)
2-year survival	300 per 1,000	**525 per 1,000** (507 to 544)	**OR 2.58** (2.40 to 2.78)	24313 (20 observational studies)
3-year survival	215 per 1,000	**306 per 1,000** (292 to 319)	**OR 1.61** (1.51 to 1.71)	31926 (18 observational studies)
4-year survival	165 per 1,000	**238 per 1,000** (184 to 304)	**OR 1.58** (1.14 to 2.21)	1483 (8 observational studies)
5-year survival	174 per 1,000	**217 per 1,000** (199 to 237)	**OR 1.32** (1.18 to 1.48)	9956 (12 observational studies)

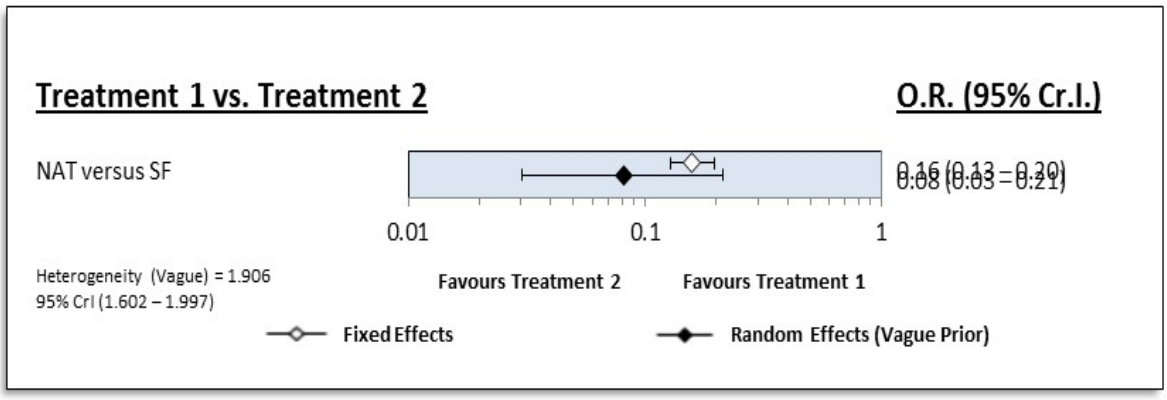
* The risk in the intervention group (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI). 'Risk' is the risk of the event occurring i.e 'risk' of being alive at the set time interval.

Resection Rates

Pairwise comparison of surgical resection rates was based on a network constructed from 9 phase II/III studies (n=621; neoadjuvant therapy: n=321; upfront surgery plus adjuvant therapy n=300). Upfront surgery plus adjuvant therapy approach was found to be superior with aggregate rate 0.8767 (0.6970-1.000) *versus* 0.3489 (0.1500-0.6818). Fixed and random effects models supported this finding with Odds Ratio (O.R) 0.07 (95% CI: 0.04-0.11) and O.R 0.04 (95% CI 0.01-0.15) respectively (Appendix K).

Inclusion of cohort studies created a network analysis based on 18 studies (neoadjuvant therapy: n=1368, upfront surgery plus adjuvant therapy n=1573) and did not alter overall outcome. Aggregate rate was superior for upfront surgery plus adjuvant therapy 0.8843 (0.6970-1.0000) *versus* 0.6060 (0.1500-0.9038). Both fixed and random effects models supported this finding with O.R 0.16 (95% CI: 0.13-0.20) and O.R. 0.08 (95% CI 0.03-0.21) respectively (Figure 8; Appendix K).

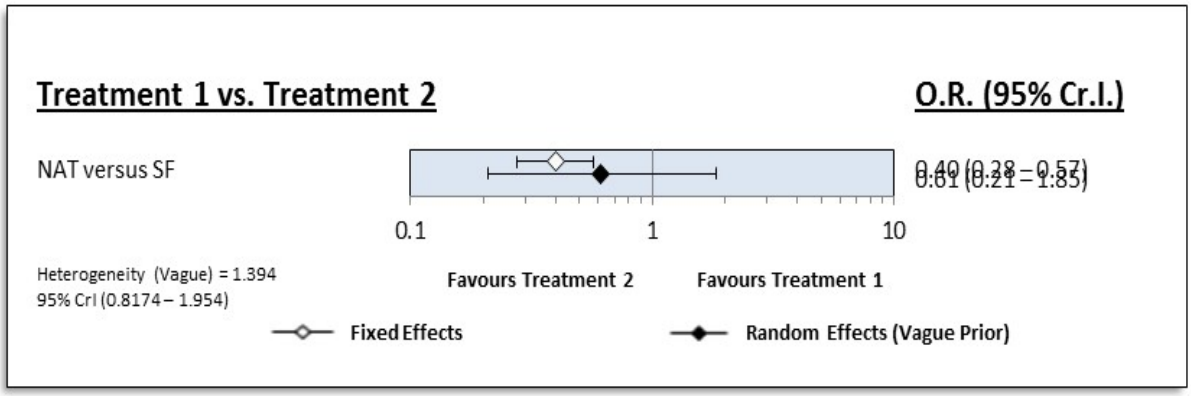
Figure 8: Results of fixed effects and random effects (vague prior) models



R0 Resection Rates

Network analysis based on 8 phase II/III trials (neoadjuvant therapy n=301; upfront surgery plus adjuvant therapy n=279) gave R0 aggregate rate 0.5090 (0.1707-0.7759) for upfront surgery plus adjuvant therapy *versus* 0.2957 (0.1570-0.5185). Fixed and random effects models favoured upfront surgery plus adjuvant therapy (O.R 0.40; 95% CI 0.28-0.57 and O.R 0.61; 95% 0.21-1.85 respectively) (Figure 9; Appendix K).

Figure 9: Results of fixed effects and random effects (vague prior) models



When cohort studies were included in the network (21 studies; neoadjuvant therapy n= 4727; upfront surgery plus adjuvant therapy n=14642) neoadjuvant therapy was superior with aggregate rate 0.7389 (0.1570-0.9144) *versus* 0.7306 (0.1600-0.8611). Both fixed and random effects models favoured neoadjuvant therapy (O.R 1.33; 95% CI 1.22-1.45 and O.R 1.12; 95% CI 0.60-2.08 respectively) (Appendix K).

1-year Survival

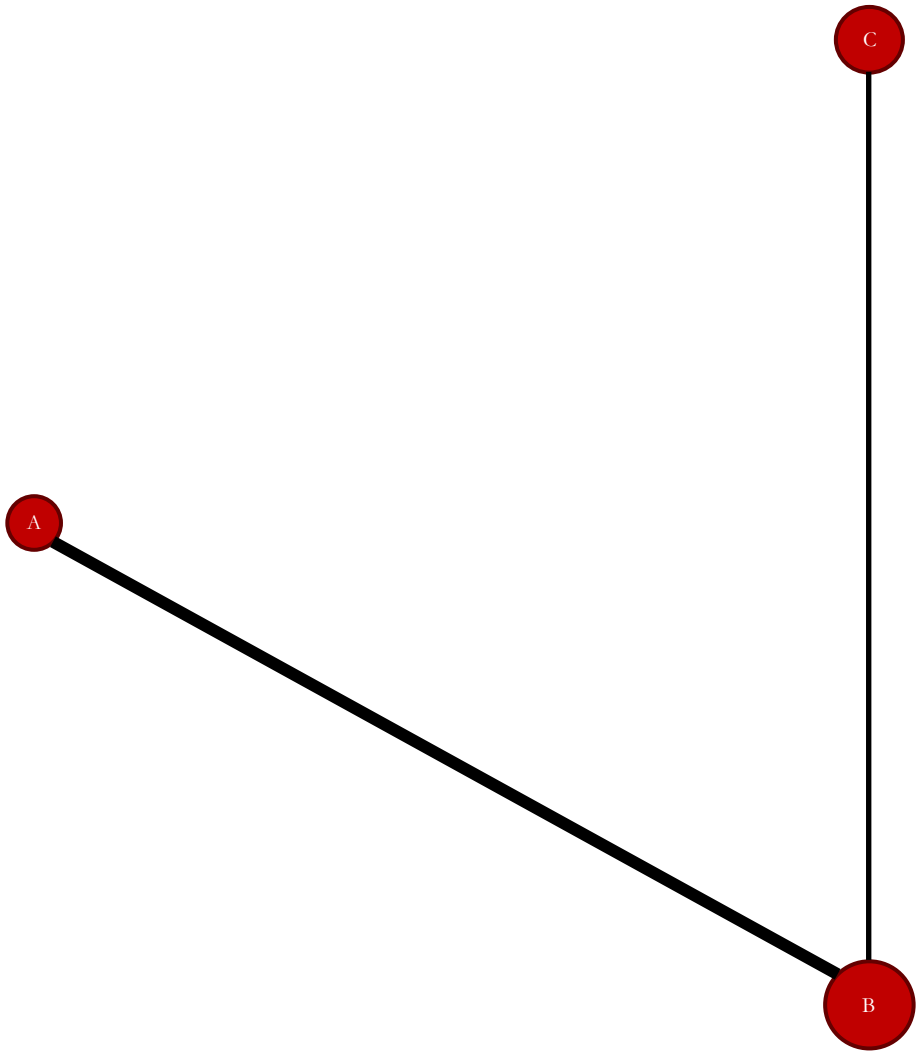
Pairwise comparison of neoadjuvant therapy *versus* upfront surgery plus adjuvant therapy through network meta-analysis based on 6 phase II/III trials (neoadjuvant therapy n= 154; upfront surgery plus adjuvant therapy n=178) favoured neoadjuvant therapy with aggregate rate 0.7466 (0.5200-1.0000) *versus* 0.6137 (0.4778-0.7200). Fixed and random effects models favoured neoadjuvant

therapy (O.R 2.12; 95%CI 1.30-3.49 and O.R 2.26; 95% CI 0.90-7.23 respectively) (Appendix K).

With inclusion of cohort studies in the network (21 studies; neoadjuvant therapy n=5988; upfront surgery plus adjuvant therapy n=26106) results did not alter with aggregate rate 0.8109 (0.5200-1.0000) *versus* 0.6403 (0.4400-0.8500). Fixed and random effects models continued to favour neoadjuvant therapy (O.R 2.37; 95% CI 2.19-2.55 and O.R. 2.12; 95% CI 1.59-2.93 respectively) (Appendix K).

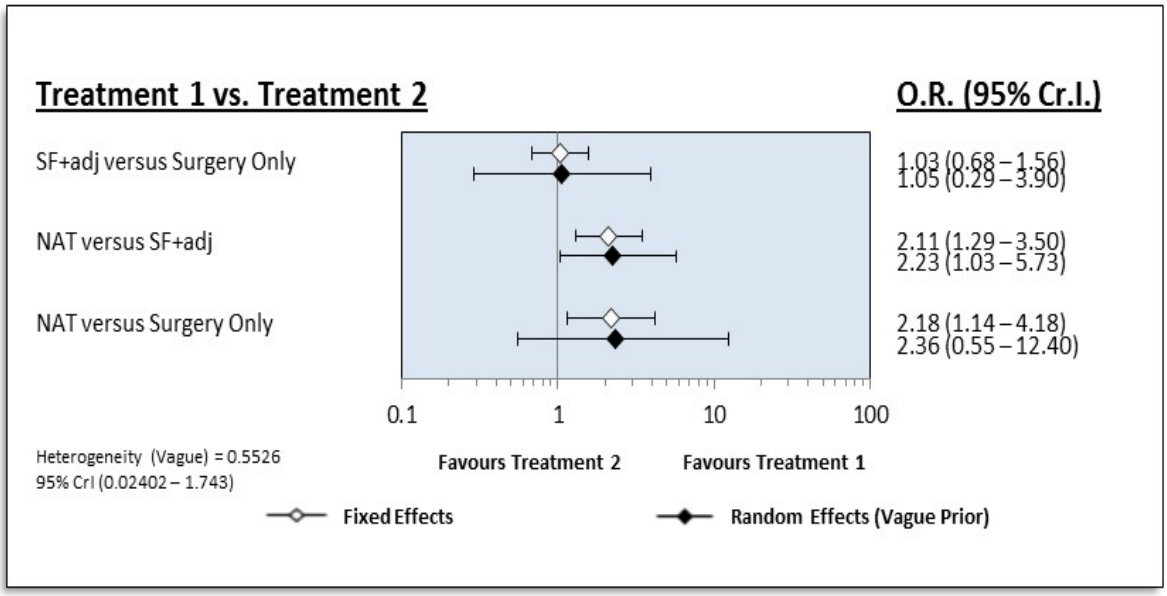
A sensitivity network analysis including phase II/III trials comparing neoadjuvant therapy and upfront surgery plus adjuvant therapy and RCTs comparing upfront surgery plus adjuvant therapy and surgery only (8 studies; neoadjuvant therapy n=154; upfront surgery plus adjuvant therapy n=415, surgery only n=235) (Figure 10) favoured neoadjuvant therapy with aggregate rate 0.7466 (0.5200-1.0000) *versus* 0.7314 (0.7250-0.7500) for surgery only and 0.6845 (0.4778-0.7760) for upfront surgery plus adjuvant therapy. Neoadjuvant therapy was found to be superior in both fixed and random effects models (Figure 11; Appendix K).

Figure 10: Bayesian Network Meta-analysis of Neoadjuvant therapy versus Upfront surgery plus adjuvant therapy versus surgery only



Drug	Abbreviation
NAT	A
SF+adj	B
Surgery Only	C

Figure 11: Results of fixed effects and random effects (vague prior) models



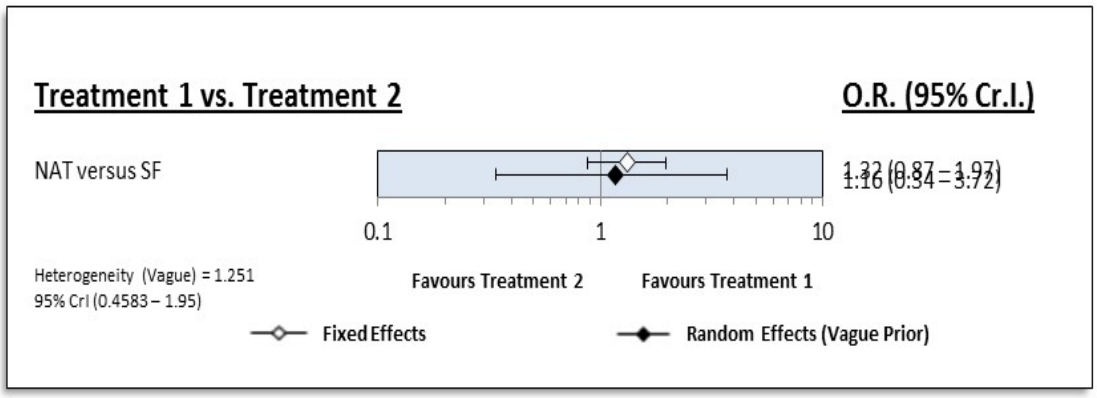
Cohort studies comparing neoadjuvant therapy and upfront surgery plus adjuvant therapy were then included in the sensitivity network analysis (23 studies; neoadjuvant therapy n=5988; upfront surgery plus adjuvant therapy n=26343; surgery only n=235). This supported superiority of neoadjuvant therapy with aggregate rate 0.8109 (0.5200-1.0000) *versus* 0.7314 (0.7250-0.7500) for surgery only and 0.6413 (0.4400-0.8500) for upfront surgery plus adjuvant therapy (Appendix K).

2-year Survival

Pairwise comparison of 2-year survival from 6 phase II/III trials (neoadjuvant therapy n=246; upfront surgery plus adjuvant therapy

n=214) showed superiority of neoadjuvant therapy with aggregate rate 0.4454 (0.0600-0.7500) *versus* 0.3475 (0.2609-0.4660). Fixed effects (O.R. 1.32; 95% CI 0.87-1.97) and random effects model (O.R 1.16; 95%CI 0.34-3.72) favoured neoadjuvant therapy (Figure 12; Appendix K).

Figure 12: Results of fixed effects and random effects (vague prior) models



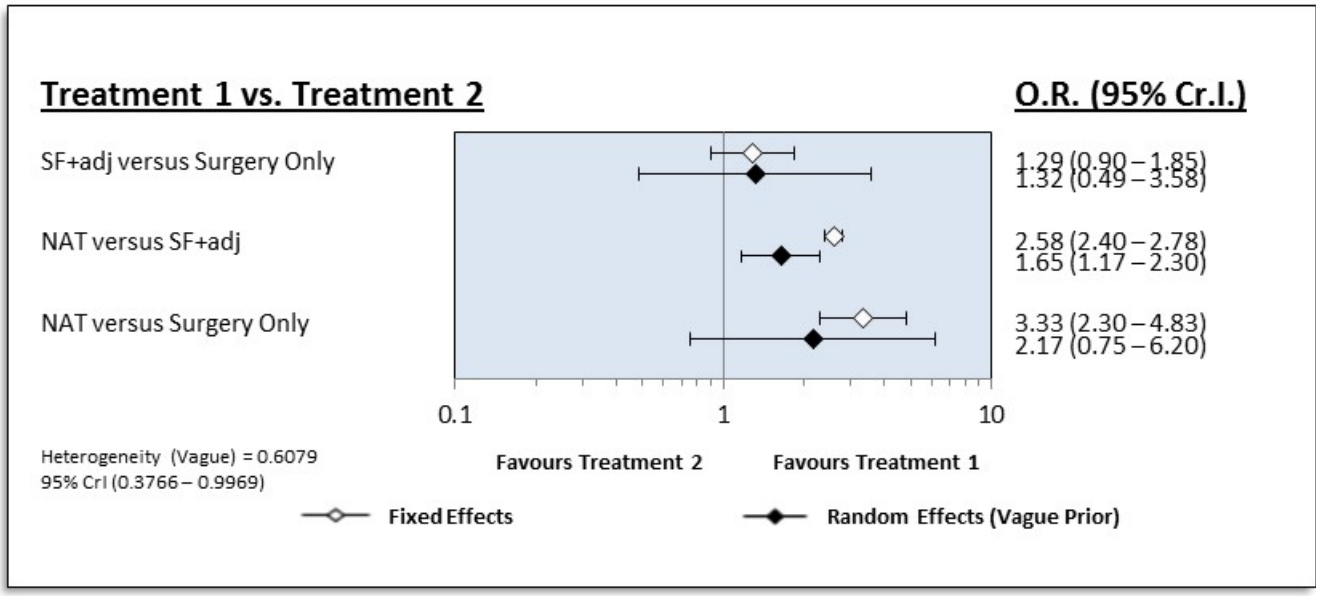
Inclusion of cohort studies within the network (20 studies; NAT n=4199; SFadj n=20114) produced corroborating results with NAT aggregate rate 0.5135 (0.0600-0.7500) *versus* 0.3002(0.1268-0.5800) and fixed and random effects models supporting NAT (O.R 2.58; 95% CI 2.40-2.78 and O.R. 1.65; 95%CI 1.16-2.34 respectively) (Appendix K).

Network sensitivity analysis including phase II/III trials comparing neoadjuvant therapy and upfront surgery plus adjuvant therapy (6 studies) and RCTs comparing upfront surgery plus adjuvant therapy and surgery only (2 studies) (neoadjuvant therapy n=246; upfront

surgery plus adjuvant therapy n=451; surgery only n=235) showed marginal superiority of neoadjuvant therapy with aggregate rate 0.4454 (0.0600-0.7500) *versus* 0.4155 (0.2609-0.4829) for upfront surgery plus adjuvant therapy and 0.4149 (0.4000-0.4200) for surgery only. Fixed and random effects models favoured neoadjuvant therapy (Appendix K).

Cohort studies were then included in network sensitivity analysis (22 studies; neoadjuvant therapy n= 4199; upfront surgery plus adjuvant therapy n=20351; surgery only n=235). Neadjuvant therapy remained superior with aggregate rate 0.5135 (0.0600-0.7500) but upfront surgery lus adjuvant therapy aggregate rate dropped to 0.3025 (0.1268-0.5800) with surgery only aggregate rate 0.4149 (0.4000-0.4200). Fixed and random effects models corroborated overall treatment ranking favouring neoadjuvant therapy (Figure 13; Appendix K).

Figure 13: Results of fixed effects and random effects (vague prior) models



3-year survival

Pairwise comparison within a network based on 4 phase II/III trials (neoadjuvant therapy n=107; upfront surgery plus adjuvant therapy n=134) marginally favoured neoadjuvant therapy but this was not statistically significant with aggregate rate 0.2642 (0.1212-0.3900) *versus* 0.2530 (0.1100-0.4700). Fixed and random effects models demonstrated no significant difference between treatment pathways (O.R. 1.00; 95% CI 0.54-1.86 and O.R 0.99; 95% CI 0.34-2.89 respectively) (Appendix K).

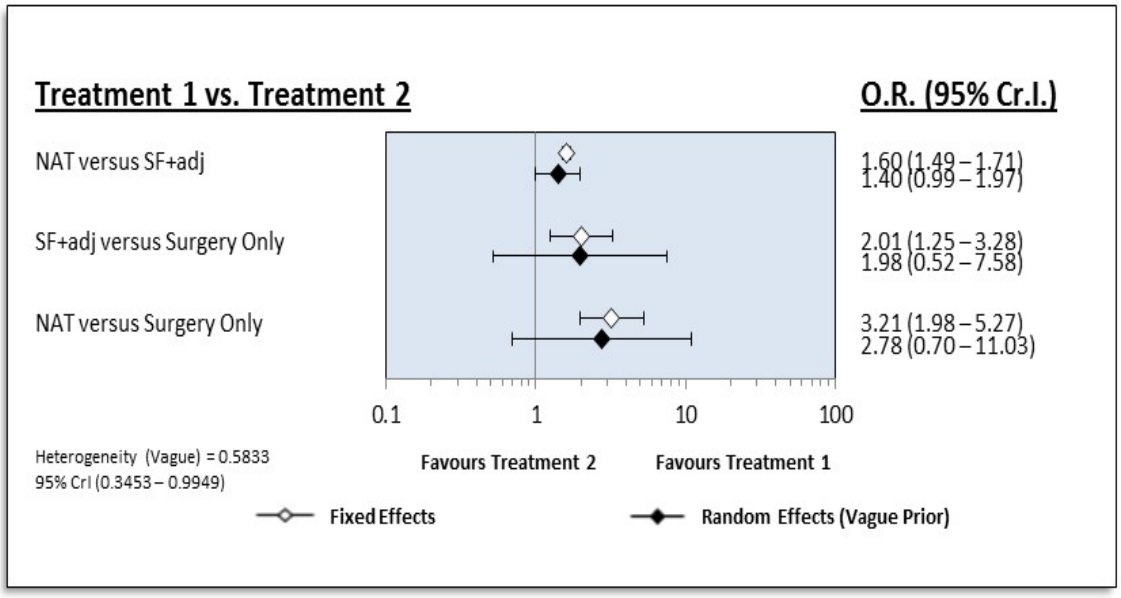
Inclusion of cohort studies increased the network to 18 studies (neoadjuvant therapy n= 5889; upfront surgery plus adjuvant therapy n=26037) and favoured neoadjuvant therapy with aggregate

rate 0.3151 (0.1212-0.4800) compared to 0.2147 (0.0563-0.4700). Fixed and random effects models favoured neoadjuvant therapy with O.R increasing to 1.61 (95%CI 1.51-1.72) and 1.50 (95% CI 1.10-2.04) respectively (Appendix K).

RCTs comparing upfront surgery plus adjuvant therapy and surgery only (1 study) were combined with the 4 phase II/III studies comparing neoadjuvant therapy and upfront surgery plus adjuvant therapy (neoadjuvant therapy n= 107; upfront surgery plus adjuvant therapy n=313; surgery only n= 175) in a network sensitivity analysis. Upfront surgery plus adjuvant therapy was superior with aggregate rate 0.3463 (0.1100-0.6329), followed by 0.2709 (0.1212-0.3900) for neoadjuvant therapy and 0.2050 (0.2050-0.2050) for surgery only with this ranking corroborated in fixed and random effects models (Appendix K).

Inclusion of cohort studies comparing neoadjuvant therapy and upfront surgery plus adjuvant therapy in further sensitivity analysis produced a network of 19 studies (neoadjuvant therapy n=58899; upfront surgery plus adjuvant therapy n=26216; surgery only n=175). Superiority altered from upfront surgery plus adjuvant therapy with aggregate rate 0.2160 (0.0563-0.6329) to neoadjuvant therapy with aggregate rate 0.3151 (0.1212-0.4800). Surgery only still held lowest ranking with aggregate rate 0.2050 (0.2050-0.2050). Ranking order was corroborated in fixed and random effects models (Figure 14; Appendix K).

Figure 14: Results of fixed effects and random effects (vague prior) models



4-year survival

Two phase II/III trials (n=118) reported data on 4-year survival (neoadjuvant therapy n=50; upfront surgery plus adjuvant therapy n=68). Neoadjuvant therapy was superior with aggregate rate 0.1016 (0.0303-0.2400) compared to 0.0860 (0.0606-0.1100) and O.R 1.56 (95% CI: 0.39-6.01) and 1.35(95%CI: 0.13-10.82) in fixed and random effects models respectively (Appendix K)

Inclusion of cohort studies increased the network to 8 studies (neoadjuvant therapy 414; upfront surgery plus adjuvant therapy n=1069). Neoadjuvant therapy maintained superiority with aggregate rate 0.2114 (0.0303-0.4000) *versus* 0.1647 (0.0423-0.3200) and O.R 1.59 (95% CI 1.14-2.21) and 1.57 (95%CI 0.80-2.99) in fixed and random effects models respectively. RCTs comparing

upfront surgery plus adjuvant therapy and surgery only did not report 4-year survival (Appendix K).

5-year survival

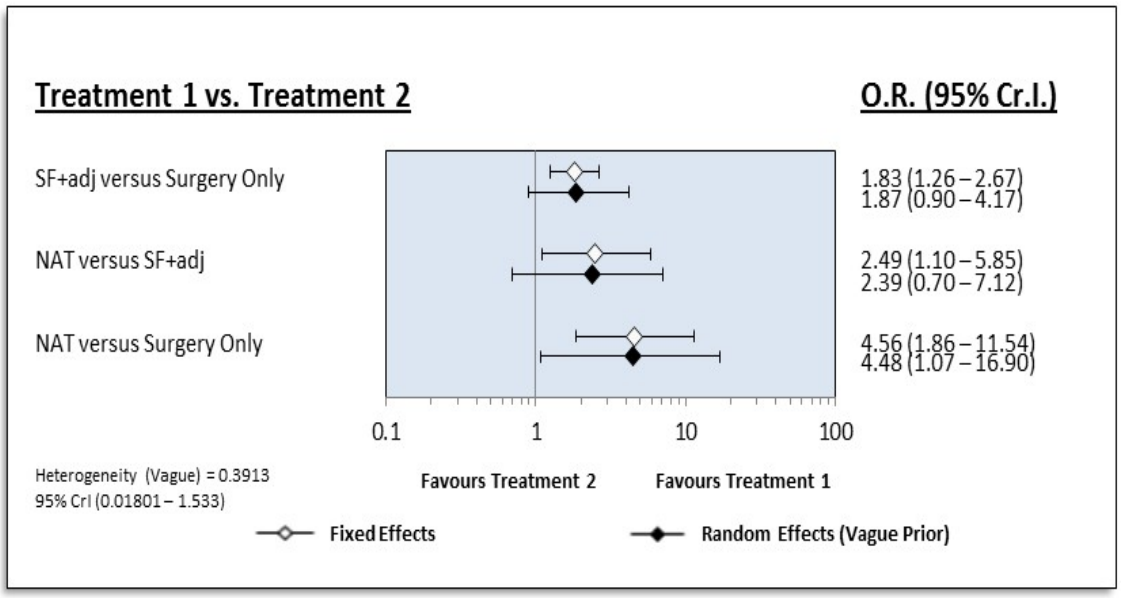
Three phase II/III trials (neoadjuvant therapy n=90; upfront surgery plus adjuvant therapy n=99) reported 5-year survival and were included in network meta-analysis. Neoadjuvant therapy held superiority with aggregate rate 0.2240 (0.0303-0.3400) *versus* 0.1156 (0.0606-0.2300) and O.R 2.50 (95%CI 1.10-5.95) and 2.20 (95%CI 0.38-10.06) in fixed and random effects models respectively (Appendix K).

Inclusion of cohort studies increased the network to 12 studies (neoadjuvant therapy n=2885; upfront suregry plus adjuvant therapy n=7071) and neoadjuvant therapy held superiority with aggregate rate 0.2118 (0.0303-0.7692) compared to 0.1736 (0.0500-0.2300) and O.R 1.32 (95% CI 1.18-1.48) and 1.65 (95% CI 0.68-3.73) in fixed and random effects models respectively (Appendix K).

RCTs comparing upfront surgery plus adjuvant therapy and surgery only (4 studies) and phase II/III trials comparing neoadjuvant therapy and upfront surgery pluc adjuvant therapy (3 studies) created a sensitivity network analysis of 7 studies (neoadjuvant therapy n=90; upfront surgery plus adjuvant therapy n=491; surgery only n=387). This did not alter the ranking of treatment pathways with neoadjuvant therapy aggrgate rate 0.2240 (0.0303-0.3400)

versus 0.2072 (0.0606-0.2640) for upfront surgery plus adjuvant therapy and 0.1418 (0.1040-0.2200) for surgery only. This was also the case in both fixed and random effects models (Figure 15; Appendix K).

Figure 15: Results of fixed effects and random effects (vague prior) models



Cohort sudies were then included increasing the sensitivity network analysis to 16 studies (neoadjuvant thaerpy n=2885; upfront surgery plus adjuvant therapy n=7463; surgery only n=387). This did not alter the ranking of treatments with aggregate rates for neoadjuvant therapy 0.2054 (0.0303-0.3400) *versus* 0.1779 (0.0500-0.3200) for upfront surergy plus adjuvant threapy and 0.1418 (0.1040-0.220) for surgery only (Appendix K).

Assessment of Impact of FUPS characteristics of Data

Using the GRADE assessment criteria the certainty of recommendations from the network analysis showed that although neoadjuvant therapy was marginally favoured overall, this was not always statistically signifinat and uncertainty was identified in the evidence synthesised (Figure 16).

Figure 16: GRADE assessment of certainty of network recommendations

CRITERIA	SUMMARY OF JUDGEMENTS				
PROBLEM	No	Probably no	Probably yes	**Yes**	
DESIRABLE EFFECTS	Trivial	Small	**Moderate**	Large	
UNDESIRABLE EFFECTS	Large	Moderate	**Small**	Trivial	
CERTAINTY OF EVIDENCE	Very low	**Low**	Moderate	High	
VALUES	Important uncertainty or variability	**Possibly important uncertainty or variability**	Probably no important uncertainty or variability	No important uncertainty or variability	
BALANCE OF EFFECTS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	**Favors the intervention**
ACCEPTABILITY	No	Probably no	**Probably yes**	Yes	
FEASIBILITY	No	Probably no	Probably yes	**Yes**	

This degree of uncertainty in the synthesised evidence was further highlighted in the assessment of risk-of-bias of each included study (Figure 17; Figure 18; Appendix K).

Figure 17: Overall Assessment of Risk of Bias of included trials of neoadjuvant versus upfront surgery plus adjuvant therapy

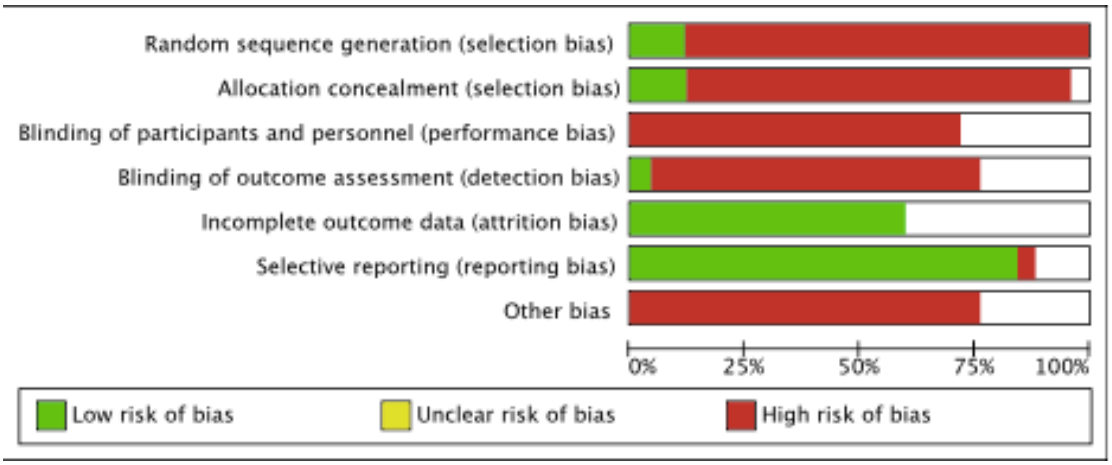
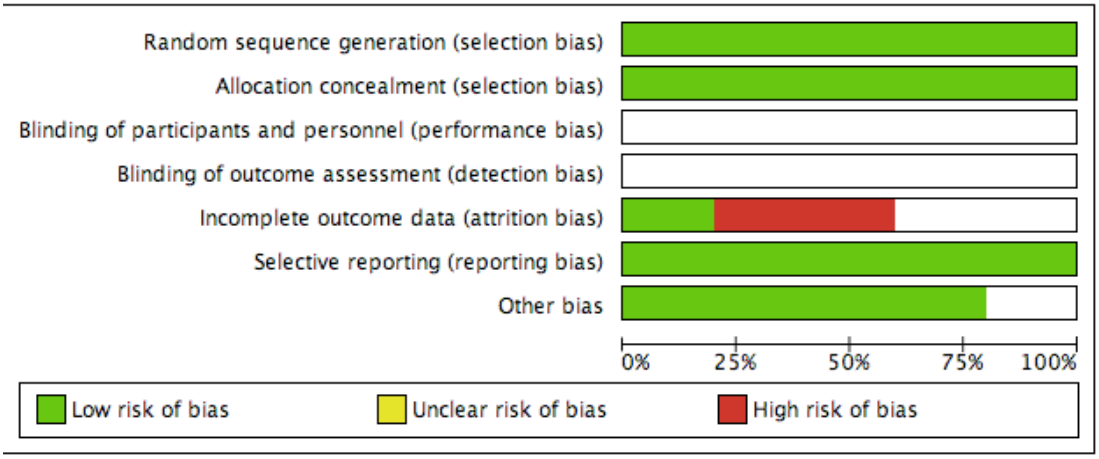


Figure 18: Overall Assessment of Risk of Bias of Randomised Controlled Trials comparing Upfront Surgery plus Adjuvant Therapy versus Surgery Only.



To further assess the impact of FUPS characteristics of the data consistency and inconsistency assessment was undertaken. Overall consistency was achieved across all Bayesian network meta-analysis

survival year models with no issues of inconsistency identified (Appendix K).

Results: Bayesian network meta-analysis of treatment options for resectable pancreatic cancer

Eligible Studies

Nine studies were identified that offered comparison between neoadjuvant therapy and upfront surgery plus adjuvant therapy for the treatment of resectable pancreatic cancer (Figure 19). As only 2 of these studies were phase II trials, one of which was randomised all studies were therefore included in the network meta-analysis. 4 studies were prospective and 3 studies were retrospective (Table 15).

Figure 19: PRISMA flow chart for neoadjuvant therapy versus upfront surgery plus adjuvant therapy

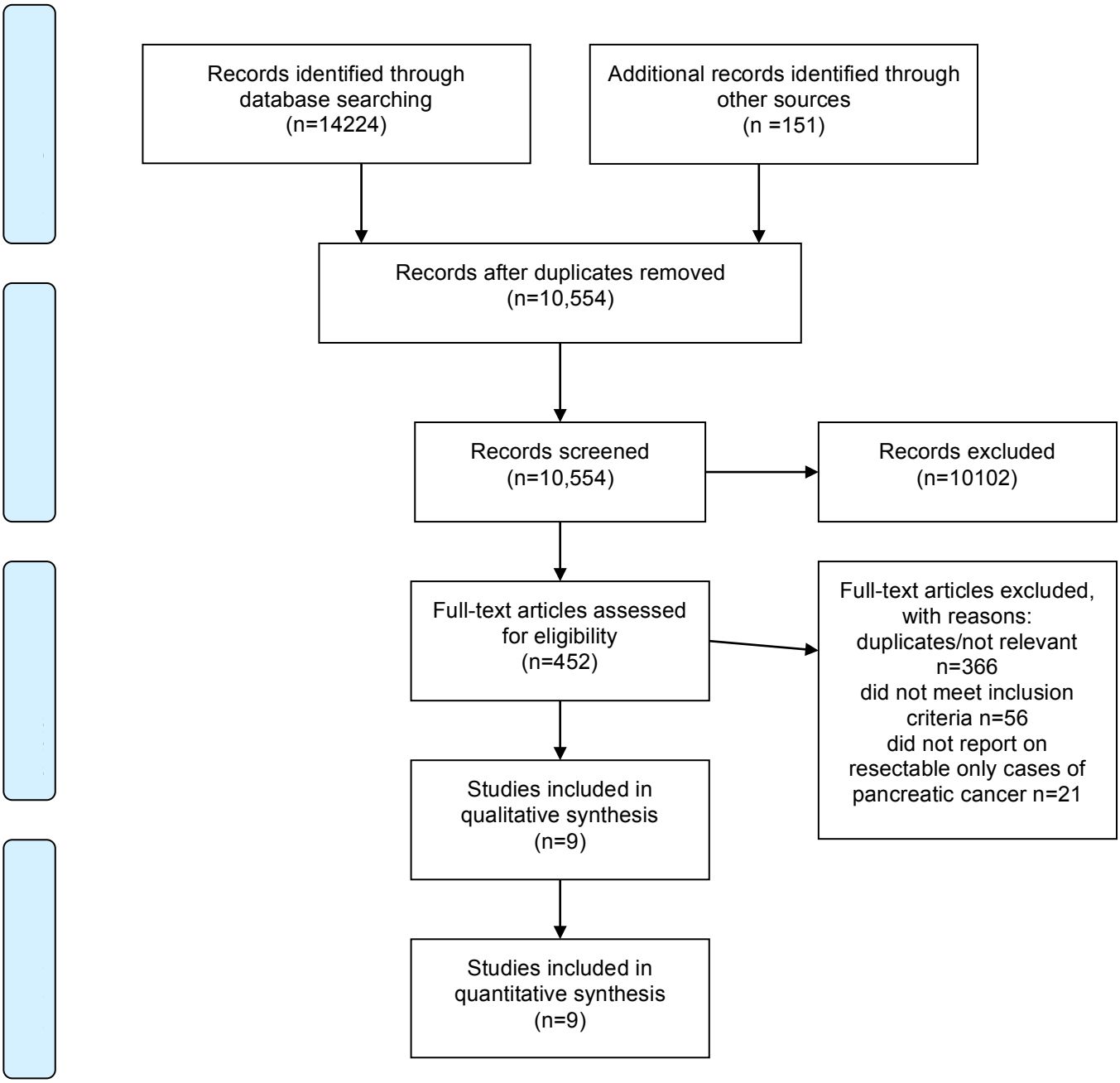


Table 15: Summary of Studies Comparing Neoadjuvant Therapy versus Upfront Surgery plus Adjuvant Therapy for Resectable Pancreatic Cancer.

Study	**Study Type**	**Randomised**	**Centre**	**Neoadjuvant treatment Regime in addition to radiotherapy**	**Total No. patient in Neoadjuvant arm**	**Neoadjuvant arm Overall Survival in months for RPC**	**Total No. patients Upfront surgery lus adjuvant therapy arm**	**Upfront surgery lus adjuvant therapy arm Overall Survival in months**	**ROBINS-I risk of bias assessment**
Golcher *et al.*, 2015	Phase II	Yes	Multiple	Gemcitabine/ cisplatin	31	17.4	33	14.4	Low
Vento *et al.*, 2007	Phase II	No	Single	Gemcitabine	22	30.2	25	35.9	Moderate
Ielpo *et al.*, 2017	Prospective	No	Single	Gemcitabine +Nabpaclitaxel	19	21.65	36	22.1	Moderate
Roland *et al.*, 2015	Prospective	No	Single	Gemcitabine, 5-FU or capecitabine	222		85		Moderate
DeGus *et al.*, 2017a	Retrospective	No	Multiple (cancer registry)	NAT: no further details given	332	26	11316	24.5	Moderate/Serious
Mokdad *et al.*, 2017	Retrospective	No	Multiple (cancer registry)	NAT: no further details given	2005	26	6015	21	Moderate/Serious
Tzeng *et al.*, 2014	Prospective	No	Single	NAT: no further details given	115	28	62	25.3	Moderate/Serious
Fujii *et al.*, 2016	Prospective	No	Single	S1+5-FU+oteracil and gimeracil	40	24	416	23	Moderate/Serious
Papalezova *et al.*, 2012	Retrospective	No	Single	5-FU	144	15	92	13	Moderate/Serious

6 studies (n=371) reported the number of cases of resectable pancreatic canncer who received neoadjuvant therapy and progressed to surgery (Golcher *et* al., 2015; Vento *et al*., 2007; Ielpo *et al*., 2017; Tzeng *et al*., 2014; Fuijii *et al*, 2016; Papalezova *et al*., 2012) giving a pooled proportion of 76.08% (95% Confidence Interval: 60.826-88.509). Two studies reported response to neoadjuvant therapy (Golcher *et al*., 2015; Ielpo *et al*., 2017). One

study reported responses for resectable cases (complete response: 0; partial response: 4/31; stable disease 8/31; disease progress 12/31; 7 unrecorded) (Golcher *et al*., 2015). The study by Ielop *et al*. (2017) did not report this outcome separately for resectable only cases but included borderline cases also in reporting the outcomes of response to neoadjuvant therapy (complete response: 5/45; partial response: 13/45; stable disease 5/45). 6 studies (n= 17596) reported the number of patients in the upfront surgery plus adjuvant therapy pathway who received adjuvant therapy (Ielpo *et al*., 2017; Roland *et al*., 2015; Tzeng *et al*., 2014; DeGus *et al*., 2017a; Mokdad *et al*., 2017; Papalezova *et al*., 2012) giving a pooled proportion of 63.01% (95% Confidence Interval: 59.452-66.489).

For sensitivity analysis, RCTs offering comparison between upfront surgery plus adjuvant therapy *versus* surgery alone were also included in a separate network meta-analysis. Electronic database search identified 25332 studies (Figure 20). 15 studies were RCTs, 5 of which offered comparison between adjuvant therapy and surgery alone and were included in the sensitivity analysis (Table 16).

Figure 20: PRISMA flow chart of upfront surgery plus adjuvant therapy versus surgery only

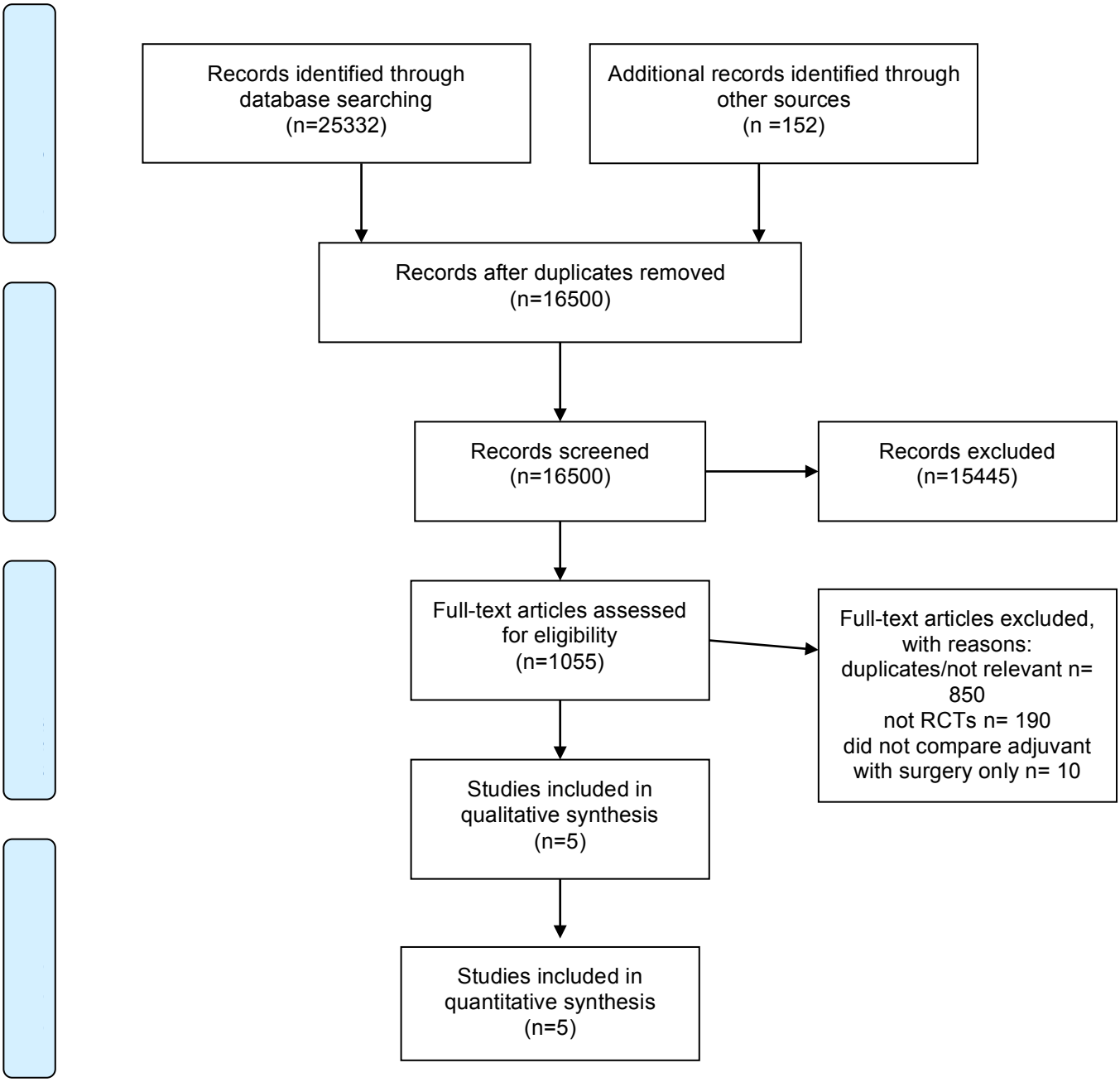


Table 16: Summary of included studies. Summary of randomised controlled trials comparing upfront and adjuvant therapy versus surgery only.

Study	Adjuvant Regime *CT=chemotherapy CRT=chemoradiotheapy	Adjuvant chemotherapy agents	No. Upfront surgery plus adjuvant arm	Overall survival in Upfront surgery plus adjuvant arm in months	No. Surgery Only arm	Overall survival in surgery only arm
Ueno *et al.*, 2009	CT	Gemcitabine	58	22.3	60	18.4
Oettle *et al.*, 2013	CT	Gemcitabine	179	22.8	175	20.2
Kosuge *et al.*, 2006	CT	Cisplatin + 5-FU	45	12.5	44	15.8
Smeenk *et al.*, 2007	CRT	5-FU	110	21.6	108	19.2
Morak *et al.*, 2008	CRT	5-FU+folic acid+ mitoxantrone + cisplatin	59	19	61	18

A summary of overall findings for each outcome measure is provided in Figure 21.

Figure 21: Summary of results of Bayesian network meta-analysis comparing upfront surgery plus adjuvant therapy with neoadjuvant therapy for the management of resectable pancreatic cancer.

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	№ of participants (studies)
	Risk with Upfront Surgery plus adjuvant therapy	Risk with Neoadjuvant		
R0 resection	751 per 1,000	**818 per 1,000** (800 to 836)	**OR 1.49** (1.32 to 1.68)	9197 (9 observational studies)
1-year survival	748 per 1,000	**813 per 1,000** (796 to 829)	**OR 1.46** (1.31 to 1.63)	12039 (8 observational studies)
2-year survival	513 per 1,000	**562 per 1,000** (518 to 606)	**OR 1.22** (1.02 to 1.46)	4279 (7 observational studies)
3-year survival	294 per 1,000	**343 per 1,000** (322 to 365)	**OR 1.25** (1.14 to 1.38)	12039 (8 observational studies)
4-year survival	127 per 1,000	**144 per 1,000** (91 to 220)	**OR 1.16** (0.69 to 1.94)	656 (4 observational studies)
5-year survival	178 per 1,000	**208 per 1,000** (188 to 229)	**OR 1.21** (1.07 to 1.37)	8896 (7 observational studies)

* The risk in the intervention group (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI). 'Risk' is the risk of the event occurring i.e 'risk' of being alive at the set time interval.

R0 Resection Rates

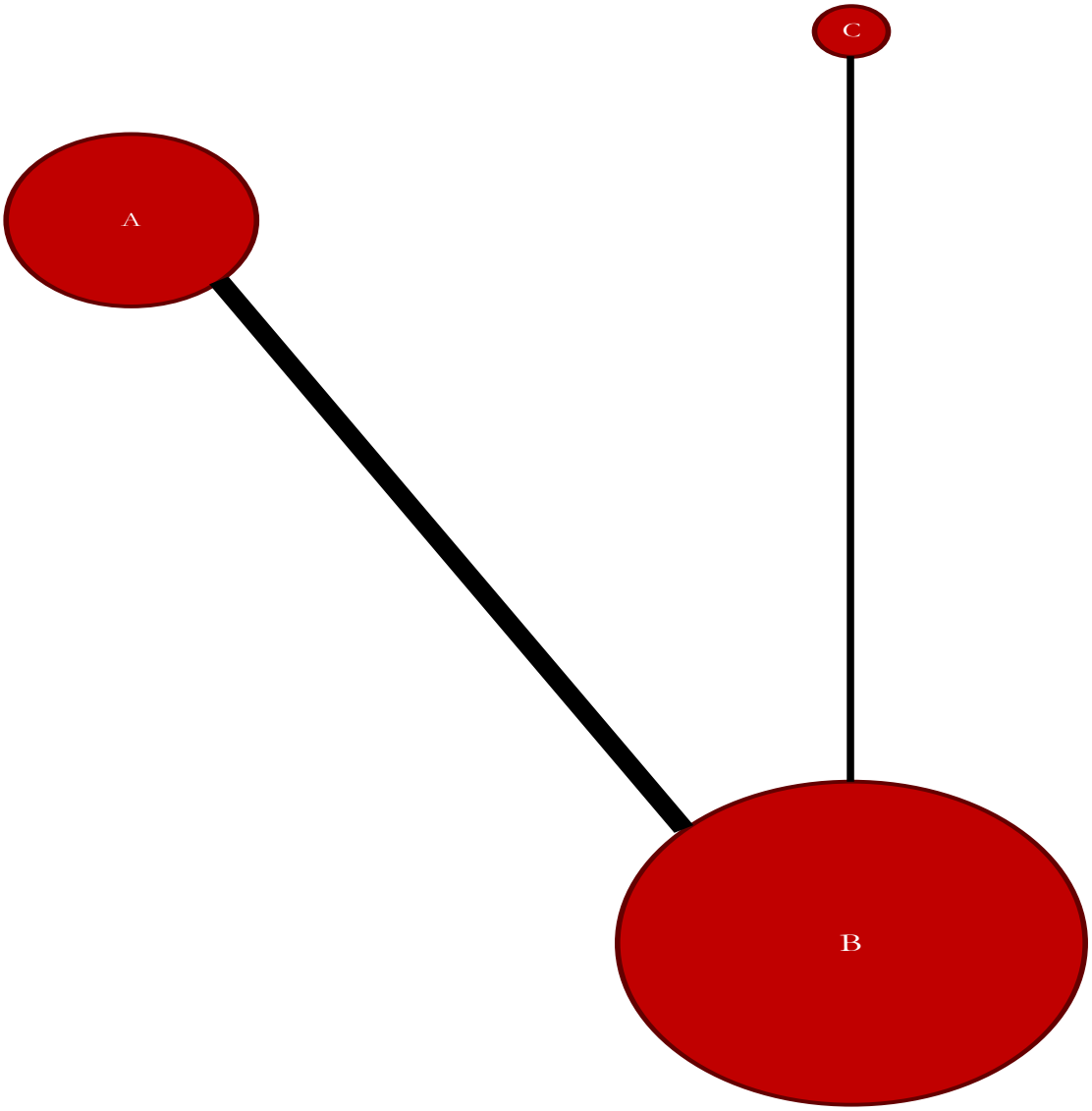
The network offering pairwise comparison of rates of R0 resection between neoadjuvant therapy and upfront surgery plus adjuvant therapy included 8 studies and 9197 participant (neoadjuvant therapy: n=2626; Upfront surgery plus adjuvant therapy: n=6571). The aggregate rate of R0 resection for neoadjuvant therapy was 0.8008 (0.3636-0.9144) compared to 0.7515 (0.2826-0.8611) for upfront surgery plus adjuvant therapy. Both fixed effects (O.R. 1.49; 95% CI 1.32-1.68) and random effects (O.R. 1.27; 95% CI 0.60-1.96) models favoured neoadjuvant therapy. Neoadjuvant therapy was found to have superior positive impact on outcome of R0 resection (SUCRA: 0.8124 *versus* 0.1876).

1-year Survival

Pairwise comparison for 1-year survival of neoadjuvant therapy versus upfront surgery plus adjuvant therapy was based on 8 studies and 12011 participants (neoadjuvant therapy: n=2708; upfront surgery plus adjuvant therapy: n=9303). Aggregate rate of 1-year survival was higher in neoadjuvant therapy at 0.7969 (0.6061-0.9500) *versus* 0.7481 (0.4848-0.8500). Both fixed effects (O.R. 1.46; 95% CI: 1.31-1.63) and random effects (O.R. 1.38 95%; CI: 0.69-2.96) models favoured neoadjuvant therapy. Neoadjuvant therapy also has a stronger positive impact on the outcome of 1-year survival (SUCRA: 0.84 v 0.16) (Appendix L).

For sensitivity analysis a network also including RCTs of upfront surgery plus adjuvant therapy *versus* surgery only was constructed based on a total of 10 studies and 12483 patients (neoadjuvant therapy: n=2708; upfront surgery plus adjuvant therapy: n=9540; surgery only: n=235) (Figure 22). 8 studies compared neoadjuvant therapy and upfront surgery plus adjuvant therapy (n=12011) and 2 studies compared upfront surgery plus adjuvant therapy and surgery only (n=472).

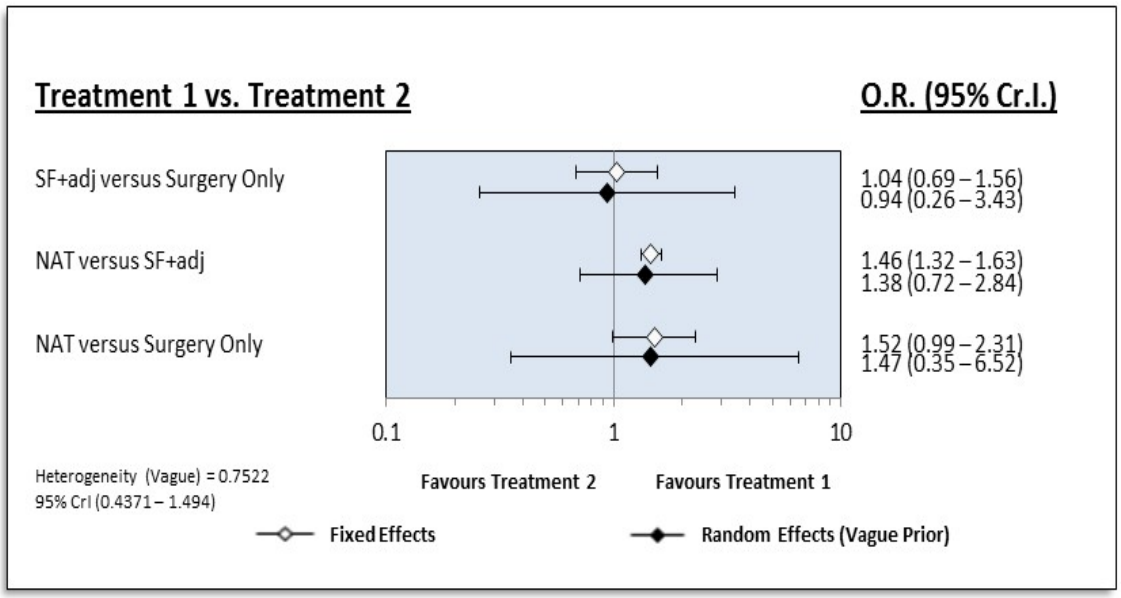
Figure 22: Bayesian Network Meta-analysis of Neoadjuvant therapy versus Upfront surgery plus adjuvant therapy versus surgery only



Drug	Abbreviation
NAT	A
SF+adj	B
Surgery Only	C

Neoadjuvant therapy was found to be superior in both fixed and random effects models. Aggregate rate of 1-year survival was highest in neoadjuvant therapy (0.7957; range 0.6205-0.9500) followed by upfront surgery plus adjuvant therapy (0.7478; range 0.4848-0.8500) then surgery only (0.7314; range 0.7250-0.7500). Again neoadjuvant therapy was found to have strongest positive impact on outcome of 1-year survival (Figure 23; Appendix L).

Figure 23: Results of fixed effects and random effects (vague prior) models



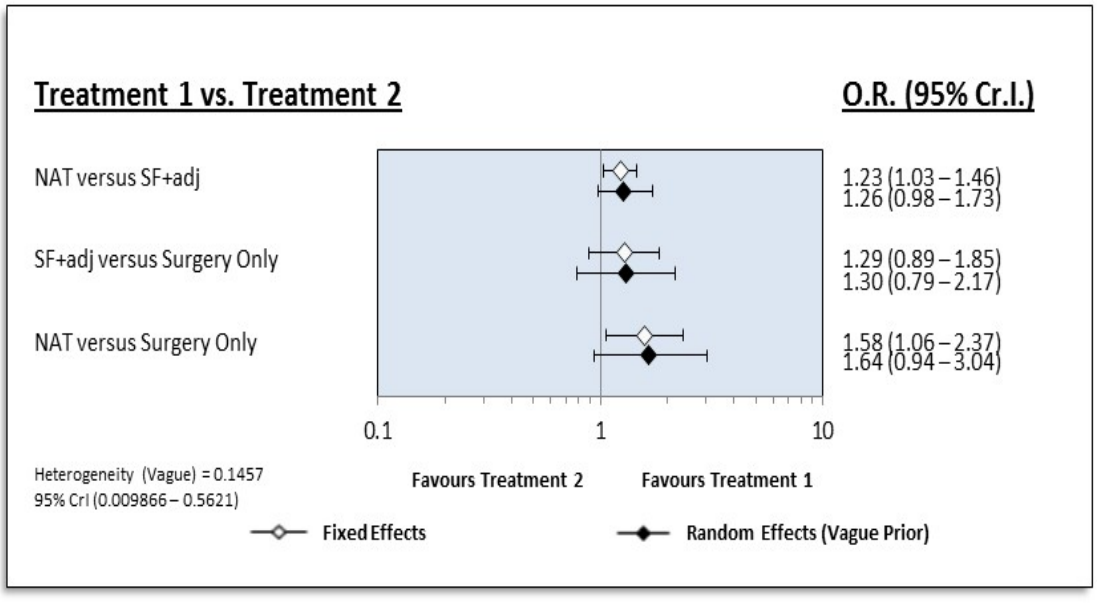
2-year Survival

Network pairwise comparison of neoadjuvant therapy and upfront surgery plus adjuvant therapy for 2-year survival was based on 7 studies (n=4251; neoadjuvant therapy n=903; upfront surgery plus adjuvant therapy n= 3348). Aggregate rate of 2-year survival was

0.5178 (0.3000-0.5970) versus 0.5131 (0.2727-0.5346) in favour of neoadjuvant therapy. Both fixed effects (O.R. 1.22; 95% CI 1.02-1.46) and random effects model (O.R. 1.26; 95% CI 0.94-1.74) favoured neoadjuvant therapy with SUCRA 0.95 for neoadjuvant therapy (Appendix L).

Inclusion of upfront surgery plus adjuvant therapy *versus* surgery only RCTs in a network based on 9 studies (n=4723; neoadjuvant therapy: n=903; upfront surgery plus adjuvant: n=3585; surgery only: n=235) also demonstrated superiority of neoadjuvant therapy for 2-year survival in both fixed and random effects model. Aggregate of 2-year survival was 0.5217 (0.3000-0.5970) for neoadjuvant therapy compared to 0.5107 (0.2727-0.5346) for upfront surgery plus adjuvant therapy and 0.4149 (0.4000-0.4200) for surgery only (Figure 24; Appendix L).

Figure 24: Results of fixed effects and random effects (vague prior) models

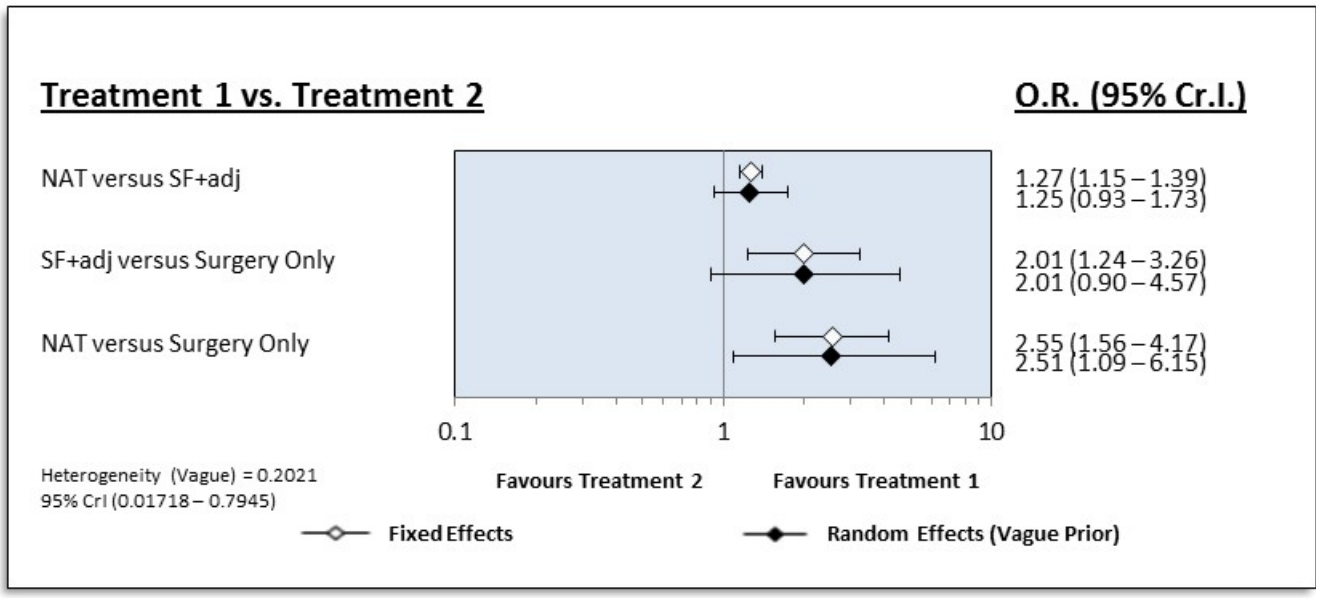


3-year Survival

Pairwise comparison of neoadjuvant therapy versus upfront surgery plus adjuvant therapy was based on a network comprising 8 studies (n= 12011; neoadjuvant therapy: n=2708; upfront surgery plus adjuvant therapy: n=9303) and demonstrated superiority of neoadjuvant therapy with aggregate rate of 0.3367 (0.1212-0.3900) to 0.2943 (0.1800-0.4700). Again both fixed effect (O.R. 1.25 95% CI 1.14-1.38) and random effects (O.R. 1.19 9% CI 0.86-1.51) models favored neoadjuvant therapy with SUCRA 0.9 demonstrating stronger positive effect with neoadjuvant therapy on outcomes of 3-year survival (Appendix L).

Inclusion of upfront surgery plus adjuvant therapy *versus* surgery only RCTs in a network produced comparisons based on 9 studies (n=12365; neoadjuvant therapy: 2708; upfront surgery plus adjuvant therapy: n=9482; surgery only: n= 175). Neoadjuvant therapy was superior in both fixed and random effects models with aggregate rate 0.3400 (0.2000-0.4194) compared to 0.2951 (0.1800-0.4700) for upfront surgery plus adjuvant therapy and 0.2050 (0.2050-0.2050) for surgery only (Figure 25; Appendix L).

Figure 25: Results of fixed effects and random effects (vague prior) models



4-year Survival

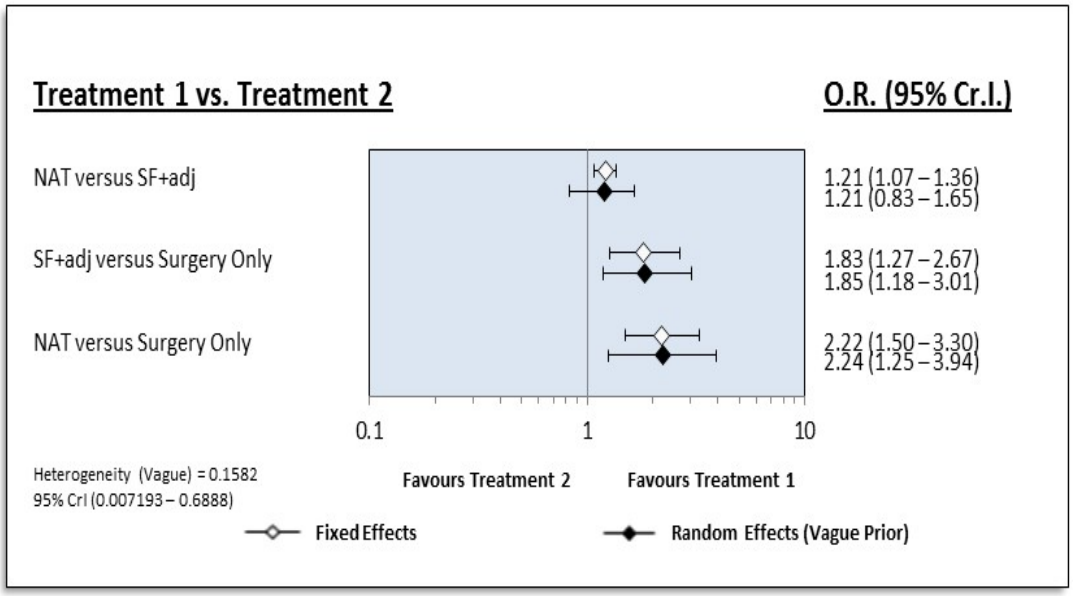
Only pairwise comparison of neoadjuvant therapy and upfront surgery plus adjuvant therapy could be offered, as upfront surgery plus adjuvant therapy *versus* surgery only RCTs did not report 4-year survival rates. This network was based on 4 studies (n=656). Neoadjuvant therapy was superior with aggregate rate 0.1416 (0.0303-0.2500) compared to 0.1269 (0.0606-0.2000). Fixed effects (O.R. 1.16; 95% CI 0.69-1.94) and random effects model (O.R 1.03; 95% CI 0.27-3.13) favoured neoadjuvant therapy (Appendix L).

5-year Survival

Network pairwise comparison of 5-year survival for neoadjuvant therapy and upfront surgery plus adjuvant therapy was based on 7 studies (n=8896; neoadjuvant therapy n=2558; upfront surgery plus adjuvant therapy n=6338). Aggregate rate for neoadjuvant therapy was 0.2069 (0.0323-0.3300) compared to 0.1783 (0.0606-0.2300). Fixed effects (O.R 2.21; 95% CI 1.07-1.37) and random effects (O.R. 1.19; 95% 0.65-1.73) favoured neoadjuvant therapy with SUCRA 0.82 for neoadjuvant therapy (Appendix L).

Inclusion of upfront surgery plus adjuvant therapy *versus* surgery only RCTs was based on 11 studies (n=9675; neoadjuvant therapy n=2558; upfront surgery plus adjuvant therapy n=6730; surgery only n=387). Neoadjuvant therapy was superior across fixed effects and random effects models with aggregate rate 0.2069 (0.0323-0.3300) followed by 0.1814 (0.0606-0.2640) for upfront surgery plus adjuvant therapy and 0.1418 (0.1040-0.2200) for surgery only (Figure 26; Appendix L).

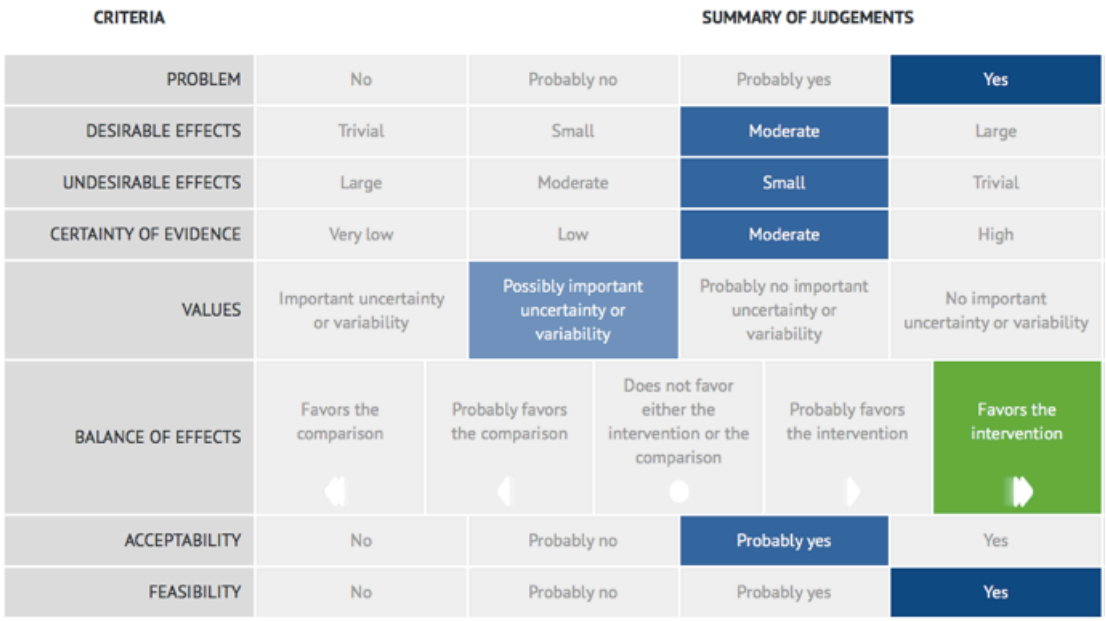
Figure 26: Results of fixed effects and random effects (vague prior) models



Assessment of Impact of FUPS characteristics of Data

Using the GRADE assessment criteria again the certainty of recommendations from the network analysis showed that although neoadjuvant therapy was marginally favoured overall, uncertainty was identified in the evidence synthesised (Figure 27). This was corroborated by the risk-of-bias assessment of included trials (Table 15; Appendix K).

Figure 27**:** An assessment of the strength of overall recommendations from the network meta-analysis according to the GRADE assessment criteria.



Convergence was achieved across all models and no issues were identified with inconsistency. In 2-year survival analysis and 5-year survival analysis there was a marginal preference towards fixed effects model as determined by the DIC statistic (Appendix L).

Discussion

Data as a Partial Remnant

Upfront surgery plus adjuvant therapy is a well established treatment pathway for resectable pancreatic cancer (Neoptolemos *et al*., 2001). Neoadjuvant therapy is supported by current guidelines for borderline resectable and locally advanced pancreatic cancer but its role in the management of resectable pancreatic cancer remains controversial (Tempero et al., 2014; de Geus et al., 2016). In the absence of conclusive results from large multi-centered RCTs this study, the first of its kind, utilises existing studies comparing neoadjuvant therapy and upfront surgery plus adjuvant therapy for the treatment of potentially resectable, and separately resectable pancreatic cancer, in a Bayesian network meta-analysis to offer an important interim analysis to inform the ongoing debate regarding the best treatment for potentially resectable, and in particular resectable, pancreatic cancer.

Overall this analysis marginally favours neoadjuvant therapy. When analysing its use in all potentially resectable pancreatic cancer neoadjuvant therapy compared favourably with traditional upfront surgery plus adjuvant therapy approach and demonstrated survival benefit across 1,2,4 and 5-year survival outcomes. There was no difference in 3-year survival but inclusion of cohort studies and RCTs demonstrated benefit with neoadjuvant therapy. For the treatment of resectable pancreatic cancer a marginal benefit was found with neoadjuvant therapy across outcomes of R0 resection, 1,2,3,4 and 5-

year survival. This is based on the best available studies and did not alter on sensitivity analysis. However, issues pertaining to quality and level of bias of available studies are an issue that weakens the strength and level of certainty of any such recommendations.

Transparency of Analysis

A strength of this study is that it included a separate analysis of only studies of resectable pancreatic cancer, identified through comprehensive literature search, to offer a true like-for-like comparison based on currently available evidence. Analysis of neoadjuvant therapy *versus* upfront surgery plus adjuvant therapy were based on direct comparisons to strengthen certainty of findings with indirect comparisons drawn from inclusion of upfront surgery plus adjuvant therapy *versus* surgery only trials, and the inclusion of cohort observational studies only in sensitivity analysis which did not alter network findings. However, this study also shares the limitations of the existing body of evidence pertaining to the treatment of potentially resectable pancreatic cancer: heterogeneity and small underpowered sample size (Andriulli *et al*., 2012). Although random effects modeling was employed to counter heterogeneity, overall there is a lack of RCTs comparing neoadjuvant therapy and upfront surgery plus adjuvant therapy (Lee *et al*., 2016; Andriulli *et al*., 2012; Sharma *et al*., 2015; de Geus *et al*., 2016). Only one of the two phase II trials for resectable pancreatic cancer were randomised (Golcher *et al*., 2015) with the remaining studies being either prospective or retrospective studies which raises serous

concerns about bias and reduced certainty in the recommendations drawn from the network meta-analysis.

Unlike the majority of existing network meta-analysis (Bafeta *et al*., 2013; Hutton *et al*., 2014; Zarin *et al*., 2017), this study went beyond only assessing bias of included trials to utilise GRADE approach to rate the certainty in estimates from our network meta-analysis (Brown *et al*., 2014; Faltinsen *et al*., 2018; Guyatt *et al*., 2008; Puhan *et al*., 2014; Salanti *et al*., 2014). Rather than dismiss the entirety of the existing body of research on the basis of its FUPS characteristics and simply conclude that RCTs are awaited, statistical techniques were utilised to quantify the perceived limitations of the existing data. This included assessing convergence using the Brooks-Gelman-Rubin method and by checking whether the Monte Carlo error is less than 5% of the standard deviation of the effect estimates and between-study variance (Brown *et al*., 2014). Furthermore, the MCMC Bayesian network meta-analysis was fitted with three chains as a means of checking MCMC convergence (Brown *et al*., 2014). Inconsistency assessment, the conflict between direct and indirect evidence, is crucial to any network meta-analysis (Dias *et al*., 2013) and was measured by comparing deviance residuals and DIC statistic in fitted consistency and inconsistency models (Brown *et al*., 2014; Dias *et al*., 2013; Spiegelhater *et al*., 2002).

This means that this study can go further than existing studies to provide an important interim analysis that adds a further dimension to the debate regarding the best treatment approach for potentially resectable pancreatic cancer. Firstly it offers an indirect comparison

of outcomes between neoadjuvant treatment pathway and outcomes for those who receive surgery only. Therefore rather than only comparing upfront surgery plus adjuvant therapy and neoadjuvant pathways this helps to explore more individualised outcomes for those patients who do not progress to receiving adjuvant therapy in the upfront surgery pathway and therefore are treated with surgery only. Secondly by transparently quantifying the impact of FUPS characteristics on the study outcomes, this study both highlights and assesses the impact of the limitations of the existing body of evidence on which current assumptions and beliefs regarding treatment approaches are based.

Triangulation

These findings are corroborated by previous attempts to synthesise existing evidence comparing upfront surgery plus adjuvant therapy and neoadjuvant therapy for resectable pancreatic cancer. Meta-analysis by both Xu *et al*. (2014) and Andriulli *et al*. (2012) reported marginal benefit of neoadjuvant therapy for resectable pancreatic cancer in terms of overall survival and disease free survival for resectable cases. However, neither of these reports focused solely on neoadjuvant therapy and therefore omitted significant studies from their meta-analysis (Lee *et al*., 2016). Sharma *et al*. (2015) and de Geus *et al.* (2016) synthesised published data in a Markov decision-analysis model to compared neoadjuvant therapy and upfront surgery plus adjuvant therapy for the treatment of resectable pancreatic cancer and also reported marginal benefit of neoadjuvant therapy. More recently Versteijne *et al*. (2018) reported more

significant survival benefit with neoadjuvant therapy in their meta-analysis but the reported weighted mean overall survival time included borderline resectable cases therefore captured the effect of conversion to resectability affecting overall survival time in neoadjuvant therapy pathway. The reported weighted mean overall survival time for resectable only cases was lower although still superior to upfront surgery plus adjuvant therapy (Versteijne *et al*.,2018).

The second key outcome explored through direct and indirect comparison was the rate of R0 resection, which is known to impact survival time (Howard *et al*., 2006). Once again neoadjuvant therapy was found to be superior to upfront surgery plus adjuvant therapy which is in keeping with the hypothesis that neoadjuvant therapy results in higher rates of R0 resection (Asare *et al*., 2016; Lee *et al*., 2016; Chua *et al*., 2011). However, definitions of R0 resection can vary between studies, which could potentially impact reported outcomes (Versteijne *et al*.,2018). In this study convergence was achieved across all models comparing this outcome and no issues with inconsistency were identified in our analysis.

A key clinical concern when selecting a treatment pathway for is the delivery of multimodal treatment: resection in the neoadjuvant therapy pathway and receipt of adjuvant therapy in the upfront surgery pathway. Our analysis of pooled proportions found that for resectable pancreatic cancer 63% of patients in the upfront surgery plus adjuvant therapy pathway received adjuvant therapy, and 76% in the neoadjuvant therapy pathway underwent resection. These

findings are in keeping with the results of a recent meta-analysis of pooled proportions that reported 68.6% of patients in upfront surgery pathway received adjuvant therapy and 76.8% of resectable cases in neoadjuvant therapy pathways underwent resection (Versteijne *et al*.,2018).

Conclusion: emergence, boundary setting, lack of complete knowledge, ethics and future direction of research

To conclude our Bayesian network meta-analysis shows that neoadjuvant therapy is no worse than traditional upfront surgery plus adjuvant therapy approach and may even hold benefit across outcomes of: R0 resection, 1,2,3,4,and 5-year survival for potentially resectable and resectable cases of pancreatic cancer. This finding in the context of the limitations of existing studies means that conclusive superiority of one approach over another cannot be determined without a degree of uncertainty. Furthermore the boundaries of this meta-analysis are determined by how outcomes are reported in studies and therefore emergence as multiple factors dynamically interact within the complex system of pancreatic cancer management delivery has not yet been explored. A lack of complete knowledge regarding the system therefore remains and it would be unethical to conclude that either pathway has superiority. In light of these findings the possibility is raised that superior pathway selection may be determined at individual patient level. The potential of Bayesian statistical approach in testing this hypothesis through Markov decision analysis was therefore undertaken.

4.2 Markov Decision Analysis

Publications resulting from this analysis:

Bradley, A. *et al.* (2018). 'Markov decision analysis of neoadjuvant treatment pathway versus surgery first pathway for resectable pancreatic cancer'. *Journal of Clinical Oncology,* 36 (4). pp. 456-456

Bradley, A. and Van Der Meer, R. (2019). 'Neoadjuvant therapy versus upfront surgery for potentially resectable pancreatic cancer: a Markov decision analysis'. *PLoS One*, 14(2):e0212805. doi:10.1371/journal.pone.0212805

Abstract

Background: Neoadjuvant therapy has emerged as an alternative treatment strategy for potentially resectable pancreatic cancer. In the absence of large RCTs offering a direct comparison, this study aims to use Markov decision analysis to compare efficacy of traditional upfront surgery plus adjuvant therapy (which will be referred to as surgery first pathway (SF) within this model) and neoadjuvant treatment pathways (which will be referred to as NAT within this model) for potentially resectable pancreatic cancer. Competing pathways will then also be compared solely for cases that are resectable at presentation and the results of this analysis will be triangulated with the results of a Markov decision-analysis based on a prospectively maintained patient database from a tertiary referral centre pancreatic unit.

Methods: An advanced Markov decision analysis model was constructed to compare SF and NAT pathways. Transition probabilities were first calculated from RCTs and phase II/III trials after comprehensive literature search. The model was then populated with data from a prospectively maintained tertiary referral centre database. Utility outcomes were measured in overall and quality-adjusted-life months (QALMs) on an intention-to-treat basis as the primary outcome. Markov cohort analysis of treatment received was the secondary outcome. Model uncertainties were tested with one and two-way deterministic and probabilistic Monte Carlo sensitivity analysis.

Results Using Synthesised Data from Published Studies for Potentially Resectable Pancreatic Cancer: SF gave 23.72 months (18.51 QALMs) *versus* 20.22 months (16.26 QALMs). Markov cohort analysis showed that where all treatment modalities were received NAT gave 35.05 months (29.87 QALMs) *versus* 30.96 months (24.86QALMs) for R0 resection and 34.08 months (29.87 QALMs) *versus* 25.85months (20.72 QALMs) for R1 resection. One-way deterministic sensitivity analysis showed that NAT was superior if the resection rate was greater than 51.04% or below 75.68% in SF pathway. Two-way sensitivity analysis showed that pathway superiority depended on obtaining multimodal treatment in either pathway.

Results Using Synthesised Data from Published Studies for Resectable Pancreatic Cancer: NAT pathway yielded 26.41 months (22.54 QALMs) compared to 23.72 months (18.51 QALMs). Markov

cohort analysis showed that in patients who received all treatment modalities NAT pathway yielded 39.34 months (34.63 QALMs) compared to 30.96 months (24.86 QALMs) for R0 resection and 34.94 months (31.07 QALMs) compared to 25.85 months (20.72 QALMs) for R1 resection. Deterministic sensitivity analysis demonstrated that pathway superiority depended on the probability of receiving multimodal treatment in either pathway.

Results Using Institutional Patient Database for Resectable Pancreatic Cancer: NAT yielded 32.90 months (28.51 QALMs) compared to 24.68 months (19.23 QALMs). Deterministic sensitivity analysis demonstrated the importance of receiving multimodal treatment in determining pathway superiority. Probabilistic Monte Carlo analysis reported NAT pathway superiority. Markov cohort analysis showed that greatest utility was achieved in the subgroup of patients in the SF pathway who received R0 resection and adjuvant therapy (42.38 QALMs).

Conclusion: Whilst NAT is a viable alternative to traditional SF approach, even for cases that are resectable at presentation, superior pathway selection depends on the individual patient's likelihood of receiving multimodal treatment in either pathway. Careful consideration must be given to patient selection pertaining to likelihood of receiving all treatment modalities and achieving R0 resection in either pathway. Future research must therefore focus on developing ways of engaging with the complexity to move towards personalised predictive modeling to support individualised treatment selection.

Introduction

Surgical resection followed by adjuvant therapy has become the standard of care for resectable pancreatic cancer (Neoptolemos *et al*., 2001). However, despite advances in surgical techniques and adjuvant therapies, 5-year survival for resected pancreatic cancer has been reported at between 7% and 25% (CRUK, 2019). Furthermore, up to 50% of patients do not actually receive adjuvant therapy post resection due to a combination of factors including: post-operative complications, early metastases nullifying the potential benefits of high-risk surgery (Winter *et al*., 2012) and reduced performance status due to pre-existing medical conditions (Bilimoria *et al*., 2007). This has resulted in a growing interest in neoadjuvant therapy (Asare *et al*., 2016; Lee *et al*., 2016).

There is currently a lack of RCTs comparing upfront surgery and neoadjuvant therapy treatment pathways (Versteijne *et al*., 2018). Despite promising results from cohort studies and phase II trials, existing meta-analysis corroborate the findings from section 4.1 in reporting only marginal benefit of NAT in terms of overall and disease-free survival (Lee *et al*., 2016; Sharma *et al*., 2015; Xu *et al*., 2014; Andriulli *et al*., 2012; Petrelli *et al*., 2015). Whilst the a role for NAT has been broadly accepted for cases that are borderline resectable or locally advanced, neoadjuvant therapy for resectable pancreatic cancer therefore remains an area of prime controversy.

Two previous Markov decision-analysis found marginal benefit with neoadjuvant therapy for resectable only cases (Sharma *et al*., 2015;

DeGus *et al*., 2016). Sharma *et al*. (2015) used data drawn from prospective phase II and III trials. De Gus *et al.* (2016) also included data from retrospective studies compiled from a literature search from a single search engine. Only one previous Markov decision analysis compared efficacy of both pathways for potentially resectable disease and found no conclusively superior pathway (VanHouten et al., 2012). These studies share the limitations of the existing body of evidence: heterogeneity and small underpowered sample size.

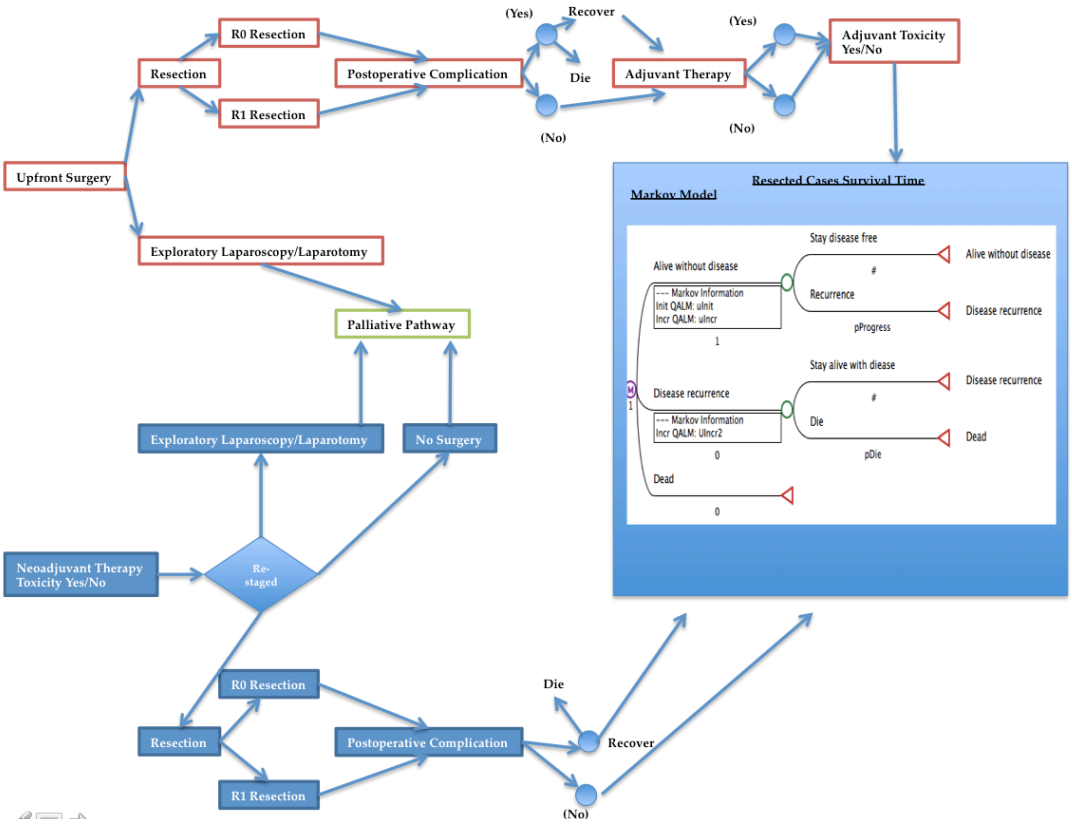
The aim of this section is to compare upfront surgery (SF) *versus* neoadjuvant therapy (NAT) for the treatment of potentially resectable pancreatic cancer (including resectable, borderline resectable and locally advanced cases) through Markov decision-analysis (Section 4.2.1). These competing treatment pathways will then be compared for only cases that are resectable at presentation (Section 4.2.2) with the results from the Markov decision-analysis using synthesised data triangulated against those from a Markov decision-analysis using patient data form a tertiary referral centre (Section 4.2.3). The objectives are to compare predicted outcomes between both pathways on an intention-to-treat basis. However, this research aims to go further by using the FUPS data in the context of complexity to attempt to uncover whether, by developing more transparent ways of statistically engaging with the complexity of the system, more individualised ways of determining treatment pathway superiority at a personalised level could emerge.

Materials and methods

Markov model

TreeAge Pro 2017 (TreeAge Software Ins., Williamstown, MA) was used to construct a Markov cohort decision analysis model in an advanced decision-tree format comparing base case, SF (with adjuvant therapy including chemotherapy, chemoradiotherapy, or both), to NAT (which included chemotherapy and/or chemoradiotherapy) followed by re-staging and, if possible, surgical resection (Figure 28). Upon completion of treatment, cohorts entered the Markov health-state transition model with possible survival states including: alive without disease, alive with disease and dead. Each Markov cycles equated to 1 month with maximum follow-up of 60-cycles or until death.

Figure 28: Overview of the structure of the Markov decision-tree.



Outcome Measures

Cumulative payoffs were calculated in life months and QALMs, which scaled survival from 0 (equivalent to death) to 1 (Sharma *et al.*, 2015; DeGus *et al.*, 2015) based on indicies taken from published literature (Ljungman *et al.*, 2011; Murphy *et al.*, 2012) and World Health Organization and European Quality of Life Survey (Eshuis *et al.*, 2015; Romanus *et al.*, 2012; Tam *et al.*, 2013).

Transparency of Analysis

Data sources and transition probabilities

Source data was identified through comprehensive literature search of MEDLINE, Embase, PubMed and Cochrane database and Cochrane database of Clinical Trials following the PRISMA checklist (Moher et al., 2009). For each of the searches, the entire database was included from the year 2000 up to and including 31st October 2018, with no further date restrictions or limits applied. Following screening, reference lists and citations of all included papers were manually searched to identify any additional articles until no new articles were identified. The following data was extracted from each study: study details (country, year, design, number, mean age, sex, co-morbidity profile and presenting disease stage of participants), details of treatment protocols, treatment outcomes (treatment completion rates, rates of tumour resection, R0 resection rates, drug toxicity data, post-operative complication rates, overall survival and disease-free survival) and risk-of-bias data.

The inclusion criteria was RCTs and prospective phase II and III studies of neoadjuvant therapy for the treatment of pancreatic cancer, published in English language since 2000, involving chemo/radiotherapy-naive human subjects over 18 years of age with preoperatively staged pancreatic cancer as potentially resectable. Included trials had to report: protocol design, number of participants per arm, median age and co-morbidities of subjects, pre-treatment staging of pancreatic cancer, toxicity profile, results of post

neoadjuvant re-staging, resection rates, post-operative complications defined by Clavien-Dindo system, and survival data. Retrospective and cohort studies, case series and case reports were excluded as were studies from identical patient cohorts and trials involving intra-operative radiotherapy and trials including disease other than pancreatic cancer. Trials matching this inclusion criteria that reported only outcomes for cases of pancreatic cancer that were resectable at presentation, or that reported the outcomes for resectable only cases separately, were also included in a separate Markov decision-analysis for the treatment of resectable pancreatic cancer.

As the majority of trials were single arm, to populate the upfront surgery pathway the same databases were searched for RCTs of surgery and adjuvant therapy, with the same inclusion and data reporting criteria. The outcomes of this group could introduce bias because by definition these patients have survived surgery and not developed early metastatic disease and also had to have adequate performance status to be randomised to adjuvant therapy even if they did not receive adjuvant therapy. To overcome this issue cohort studies comparing neoadjuvant therapy and upfront surgery, with the otherwise same inclusion criteria and data reporting requirements, were also included in the upfront surgery arm and solely used to offer comparison across outcomes of resection, R0 resection rates and receipt of adjuvant therapy.

Statistical analysis

Markov model transition probabilities were based on weighted pooled estimates of proportions from included studies, calculated using Freeman-Tukey (1950) arcsine square root transformation under random effects model to account for heterogeneity. Survival time was based on time from diagnosis. Gillen *et al*. (2010) approach to calculating weighted median survival time was used as evidence has shown that weighted averaging of medians cannot achieve unbiased pooled estimates of survival time (Rouder *et al*., 2004). This approach is based on averaging parameter estimates of a presumed density function of survival. The pooled distribution parameter is used to recalculate the estimate of the median from the pooled distribution parameter (Gillen *et al*., 2010). In this case the pooled distribution parameter is the exponential distribution, which implies a time constant hazard rate corresponding to the sole distribution parameter λ. From this the weighted estimate of median survival (m_p) is derived from the formula (Gillen *et al*., 2010):

$$mp = \left(\sum_{i=1}^{k} \frac{wi}{mi}\right)^{-1}$$

where m_i is median survival within the study population i (with i being 1 to k where k is the number of included studies) (Gillen *et al*., 2010). w_i is the study specific weight function derived from number of study participants divided by total number of evaluable patients (Gillen *et al*., 2010).

Quantifying the Limitations Stemming from FUPS Characteristics of Data

The Cochrane Collaboration's risk of bias tool (Higgins *et al*., 2011) and ROBINS-I tool (Sterne *et al*., 2016) were used to assess the quality and risk-of-bias of each included trial. Furthermore the potential impact of bias and uncertainty on all variables within the model were extensively tested through deterministic and probabilistic sensitivity analysis.

Model uncertainties for all included components were tested with one and two-way deterministic sensitivity analysis with baseline transition probabilities for each variable altered between highest and lowest reported values. Probabilistic Monte Carlo sensitivity analysis was set to 10000 iterations with model probabilities sampled from the entirety of the data distribution of each variable contained within the Markov models. Data for each variable was fitted against 55 possible distributions with the best fit determined by the Anderson Darling statistic.

4.2.1 Results: Markov Decision-Analysis for Potentially Resectable Pancreatic Cancer

Eligible Studies

50 phase II/III studies met the inclusion criteria and were included in the neoadjuvant therapy arm of the model, 4 of which were

randomised. 9 of these studies offered comparison with upfront surgery (Appendix M).

For the upfront surgery pathway 15 studies were RCTs, 10 of which offered comparison between adjuvant regimes, 5 of which offered comparison between adjuvant therapy and surgery only (Appendix M). 16 cohort studies were also included in the upfront surgery pathway to offer comparison across outcomes of resection rates, R0 resection, and rates of receiving adjuvant therapy (Appenidx M). Probability estimates and ranges and quality-of-life utilities are displayed in Table 17.

Table 17: Summary of transition probabilities, parameters of data distribution and payoff utilities for quality adjusted life months (QALMs).

Variable	**Baseline Transition Probability (95% Confidence Interval)**	**Range**	**Standard Deviation**	**Variance**	**Data Distribution: parameters (Anderson Darling Statistic)**
Grade 3+ toxicity with NAT	0.35 (0.28-0.43)	0-1.0	0.03799	0.00144	Generalised Extreme Value: k=0.45856 σ=0.01111 μ=0.00904 (0.55904)
Resection in NAT pathway	0.41 (0.33-0.49)	0-0.86	0.00848	7.1972E-5	Generalised Extreme Value: k=0.15727 σ=0.00545 μ=0.00618 (0.36129)
Exploratory Laparoscopy /Laparotomy	0.1 (0.07-0.13)	0-0.36	0.00349	1.2182E-5	Generalised Pareto: k=0.06879 σ=0.00306 μ= -5.1223E-4 (1.3525)
R0 resection NAT pathway	0.29 (0.21-0.36)	0-0.74	0.0068	4.6303E-5	Johnson SB: γ=1.7195 δ=1.0417 λ=0.04849 ξ=-0.00113 (0.35896)
Grade 3-4 post-operative complication NAT pathway	0.35 (0.19-0.53)	0.11-0.64	0.02702	7.3021E-4	Generalised Extreme Value: k=-0.45505 σ=0.03128 μ=0.04101 (0.1996)
Grade 5 post-operative complication NAT pathway	0.02 (0.01-0.03)	0-0.36	0.00097	9.4387E-7	Pareto 2: α=0.34207 β=1.3899E-13 (-13.983)
Resection SF pathway	0.94 (0.90-0.96)	0.70-1.0	0.1219	0.01486	Burr: k=0.0595 α=10.327 β=0.00112 (0.12818)
R0 resection SF pathway	0.56 (0.51-0.62)	0.16-0.86	0.09869	0.00974	Pearson 5: α=0.61636 β=7.0460E-4 (0.18259)
Grade 3-4 post-operative	0.22 (0.13-0.33)	0.04-0.54	0.01297	0.0002	Log-Pearson 3: α=66.845 β=-0.09425

complication SF pathway					γ=2.0838 (0.29235)
Grade 5 post-operative complication SF pathway	0.07(0.02-0.13)	0-0.36	0.00948	8.9795E-5	Cauchy: σ=0.00373 μ= 0.00639 (0.38658)
Receiving adjuvant therapy	0.61(0.57-0.66)	0.26-0.94	0.10088	0.01018	Burr: k=0.26048 α=2.145 β=9.2071E-4 (0.18949)
Adjuvant toxicity grade 3+	0.43(0.25-0.62)	0.09-0.98	0.02753	0.00076	Log-Pearson 3: α=1916.0 β=-0.02672 γ=47.081 (0.34508)

Survival State	**Utility for QALM**
Living with stable pancreatic cancer	0.81
Undergoing chemo/radio therapy	0.81
Experiencing chemo/radio therapy complications	0.53
Recovering from pancreatic surgery	0.59
Experiencing surgical complications	0.48
Living with unresectable disease and pre-operative quality-of-life	0.65

***NAT= Neoadjuvant Pathway; SF = Surgery First (or Upfront Surgery) Pathway**

Results of Markov decision-analysis

Intention-to-treat analysis of the treatment pathways, based on baseline transition probabilities, showed that upfront surgery pathway gave 23.72 months (18.51 QALMs) compared to 20.22 months (16.26 QALMs) for neoadjuvant therapy pathway. The results of Markov cohort analysis are outlined in Table 18 and demonstrate superiority of the NAT pathway for patients who received all treatment modalities.

Table 18: Results from Markov cohort analysis

	NAT Pathway	**SF Pathway**
R0 Resection	35.05 months (29.87 QALMs; POC =29.76 QALMs)	Received Adjuvant Therapy: 30.96 months (24.86 QALMs; POC= 24.75 QALMs; AT= 21.82 QALMs; POC and AT=21.71 QALMs) No Adjuvant Therapy: 24.03 months (20.12 QALMs; POC=20.01QALMs)
R1 Resection	34.08 months (29.87 QALMs; POC=29.76 QALMs)	Received Adjuvant Therapy: 25.85 months (20.72 QALMs; POC= 20.61 QALMs; AT= 18.20 QALMs; POC and AT=18.09 QALMs) No Adjuvant Therapy: 21.26 months (17.56 QALMs; POC=17.45 QALMs)
Exploratory Laparoscopy or Laparotomy	10.86 months (7.22QALMs)	10.48 months (6.97QALMs)
No Surgery	10.86 months (7.06 QALMs)	

POC= post-operative complication grade 3 or 4; AT= adjuvant therapy resulting in grade 3 or 4 toxicity

Deterministic sensitivity analysis

Deterministic sensitivity analysis tested the sensitivity of the results of the model to variations in parameters of specific model variables by altering the parameters between highest and lowest reported values. One-way deterministic sensitivity analysis determined the effect on the overall results of the model by varying the parameter of each variable individually. Two-way deterministic sensitivity analysis determined the effect on the model of altering the parameters of two variables simultaneous.

One-way deterministic sensitivity analysis showed that NAT was the superior treatment pathway if the probability of achieving resection in this pathway was greater than 51.04% (Figure 29) or the probability of achieving resection in the SF pathway was less than 75.68% (Figure 30).

Figure 29: One-way deterministic sensitivity analysis of the probability of resection in Neoadjuvant pathway. This figure shows the effect of altering the baseline probability of resection in the neoadjuvant pathway on overall model outcome.

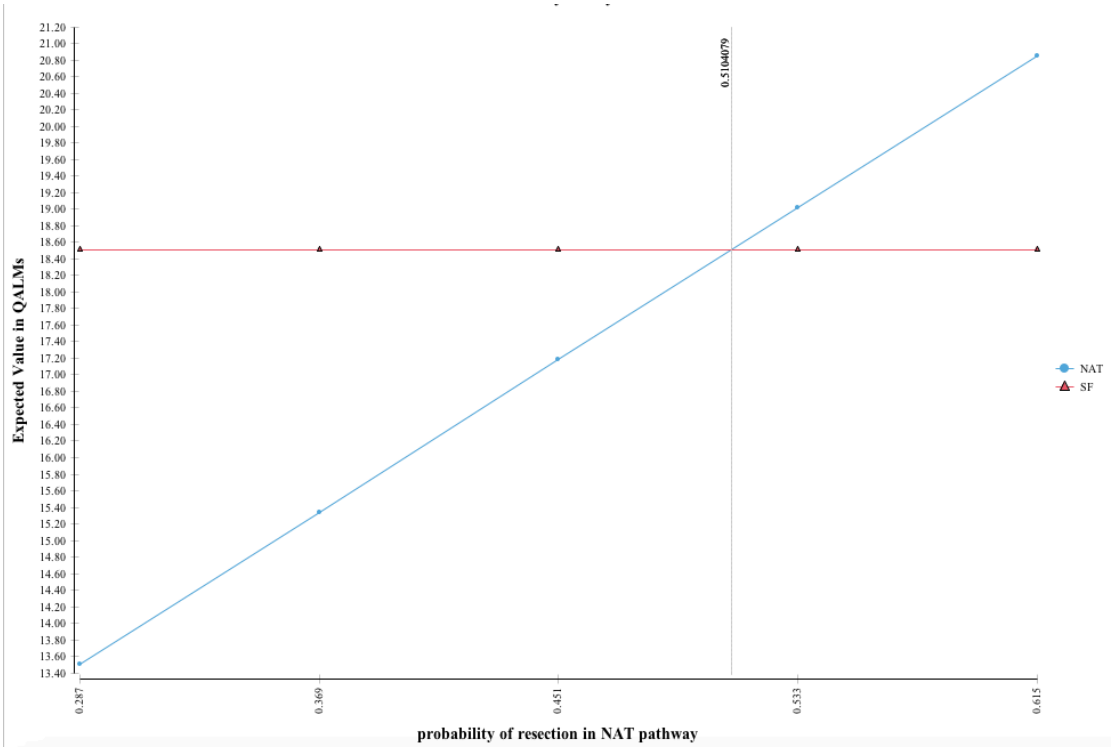
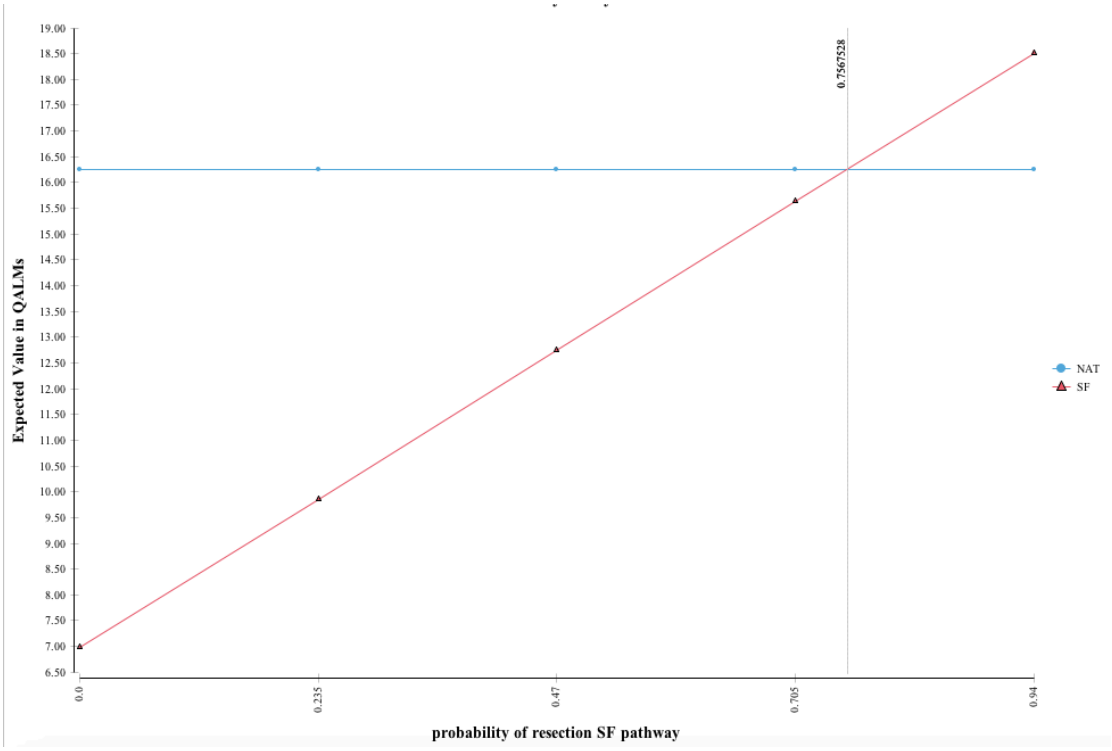


Figure 30: One-way deterministic sensitivity analysis of the probability of resection in the surgery first pathway. This figure shows the effect of altering the baseline probability of resection in the surgery first pathway on overall model outcome.



Two-way deterministic sensitivity analysis demonstrated that treatment superiority depended on receiving multimodal treatment (resection in the NAT pathway and adjuvant therapy in the SF pathway). Fig 31a shows the thresholds at which competing pathways offer superior outcomes with Figure 31b providing corresponding probability thresholds and predicted resulting quality-adjusted survival time.

Fig 31a. Two-way sensitivity analysis. Y-axis shows probability of receiving adjuvant therapy in surgery first (SF) pathway and x-axis shows probability of receiving resection in neoadjuvant (NAT) pathway. The red area represents where patients, given competing probability of receiving multimodal treatment in competing pathways, would benefit from surgery first approach. The blue area represents where neoadjuvant therapy would be the superior treatment pathway in terms of quality-adjusted survival.

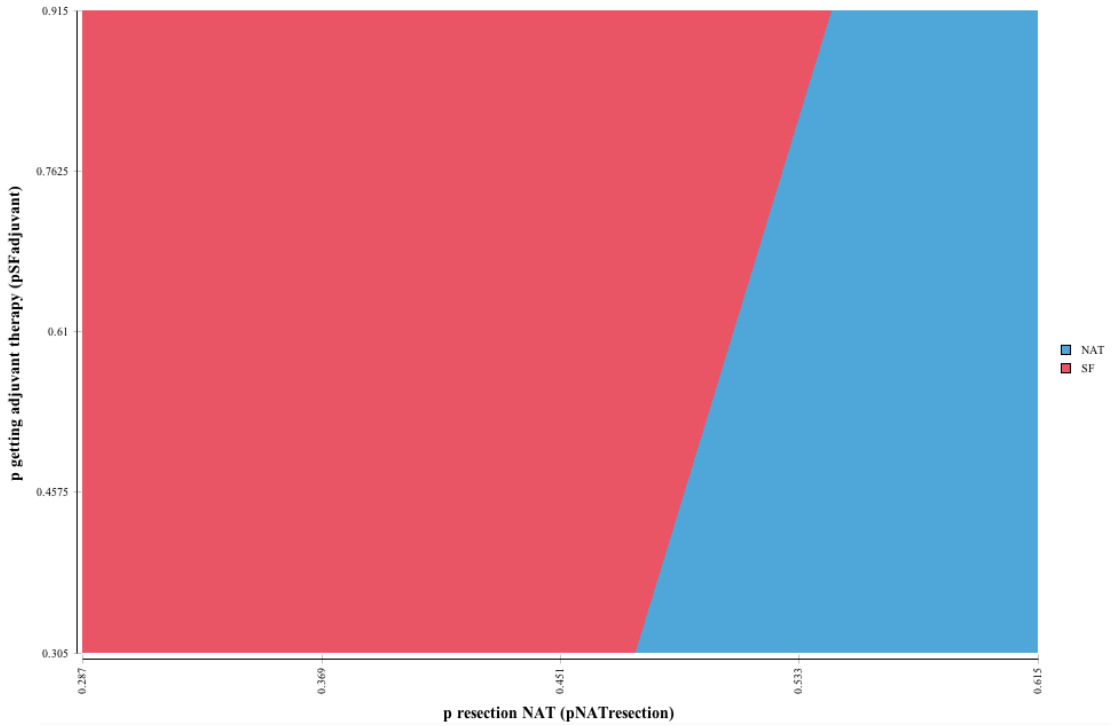


Fig 31b. Corresponding predicted survival time in QALMs. X and Y-axis provide altering probabilities of multimodal treatment in each pathway with corresponding survival time given in QALMs.

p getting adjuvant therapy \ p resection NAT	0.287	0.369	0.451	0.533	0.615
▼0.305					
SF	17.76	17.76	17.76	17.76	17.76
NAT	13.51	15.34	17.18	19.02	20.85
▼0.4575					
SF	18.13	18.13	18.13	18.13	18.13
NAT	13.51	15.34	17.18	19.02	20.85
▼0.61					
SF	18.51	18.51	18.51	18.51	18.51
NAT	13.51	15.34	17.18	19.02	20.85
▼0.7625					
SF	18.89	18.89	18.89	18.89	18.89
NAT	13.51	15.34	17.18	19.02	20.85
▼0.915					
SF	19.27	19.27	19.27	19.27	19.27
NAT	13.51	15.34	17.18	19.02	20.85

Probabilistic sensitivity analysis

Probabilistic sensitivity analysis tested the level of confidence in the model output in relation to uncertainty in model input by determining the distribution of the input data for each variable from the median, standard deviation and variance of the input data (Table 17). All possible parameter values for each variable within the model were therefore tested by drawing probabilities from the data distribution when probabilistic Monte Carlo sensitivity analysis was set to simulate 10000 patients cycling through the model.

The results of probabilistic Monte Carlo sensitivity analysis showed that SF gave a mean survival time of 19.72 months (range 5.57-

22.95) compared to 17.16 months (range 16.50-17.38) for NAT with standard deviation 2.68 and 0.19, and variance 7.17 and 0.04 in SF and NAT pathways respectively. When minimum significant difference was set at 3.65 months or greater, the model reported indifference in superior pathway selection frequency.

4.2.2 Results: Markov Decision-Analysis for Resectable Pancreatic Cancer

Eligible Studies

A total of 18825 studies on neoadjuvant therapy for pancreatic cancer were identified, of which 452 underwent full screening. 50 phase II/III studies were identified, 9 of which offered comparison with upfront surgery and were included in the SF pathway. 18 of the 50 studies reported outcomes either solely for resectable cases, or reported outcomes for resectable cases separately and were included in the NAT arm of the Markov model (Appendix M).

15 out of these 18 studies used neoadjuvant chemoradiotherapy. 12 neoadjuvant chemotherapy regimes were reported across the 18 studies with 3 studies offering comparison between neoadjuvant regimes. Gemcitabine featured in 9 neoadjuvant chemotherapy regimes either used in isolation (n=1 study) or combined with: radiotherapy only (n=6 studies), cisplatin (n=3 studies; cisplatin and radiation n=1 study), oxaliplatin (n=1 study; oxaliplatin and radiation n=1 study), capecitabine (and radiation n=1 study; capecitabine and

docetaxel n=1 study) or cetuximab and radiation (n=1 study). The remaining 3 neoadjuvant chemotherapy regimes consisted of capecitabine and radiation (n=1 study), fluorouracil, cisplatin and radiation (n=3 studies) and paclitaxel and radiation (n=1 study). 4 of the 18 studies stated that patients who underwent resection in the NAT pathway also received adjuvant therapy post resection (Appenidx M).

For comparative analysis RCTs reporting outcomes of adjuvant therapy following surgery were included. Electronic database search identified 25332 studies. 15 studies were RCTs, 10 of which offered comparison between adjuvant regimes, 5 of which offered comparison between adjuvant therapy and surgery only (Appendix M).

For comparison of outcomes of resection rates, R0 resection rates, and rates of receiving adjuvant therapy a search was undertaken of studies offering comparison between upfront surgery and neoadjuvant therapy. A total of 14375 studies were identified through search of electronic databases, 452 of which underwent full text review (Appendix M). 25 studies were identified that offered comparison between neoadjuvant therapy and upfront surgery. In addition to the 9 phase II/III trials already identified that compared neoadjuvant therapy and upfront surgery, 16 cohort studies were identified and outcomes included in the SF arm. 6 of these studies were prospective and 10 of these studies were retrospective (Appendix M).

Model Evidence

Markov Model structure is shown in Figure 28. Table 19 outlines transition probabilities within the model derived from weighted pooled estimates of proportions calculated using Freeman-Tukey arcsine square root transformation under random effects model with corresponding 95% Confidence Intervals (Freeman & Tukey, 1950). The ranges in reported literature, standard deviation and variance were used to test uncertainty in model output through both deterministic and probabilistic sensitivity analysis.

Table 19: Transition Probabilities and Payoff Utility for Quality-Adjusted-Life-Months (QALMs)

Variable	**Transition Probability (95% Confidence Interval)**	**Range**	**Standard Deviation; Variance**	**Data distribution; Parameters; (Anderson Darling Statistic)**
Grade 3+ toxicity with NAT	0.41 (0.90-0.97)	0.70-1.00	0.09037; 0.00817	Gen. Pareto; k=0.16131 σ=0.06585 μ=-0.00512 (0.37908)
Resection in NAT pathway	0.63 (0.57-0.69)	0.32-0.85	0.02102; 4.4190E-4	Gen. Extreme Value; k=0.07104 σ=0.01585 μ=0.03134 (0.30431)
Exploratory Laparoscopy/Laparotomy	0.12 (0.08-0.17)	0-0.36	0.00633; 4.0057E-5	Johnson SB; γ=2.0682 δ=1.7897 λ=0.0624 ζ=-0.00855 (0.56039)
R0 resection NAT pathway	0.49 (0.36-0.62)	0.06-0.71	0.03079; 9.4797E-4	Cauchy; σ=0.013 μ=0.05608; (0.21049)
Grade 3-4 post-operative complication NAT pathway	0.19(0.13-0.26)	0.06-0.64	0.00457; 2.0931E-5	Gen. Extreme. Value; k=-0.32622 σ=0.0048 μ=0.01075 (0.27029)
Grade 5 post-operative complication NAT pathway	0.02(0.01-0.04)	0-0.12	0.00217; 4.7206E-6	Pareto 2; α=0.22134 β=4.0418E-13 (-6.8426)
Resection SF pathway	0.94 (0.90-0.96)	0.70-1.0	0.1219; 0.01486	Burr: k=0.0595 α=10.327 β=0.00112 (0.12818)
R0 resection SF pathway	0.56 (0.51-0.62)	0.16-0.86	0.09869; 0.00974	Pearson 5: α=0.61636 β=7.0460E-4 (0.18259)
Grade 3-4 post-operative complication SF pathway	0.22 (0.13-0.33)	0.04-0.54	0.01297; 0.0002	Log-Pearson 3: α=66.845 β=-0.09425 γ=2.0838 (0.29235)

Grade 5 post-operative complication SF pathway	0.07(0.02-0.13)	0-0.36	0.00948; 0.0002	Cauchy: σ=0.00373 μ= 0.00639 (0.38658)
Receiving adjuvant therapy	0.61(0.57-0.66)	0.26-0.94	0.10088; 0.01018	Burr: k=0.26048 α=2.145 β=9.2071E-4 (0.18949)
Adjuvant toxicity grade 3+	0.43(0.25-0.62)	0.09-0.98	0.02753; 0.00076	Log-Pearson 3: α=1916.0 β=-0.02672 γ=47.081 (0.34508)

Survival State	**Utility for QALM**
Living with stable pancreatic cancer	0.81
Undergoing chemo/radiotherapy	0.81
Experiencing chemo/radiotherapy complications	0.53
Recovering from pancreatic surgery	0.59
Experiencing surgical complications	0.48
Living with unresectable disease and pre-operative quality-of-life	0.65

Results of Markov Decision-Analysis

Intention-to-treat analysis of the treatment pathways showed that NAT pathway gave 26.41 months and 22.54 QALMs, compared to 23.72 months (18.51 QALMs) for SF pathway. The results of Markov cohort analysis are outlined in Table 20 and demonstrate superiority of NAT pathway for patients who received all treatment modalities.

Table 20: Results from Markov Cohort Analysis

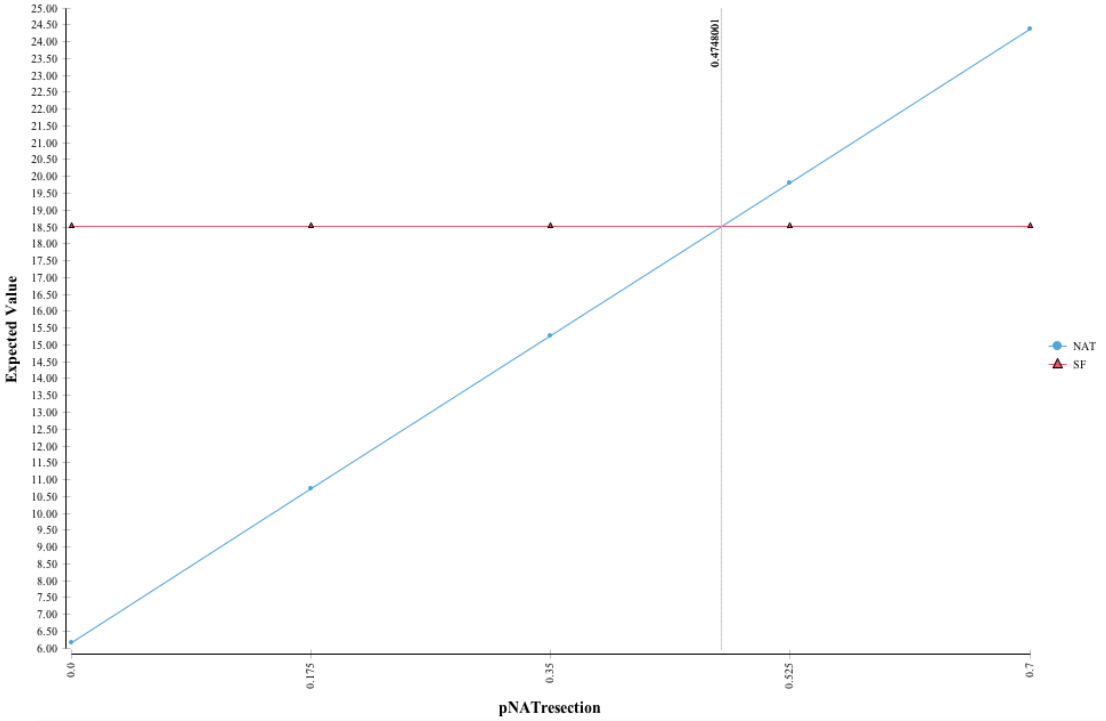
	NAT Pathway	SF Pathway
R0 Resection	39.34 months (34.63 QALMs; POC =34.52 QALMs)	Received Adjuvant Therapy: 30.96 months (24.86 QALMs; POC= 24.75 QALMs; AT= 21.82 QALMs; POC and AT=21.71 QALMs) No Adjuvant Therapy: 24.03 months (20.12 QALMs; POC=20.01QALMs)
R1 Resection	34.94 months (31.07 QALMs; POC=30.96QALMs)	Received Adjuvant Therapy: 25.85 months (20.72 QALMs; POC= 20.61 QALMs; AT= 18.20 QALMs; POC and AT=18.09 QALMs) No Adjuvant Therapy: 21.26 months (17.56 QALMs; POC=17.45 QALMs)
Exploratory Laparoscopy or Laparotomy	9.47 months (6.32QALMs)	10.48 months (6.97QALMs)
No Surgery	9.47 months (6.16 QALMs)	

NAT= Neoadjuvant Pathway; SF = surgery first/upfront surgery pathway; POC= post-operative complication grade 3 or 4; AT= adjuvant therapy resulting in grade 3 or 4 toxicity

Deterministic Sensitivity Analysis

One-way deterministic sensitivity analysis showed that NAT maintained superiority when all variables were individually altered between highest and lowest reported values with the exception of the probability of resection in NAT pathway, which had to be greater than 47.48% to maintain superiority (Figure 32).

Figure 32: One-way sensitivity analysis. Expected value on y-axis relates to months survival, pNATresection on x-axis refers to probability of undergoing surgical resection in NAT pathway.



Two-way deterministic sensitivity analysis demonstrated that treatment superiority depended on receiving multimodal treatment (surgical resection in the NAT pathway and adjuvant therapy in SF pathway). Figure 33a shows thresholds at which competing pathways offer superior outcomes with Figure 33b providing corresponding probability thresholds and predicted resulting quality-adjusted survival time.

Figure 33a: Two-way sensitivity analysis. Y-axis shows probability of receiving adjuvant therapy in the surgery first (SF) pathway and x-axis shows probability of receiving resection in the neoadjuvant (NAT) pathway

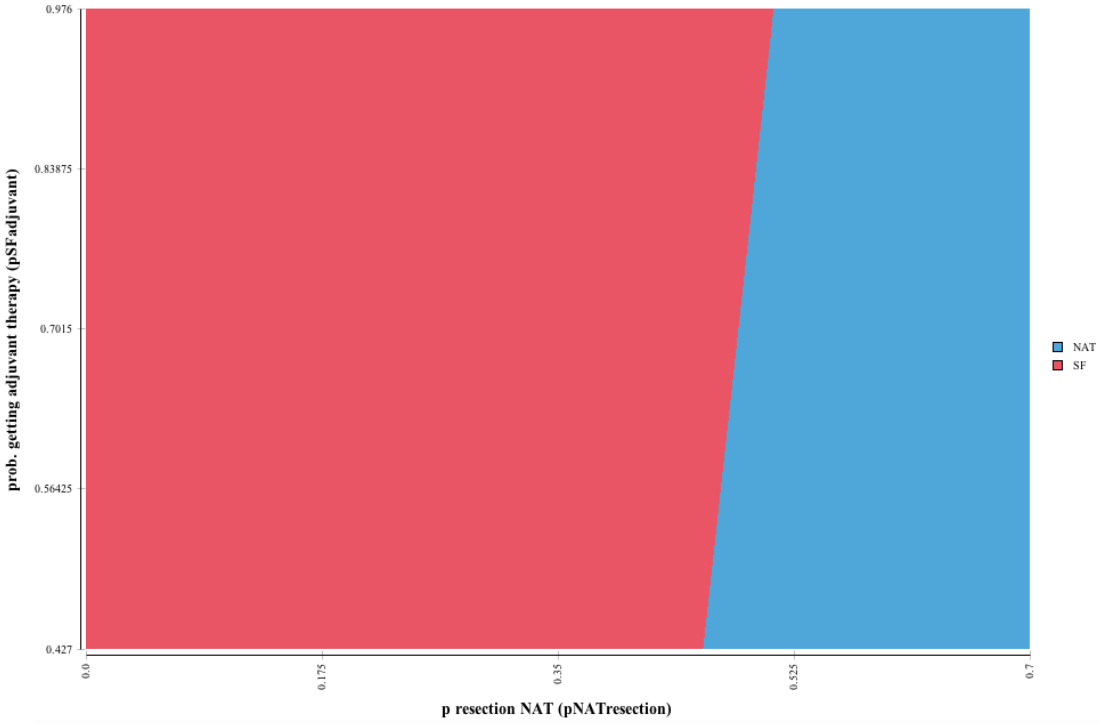


Figure 33b: predicted QALMs given altering probabilities of receiving adjuvant therapy in SF pathway (y-axis) and resection in NAT pathway (x-axis).

Two Way Sensitivity Analysis

prob. getting adjuvant therapy \ p resection NAT	0.0	0.175	0.35	0.525	0.7
▼0.427					
SF	18.06	18.06	18.06	18.06	18.06
NAT	6.18	10.72	15.27	19.82	24.36
▼0.56425					
SF	18.40	18.40	18.40	18.40	18.40
NAT	6.18	10.72	15.27	19.82	24.36
▼0.7015					
SF	18.74	18.74	18.74	18.74	18.74
NAT	6.18	10.72	15.27	19.82	24.36
▼0.83875					
SF	19.08	19.08	19.08	19.08	19.08
NAT	6.18	10.72	15.27	19.82	24.36
▼0.976					
SF	19.42	19.42	19.42	19.42	19.42
NAT	6.18	10.72	15.27	19.82	24.36

Monte Carlo Probabilistic Sensitivity Analysis

To test the level of confidence in the Markov model output, Monte Carlo probabilistic sensitivity analysis was performed to simulate 10000 patients cycling through the model, with input data for each model variable drawn from the data distribution of that individual variable (Table 19).

The results of probabilistic Monte Carlo sensitivity analysis showed that NAT pathway gave a mean survival time of 22.54 months (range 20.25-24.55months) compared to 18.50 months (range 7.24-20.58months) for SF pathway. Standard deviation was 0.56 and variance 0.31 for the NAT arm of the model. Standard deviation was

1.56 and variance 2.45 for the SF arm of the model. When minimum significant difference threshold between treatment pathways was set to 3months, the model reported that NAT was the superior pathway with a superior pathway selection frequency of 78%, and 22% frequency of indifference between pathways.

4.2.3 Triangulation of Findings: Surgery first versus neoadjuvant pathway for resectable pancreatic ductal adenocarcinoma (PDAC): a Markov decision analysis based on a tertiary referral center's 7-year experience

Patients and Treatment Strategies

Probabilities of interventions, clinical outcomes, and survival in both SF and NAT cohorts were calculated from the West of Scotland Pancreatic Unit database which recorded data prospectively for a cohort of 200 sequential patients diagnosed with non-metastatic pancreatic cancer and who were deemed fit for surgery. All patients underwent surgery in the West of Scotland Pancreatic Unit. SF pathway was exclusively performed from January 2008 to July 2012. From 1st August 2012-30th December 2015 100 patients with non-metastatic PDAC were treated in the NAT pathway. For this model, only those patients with resectable PDAC on completion of initial staging prior to commencing NAT were included. Borderline and locally advanced PDAC, as determined according to AHPBA/SSO/SSAT guidelines (Callery *et al*., 2009a) were excluded. From August 2012 working backwards, 100 sequential patients in SF

pathway who had resectable PDAC, and were deemed fit for surgery based on performance status score and formal cardiovascular fitness testing (CPET), populated the SF arm of the model. No patients were lost to follow-up (Table 21).

Table 21: Baseline Patient Characteristics

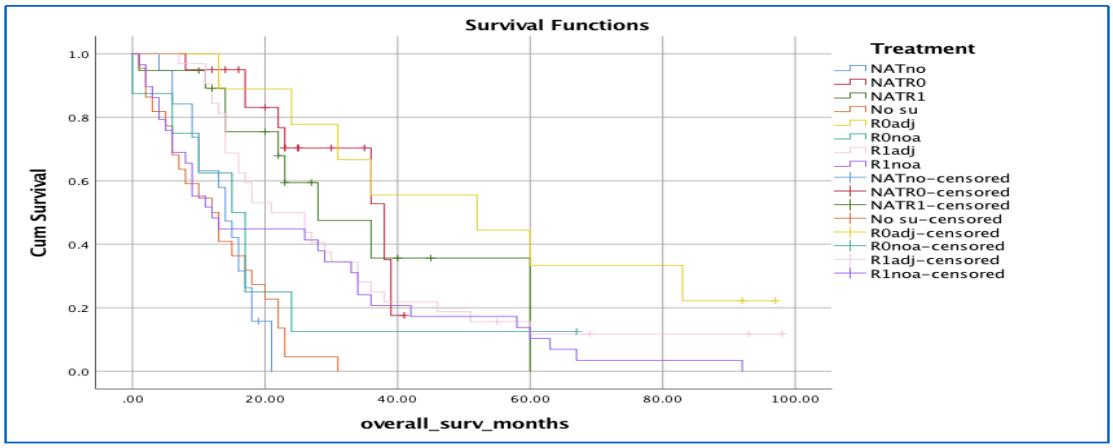
	Neoadjuvant Cohort (n=59)	Surgery First Cohort (n=100)
Age	64 (range 49-80)	63 (range 38-86)
Gender Male Female	 33 (55.9%) 26 (44.1%)	 58 (58%) 42 (42%)
ASA Grade 1 2 3	 4 (6.8%) 37 (62.7%) 18 (30.5%)	 10 (10%) 66 (66%) 24 (24%)
Tumor Site Head Neck/Body/Tail	 51 (86.4%) 8 (13.6%)	 87 (87%) 13 (135)
Resection	39 (66.1%) R0: 19 (48.7%)	78 (78%) R0: 16 (20.5%)
Exploratory Laparotomy or Bypass Surgery	6 (10.2%)	22 (22%)
Completed Adjuvant Therapy		50 (50%)

Transition nodes were based on outcomes of response to neoadjuvant therapy on re-staging CT scan (for NAT cohort only), operative intervention and outcome, post-operative complications, and receipt of adjuvant therapy. Pathological stage was defined by AJCC staging system 7th edition and complications were graded according to the Clavien-Dindo system and categorised as grade 3 or above or grade 2 and below. Utility was defined as QALMs and was

calculated from the time spent in each of the Markov states based on overall and disease free survival calculated from median survival time taken from Kaplan-Meir survival analysis based on treatment received (Figure 34; Table 22).

Figure 34: Kaplan-Meir Survival Analysis: a) overall survival time b) disease-free survival

a)



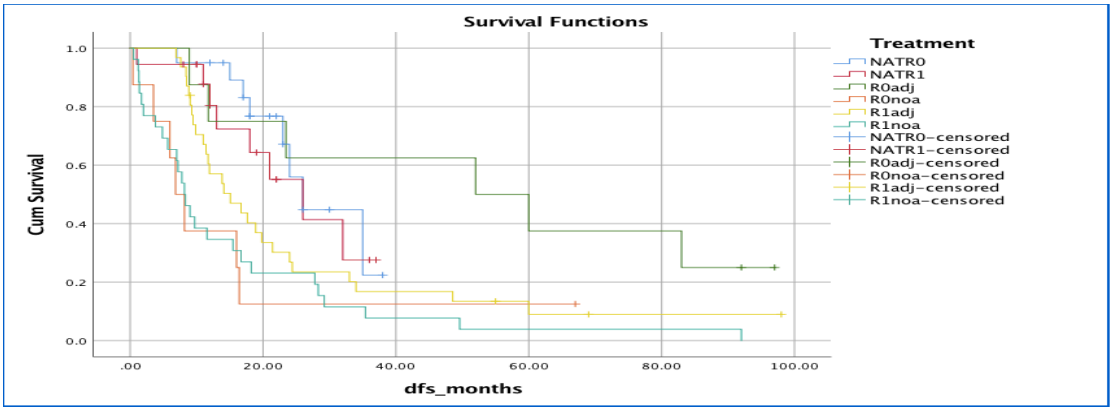
Means and Medians for Survival Time

	Mean[a]				Median			
			95% Confidence Interval				95% Confidence Interval	
Treatment	Estimate	Std. Error	Lower Bound	Upper Bound	Estimate	Std. Error	Lower Bound	Upper Bound
NATno	13.579	1.217	11.194	15.963	14.000	1.451	11.156	16.844
NATR0	32.376	2.444	27.585	37.167	38.000	7.874	22.567	53.433
NATR1	35.192	6.115	23.206	47.178	28.000	8.204	11.921	44.079
No su	12.682	1.800	9.154	16.210	12.000	2.932	6.254	17.746
R0adj	54.778	9.943	35.290	74.266	52.000	23.851	5.251	98.749
R0noa	19.500	6.803	6.166	32.834	15.000	3.300	8.532	21.468
R1adj	32.734	4.798	23.330	42.139	21.000	5.091	11.021	30.979
R1noa	24.345	4.487	15.550	33.139	12.000	3.588	4.968	19.032
Overall	28.530	2.220	24.179	32.881	20.000	1.770	16.532	23.468

a. Estimation is limited to the largest survival time if it is censored.

NAT= neoadjuvant pathway; NATno= no surgery in neoadjuvant pathway; SF= surgery first pathway; No su= no surgery/exploratory laparotomy or laparoscopy in surgery first pathway; noa= no adjuvant therapy received; adj= adjuvant therapy received

b)



Means and Medians for Survival Time

	Mean[a]				Median			
			95% Confidence Interval				95% Confidence Interval	
Treatment	Estimate	Std. Error	Lower Bound	Upper Bound	Estimate	Std. Error	Lower Bound	Upper Bound
NATR0	27.539	2.509	22.622	32.457	26.000	2.762	20.587	31.413
NATR1	24.287	3.059	18.292	30.282	26.000	5.780	14.671	37.329
R0adj	54.140	12.074	30.475	77.805	52.000	25.838	1.358	102.642
R0noa	15.543	7.120	1.588	29.497	6.830	1.556	3.781	9.879
R1adj	25.669	4.805	16.250	35.088	15.080	3.232	8.746	21.414
R1noa	15.536	3.897	7.898	23.174	8.180	1.147	5.931	10.429
Overall	29.314	3.161	23.119	35.509	18.000	1.765	14.540	21.460

a. Estimation is limited to the largest survival time if it is censored.

NAT= neoadjuvant pathway; NATno= no surgery in neoadjuvant pathway; SF= surgery first pathway; No su= no surgery/exploratory laparotomy or laparoscopy in surgery first pathway; noa= no adjuvant therapy received; adj= adjuvant therapy received

The neoadjuvant regimen, previously described by Grose *et al*. (2017), was modified FOLFIRINOX which consisted of: Oxaliplatin 85mg/m^2, Irinotecan 180mg/m^2 and 5-Flurouracil 400mg/m^2 on day 1 and 5-Flurouracil 2,400mg/m^2 on days 1-2. If patients had: poor performance status, or were aged over 70 years, or FOLFIRINOX was poorly tolerated, they received Gemcitabine 1,000 mg/m^2 on days 1,8 and 15 and Capcitabine 830 mg/m^2 BD (GemCap) on days 1 to 21. A re-staging CT of the chest, abdomen and pelvis was performed 3 months post chemotherapy (6 cycles FOLFIRINOX or 3 cycles

GemCap). For patients who had response or stable disease on re-staging CT approximately 4 weeks post completion of chemotherapy, chemoradiation (CRT) comprising Volumetric Modulated Arc Therapy was given at a dose of 50.4Gy in 28 fractions over 5.5 weeks with concurrent Capecitabine. Adjuvant therapy regime in the SF group was Gemcitabine monotherapy.

Ethical approval for data collection was granted by the West of Scotland Local Research Ethics Committee.

Quantifying the Limitations Stemming from FUPS Characteristics of Data

One and two-way deterministic sensitivity analysis determined which variables had the greatest potential to affect model outcomes by altering values between highest lowest observed values. Monte Carlo probabilistic sensitivity analysis set to 10000 cycles assessed the overall effect of uncertainty within the model. For this analysis the model probabilities were sampled from the data distribution of each variable contained within the model (Table 22).

Table 22: Transition Probabilities and Payoff Utility for Quality-Adjusted-Life-Months (QALMs)

Variable	**Transition Probability**	**Variance**	**Standard Deviation**	**Data Distribution: parameters (Anderson Darling Statistic)**
Grade 3+ toxicity with NAT	0.22	0.21607	0.0046483	D. Uniform: a=0 b=1 (34.26)
Resection in NAT pathway	0.66	0.29876	0.0054659	Poisson: λ=0.84746 (23.333)
Exploratory Laparoscopy/Laparotomy	0.10	0.29876	0.0054659	Poisson: λ=0.84746 (23.333)
No surgery	0.24	0.29876	0.0054659	Poisson: λ=0.84746 (23.333)
R0 resection NAT pathway	0.49	0.24984	0.0049984	Poisson: λ=0. 51282 (19.177)
Grade 3-4 post-operative complication NAT pathway	0.18	0.25895	0.0050887	D. Uniform: a=0 b=1 (19.928)
Grade 5 post-operative complication NAT pathway	0.03	0.25895	0.0050887	D. Uniform: a=0 b=1 (19.928)
Resection SF pathway	0.78	0.1716	0.0041425	Poisson: λ=1.22 (42.381)
R0 resection SF pathway	0.21	0.16305	0.004038	Bernoulli: p=0.79487 (16.14)
Grade 3-4 post-operative complication SF pathway	0.27	0.29692	0.005449	Poisson: λ= 0.34884 (50.791)
Grade 5 post-operative complication SF pathway	0.04	0.29692	0.005449	Poisson: λ= 0.34884 (50.791)
Receiving adjuvant therapy	0.50	0.25	0.005	Poisson: λ=0.5 (38.757)
Adjuvant toxicity grade 3+	0.36	0.24377	0.0049373	Bernoulli: p=0.57895 (33.585)

Survival State	**Utility for QALM**
Living with stable pancreatic cancer	0.81
Undergoing chemo/radiotherapy	0.81
Experiencing chemo/radiotherapy complications	0.53

Recovering from pancreatic surgery	0.59
Experiencing surgical complications	0.48
Living with unresectable disease and pre-operative quality-of-life	0.65

Results

In total 100 patients were included in the SF arm and 59 patients were included in the NAT arm of the Markov model. Baseline characteristics of both patient groups are described in Table 21. The median age was 64 years (range 38-86). The majority were male (n= 91; 57.2%). Most had an American Society of Anesthesiology (ASA) grade 2 (n= 103; 64.8%). 8 patients in the SF arm achieved R0 resection and completed adjuvant therapy.

Markov Decision-Analysis

On an intention-to-treat basis, analysis of the treatment pathways showed NAT pathway gave 32.90 months (28.51 QALMs) compared to 24.68 months (19.23 QALMs) for SF pathway. Markov cohort analysis examines the utility outcomes when cohorts are determined by the treatments received by patients. The results of Markov cohort analysis are outlined in Table 23 and demonstrate that greatest utility was achieved in the small subgroup of patients who received early R0 resection and adjuvant therapy.

Table 23: Results from Markov Cohort Analysis

	NAT Pathway	SF Pathway
R0 Resection	45.36 months (40.86 QALMs; POC =40.75 QALMs)	Received Adjuvant Therapy: 52.59 months (42.38 QALMs; POC= 42.27 QALMs; AT= 32.59 QALMs; POC and AT= 32.48 QALMs) No Adjuvant Therapy: 22.31 months (19.06 QALMs; POC=18.95 QALMs)
R1 Resection	42.29 months (38.38 QALMs; POC=38.27 QALMs)	Received Adjuvant Therapy: 33.37 months (26.81 QALMs; POC= 26.70 QALMs; AT= 22.81 QALMs; POC and AT=22.70 QALMs) No Adjuvant Therapy: 20.93 months (18.20 QALMs; POC=18.09 QALMs)
Exploratory Laparoscopy or Laparotomy	14.31 months (9.46 QALMs)	12.42 months (8.23 QALMs)
No Surgery	14.31 months (9.30 QALMs)	

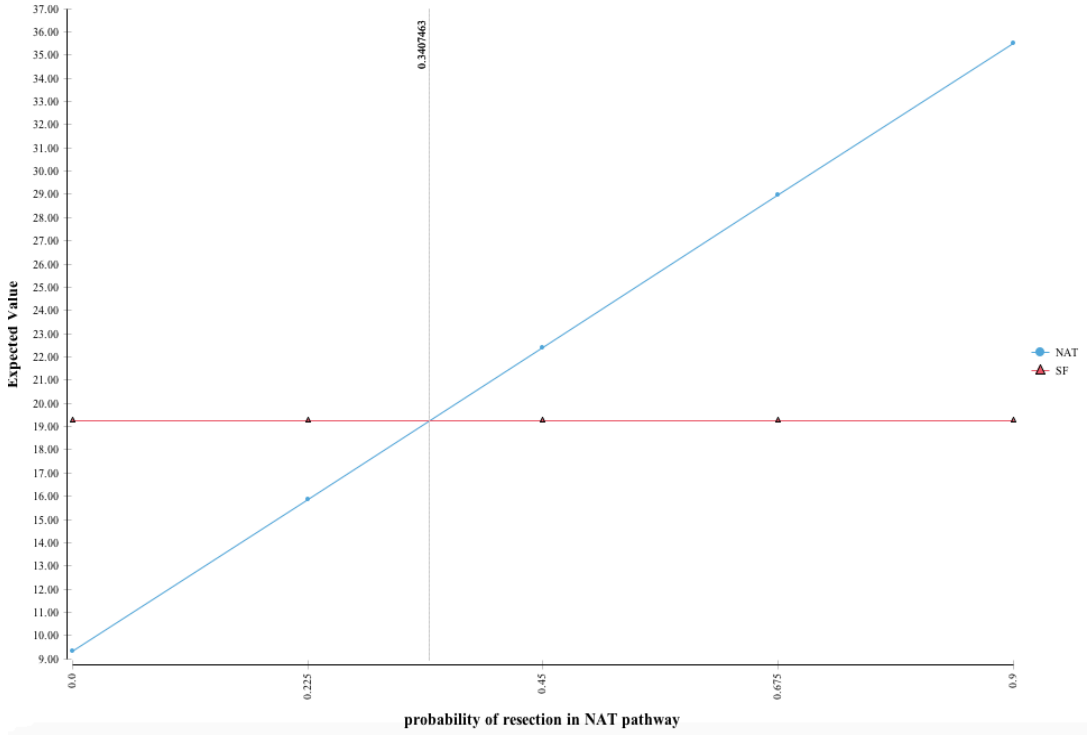
POC= post-operative complication grade 3 or 4; AT= adjuvant therapy resulting in grade 3 or 4 toxicity

Deterministic Sensitivity Analysis:

To test the sensitivity of the model's results to variations in parameters of specific variables, deterministic sensitivity analysis was performed by altering the variable parameters between highest and lowest reported values.

One-way deterministic sensitivity analysis showed that when all variables were altered between highest and lowest reported values, NAT maintained superiority with the exception of the probability of undergoing resection in the NAT pathway, which had to be above 34% to maintain superiority (Figure 35).

Figure 35: One-way sensitivity analysis. Y-axis relates to expected survival time in months. X-axis refers to probability of undergoing surgical resection in NAT pathway.



Two-way deterministic sensitivity analysis demonstrated that treatment superiority depended on receiving multimodal treatment (resection in NAT pathway and adjuvant therapy in SF pathway) (Figure 36) and on achieving R0 resection in SF pathway (Figure 37). 3-way deterministic sensitivity analysis was therefore undertaken to test the hypothesis that SF was the superior pathway for patients most likely to achieve R0 resection and receive adjuvant therapy. This involved assessing the probability of receiving adjuvant therapy in SF pathway against the probability of receiving resection in NAT

pathway, then increasingly the probability of R0 resection in SF pathway from 0 to 100% in quartiles. This demonstrated that whilst NAT pathway was superior for most patients likely to receive resection within this pathway, as the probability of R0 resection increased within SF pathway, particularly for those most likely to receive adjuvant therapy, so too did the likelihood of SF pathway being the superior pathway choice.

Figure 36 a): Markov Two-way sensitivity analysis. Y-axis shows probability of receiving adjuvant therapy in surgery first (SF) pathway and x-axis shows probability of receiving resection in neoadjuvant (NAT) pathway. The red area depicts the range whereby SF pathway would be the superior pathway. The blue area depicts the range over which NAT pathway would be the superior pathway.

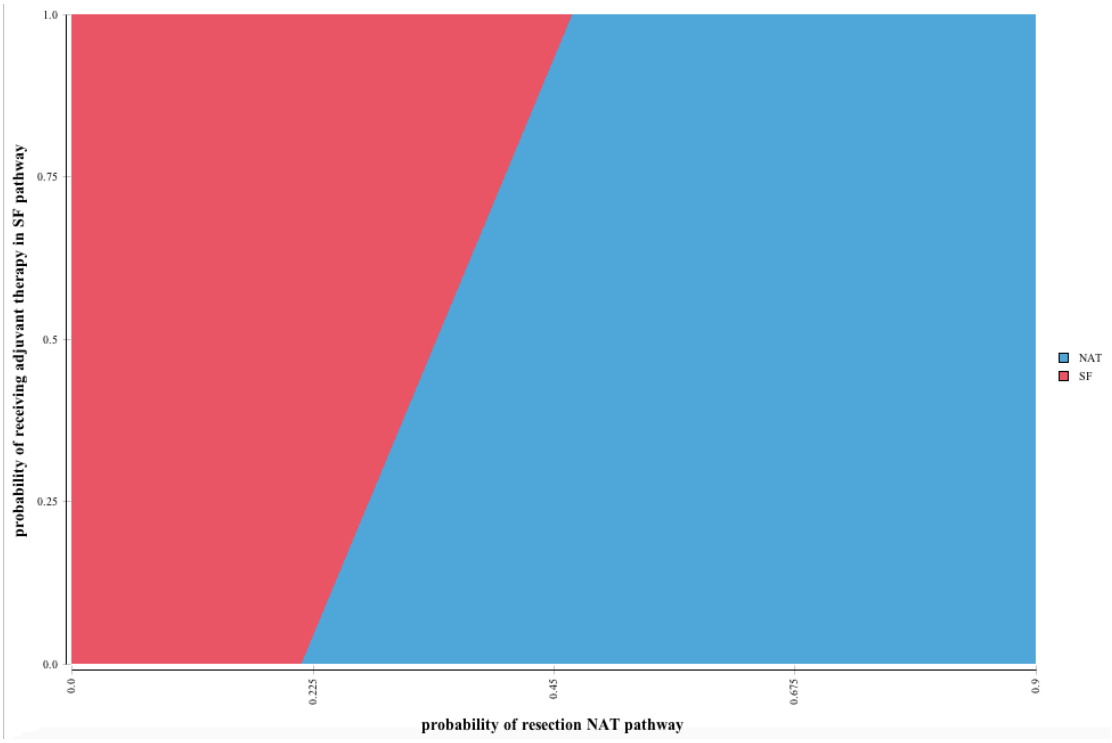


Figure 36b): predicted QALMs given altering probabilities of receiving adjuvant therapy in SF pathway (y-axis) and resection in NAT pathway (x-axis).

prob. getting adjuvant therapy \ p resection NAT	0.0	0.225	0.45	0.675	0.9
▼0.0					
SF	15.55	15.55	15.55	15.55	15.55
NAT	9.32	15.86	22.41	28.95	35.50
▼0.25					
SF	17.39	17.39	17.39	17.39	17.39
NAT	9.32	15.86	22.41	28.95	35.50
▼0.5					
SF	19.23	19.23	19.23	19.23	19.23
NAT	9.32	15.86	22.41	28.95	35.50
▼0.75					
SF	21.07	21.07	21.07	21.07	21.07
NAT	9.32	15.86	22.41	28.95	35.50
▼1.0					
SF	22.91	22.91	22.91	22.91	22.91
NAT	9.32	15.86	22.41	28.95	35.50

Figure 37 a): Two-way sensitivity analysis. Y-axis shows probability of receiving R0 resection in surgery first (SF) pathway and x-axis shows probability of receiving resection in neoadjuvant (NAT) pathway. The red area depicts the range whereby SF pathway would be the superior pathway. The blue area depicts the range over which NAT pathway would be the superior pathway.

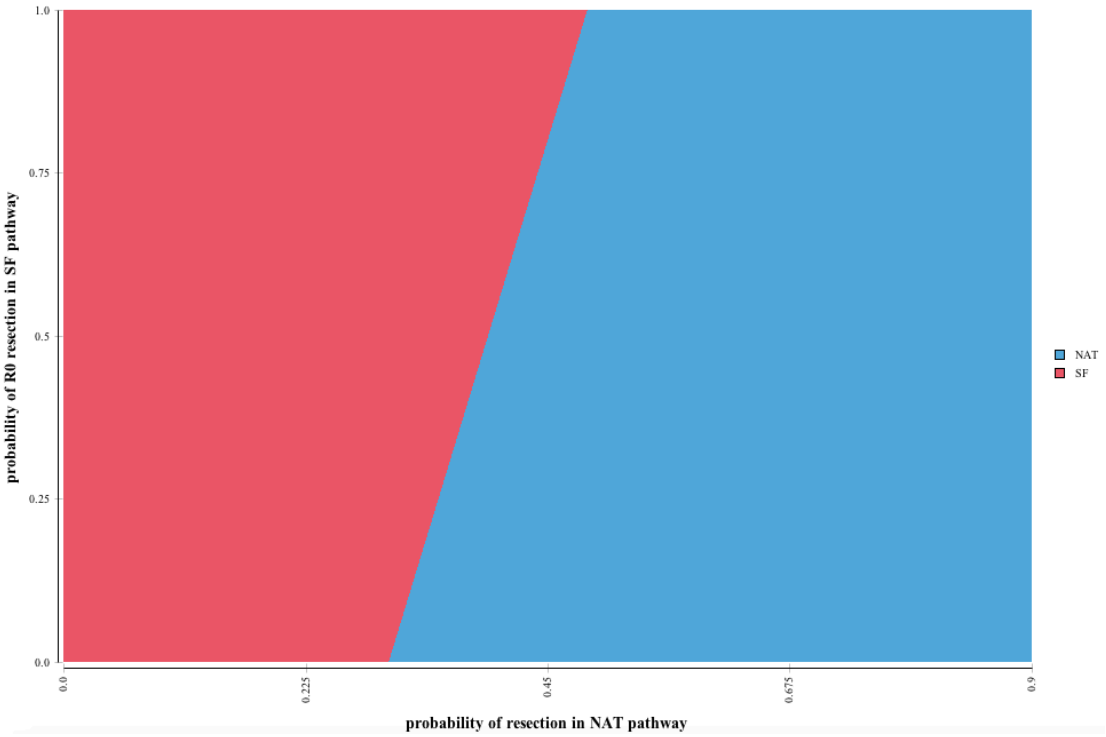


Figure 37 b): predicted QALMs given altering probabilities of receiving R0 resection in SF pathway (y-axis) and resection in NAT pathway (x-axis).

p.R0 resection in SF \ p resection NAT	0.0	0.225	0.45	0.675	0.9
▼0.0					
SF	18.10	18.10	18.10	18.10	18.10
NAT	9.32	15.86	22.41	28.95	35.50
▼0.25					
SF	19.44	19.44	19.44	19.44	19.44
NAT	9.32	15.86	22.41	28.95	35.50
▼0.5					
SF	20.78	20.78	20.78	20.78	20.78
NAT	9.32	15.86	22.41	28.95	35.50
▼0.75					
SF	22.13	22.13	22.13	22.13	22.13
NAT	9.32	15.86	22.41	28.95	35.50
▼1.0					
SF	23.47	23.47	23.47	23.47	23.47
NAT	9.32	15.86	22.41	28.95	35.50

Probabilistic Monte Carlo Sensitivity Analysis

Probabilistic Monte Carlo sensitivity analysis was set to simulate 10000 patients cycling through the model. This analysis was performed to test the level of confidence in the model's output. (Table 22). During Monte Carlo simulation probabilities were therefore drawn from the data distribution of each variable, which meant that all possible parameter values for each variable within the model was tested.

The results of probabilistic Monte Carlo sensitivity analysis showed that NAT pathway gave a mean survival time of 27.21 QALMs (range:26.43-27.70 QALMs) compared to 20.37 QALMs (range:11.41-30.62 QALMs) for SF pathway with standard deviation 0.21 and 2.74, and variance 0.04 and 7.51 in NAT and SF pathways respectively. Strategy selection frequency was 99.41% for NAT pathway and 0.59% for SF pathway when no minimum significant difference between pathways was set. When minimum significant difference was set at 6months, the model reported a selection frequency of 62.35% for NAT pathway and indifference between pathways selected with a frequency of 37.65%.

Discussion

Data As A Partial Remnant

The role of NAT in treatment of pancreatic cancer is an ongoing area of debate (Tempero *et al*., 2014). Although NAT is supported by current guidelines for borderline resectable and locally advanced pancreatic cancer, optimal treatment of resectable pancreatic cancer remains controversial (deGeus *et al*., 2016; Tempero *et al*., 2014). Markov decision analysis is a powerful tool offering analysis of complex medical decisions therefore this study utilises current evidence in a Markov decision-analysis model, and triangulates findings with an institutional database to offer an important interim source of information to inform the ongoing debate regarding the best treatment approach for resectable pancreatic cancer (deGeus *et al*., 2016; VanHouten *et al*., 2012). Importantly by utilising FUPS data

within the context of complexity this research adds a further dimension to this debate by demonstrating for the first time that optimal treatment selection depends on individual patient and tumor factors.

The Markov decision-analysis based on synthesised data from trials including all cases of potentially resectable pancreatic cancer showed that SF pathway gave an additional 3.5 months (2.25 QALMs) but neither pathway was conclusively superior. When the Markov decision-analysis model was populated with synthesised data from trials that included only cases that were resectable at presentation NAT pathway gave an additional 4.03 QALMs (26.41 months (22.54 QALMs) *versus* 23.72 months (18.51 QALMs) for SF pathway. When these findings were triangulated by populating the model with data from an institutional patient database containing only cases that presented with resectable disease the tendency towards superiority of NAT pathway was upheld. NAT pathway gave an additional 9.29 QALMs (32.90 months *versus* 24.68 months and 28.51 QALMs *versus* 19.23 QALMs). Markov cohort analysis of outcomes where multimodal treatment was received in both pathways (resection in NAT pathway and adjuvant therapy in SF pathway) demonstrated superior outcomes with NAT pathway when populating the model with synthesised data. Yet the base case probability of undergoing resection in the NAT pathway was consistently lower than in the SF pathway across all Markov decision analysis models. This could demonstrate that NAT allowed a period of time for more aggressive tumors, for which surgery would be ultimately futile, to declare themselves (Asare *et al*., 2016; Lee *et al*., 2016). Conversely

opponents of NAT would argue that this demonstrates losing the window of resectability (Asare *et al*., 2016; Lee *et al*., 2016).

Deterministic and Monte Carlo probabilistic sensitivity analysis allowed a deeper level of engagement with the decision-making process being modeled as well as a more detailed analysis of the impact of uncertainty associated with the FUPS characteristics of the data. This revealed that superior pathway selection was actually determined by individual patient and tumour factors affecting the probability of receiving multimodal treatment. This raised the possibility that although NAT would benefit most patients with potentially resectable as well as resectable pancreatic cancer, SF pathway could have maximum utility for those patients with the earliest stages of the disease who had the highest probability of receiving adjuvant therapy.

This hypothesis was tested further when modeling institutional data for resectable only cases. This revealed that cumulative utility was greatest for the minority subgroup of patients who received R0 resection followed by adjuvant therapy in the SF pathway. Furthermore, this is the first study that has been able to mathematically demonstrate the degree to which altering the probabilities of receiving multimodal treatment in competing pathways, whilst simultaneously altering the probability of achieving R0 resection in the SF pathway, affects selection of superior treatment pathway at an individualised level. This marks an important step in widening the current debate beyond SF *versus* NAT

management pathways towards achieving personalised predictive medicine and delivering personalised pancreatic cancer care with the associated benefits to patients and for resource allocation and utilisation.

Triangulation

These findings are in keeping with the few existing meta-analysis comparing NAT and SF treatment approaches, which report modest superior survival benefit with NAT pathway (Versteijne *et al*., 2018). The results of our Markov cohort analysis corroborate findings of prospective and retrospective experiences of NAT, which have demonstrated favorable survival outcomes ranging from 26 to 45 months (deGeus *et al*., 2016; Talamonti *et al*., 2006; Faith *et al*., 2015; Kharofa *et al*., 2014). Three Markov Decision-Analysis studies comparing NAT and SF pathways for pancreatic cancer exist (Sharma *et al*., 2015; deGeus *et al*., 2016; VanHouten *et al*., 2012). One of these studies focuses on potentially resectable pancreatic cancer, therefore including borderline and locally advanced cases in the NAT pathway to capture the effect of conversion to resectability on overall pathway analysis (VanHouten *et al*., 2012). As with this study's findings it did not demonstrate an overall conclusively superior pathway on an intention-to-treat basis (NAT 18.6 months *versus* 17.1 months) (VanHouten *et al*., 2012). Two other Markov decision analysis studies have reported superiority of NAT pathway but they focus on resectable only cases and based their model on literature from a single search engine (deGeus *et al*., 2016; Sharma *et al*., 2015). One

study, based on phase I/II trials, reported 22 months for NAT *versus* 20 months for SF (Sharma *et al*., 2015). De Geus *et al.* (2016) demonstrated larger survival gain with NAT (32.2 versus 26.7 months) but their analysis mostly included retrospective studies. Preliminary results from the PREOPANC-1 trial, a multicenter phase III RCT comparing NAT and SF for borderline resectable pancreatic cancer, had reported improved survival with NAT on an intention-to-treat basis (17.1 months *versus* 13.5 months) but full results have demonstrated no statistically significant difference in terms of overall survival (Van Tienhoven *et al.,* 2018). Although PREOPANC-1 is a very different study to the one presented here in terms of design and statistical methodology, the results do echo our findings in reporting that higher reported resection rates in the SF pathway do not equate with superior overall survival time for patients treated in a SF pathway. Furthermore the subgroup analysis of resected cases in the PREOPANC-1 trial reported superior survival time with NAT (29.9months *versus* 16.6months), which further corroborates the results our Markov cohort analysis. Preliminary results from the Prep-02/JSAP-05 randomised phase II/III trial comparing NAT *versus* SF for resectable pancreatic cancer also corroborate our findings, reporting an overall survival time of 36.7months with NAT compared to 26.6months for SF approach (Unno et al., 2019).

Transparency of Analysis

Like existing Markov decision analysis based on data from published studies, this study also shares the limitations of the existing body of

evidence: heterogeneity, lack of randomisation, potential bias, small and underpowered studies (Asare *et al*., 2016; Lee *et al*., 2016). Furthermore definitions of radiological and surgical resectability, R0 resection and staging protocol can vary across trials further confounding the issues of heterogeneity of synthesised data (Versteijne *et al*., 2018). Such heterogeneity could account for why, at base-case analysis, the probability of R0 resection in the NAT pathway was smaller than anticipated particularly when considering that the PREOPANC-1 trial has reported higher rates of R0 resection (65%) in the NAT arm of their trial (Versteijne *et al*., 2018; VanTienhoven *et al*., 2018). Although the uncertainty of this variable was extensively tested through sensitivity analysis and found not to affect the overall model outcomes, this highlights the potential impact of heterogeneity of data on model output. To address the issue of heterogeneity, this study based transition probabilities on weighted pooled proportion estimates calculated using Freeman-Tukey arcsine square root transformation under random effects model (Freeman & Tukey, 1950). Furthermore probabilistic Monte Carlo sensitivity analysis sampled model probabilities from the entire range of the data distribution and provided assessment of the extent of variance and standard deviation within the model. Unlike previous Markov decision-analysis, weighted survival times were based on the Gillen *et al.* (2010) formulae as evidence has shown that unbiased pooled estimates of median survival times cannot be achieved by weighted averaging of medians (Gillen *et al*., 2010; Rouder *et al*., 2004). Quality adjusted survival time is limited by the lack of studies measuring quality-of-life across the treatment trajectory for pancreatic cancer. This study utilised best available

data that was shared with existing decision-analysis studies, which enhanced comparability. Our model did not assume return to full health after an intervention or event but accounted for the impact of therapy, surgery, complications and disease recurrence when calculating quality-adjusted survival. For both survival and quality-adjusted survival, uncertainty was rigorously tested across every variable in the model through probabilistic and deterministic sensitivity analysis.

The majority of studies meeting the inclusion criteria for the NAT arm of the model were gemcitabine based. Neoadjuvant mFOLFIRINOX is increasingly being used therefore, accepting this study as providing important interim analysis, we envisage that as the body of evidence surrounding mFOLFIRINOX matures it would be pertinent in the future to repeat this study to assess the impact of both efficacy and toxicity of emerging neoadjuvant regimes.

Conclusion

In conclusion the Markov decision analysis of SF and NAT pathways for the management of potentially resectable and resectable pancreatic cancer demonstrated marginal superiority of NAT pathway on an intention-to-treat basis. However, patients with early resectable disease who are most likely to receive adjuvant therapy could benefit from SF pathway. By engaging with the complexity of the systems being modeled, these findings have helped to evolved the contemporary narrative beyond SF *versus* NAT. Optimal treatment selection depends on receiving all treatment modalities (resection in

both pathways and adjuvant therapy in SF pathway). Yet, on an intention-to-treat basis a significant number in each pathway failed to receive all intended treatment modalities. What has begun to emerge from these findings is that individual patient and tumour factors interact in a dynamic way that determines the probability of receiving multimodal treatment hence determining the selection of the optimal treatment pathway at an individual level.

This highlights two important directions for future research based on our Markov decision analysis: 1) exploring methods of predictive statistical modeling to identify patients who are more likely to receive and benefit from differing treatment modalities and 2) cost-effectiveness analysis of neoadjuvant *versus* upfront surgery. The following sections will therefore take the analysis further through DES modeling to determine whether microsimulation can uncover what anticipated outcomes individuals who did not receive multimodal treatment could have expected if they were treated in the competing pathway. Specifically this will seek to explore whether those who failed to undergo resection in the NAT pathway lost the window of resectability or were successfully filtered away from futile surgery. This will also allow the results of Markov modeling to be further triangulated with DES modeling methods. Section 4.2 and 4.3 therefore will pave the way for cost-effectiveness analysis of SF *versus* NAT for resectable pancreatic cancer using both the Markov and DES models developed in these sections.

4.3. Discrete Event Simulation Decision (DES) Analysis

Publication resulting from this analysis:

Bradley, A., Van Der Meer, R., McKay, C.J. (2020) 'Computer simulated comparison of neoadjuvant versus upfront surgery for resectable pancreatic cancer: the application of machine-learning algorithms to support personalised decision-making'. *British Journal of Surgery*: accepted

Abstract

Background: Pancreatic cancer is a challenging malignancy with poor survival outcomes. Markov decision-analysis modeling has previously suggested that superior treatment pathway selection depends on individual patient and tumour factors. However this approach carries methodological limitations due to the memory-less property of Markov models and the depletion of susceptibles that could affect the accuracy of model output. The aim of this study was to triangulate the results of Markov decision analysis with that of DES decision analysis and test the hypothesis that this approach to microsimulation modeling could reveal new insights regarding individualised treatment pathway selection.

Methods: A combined systems level approach and DES model was created to established how alternative options for treatment pathway delivery affects outcomes at individual patient level

depending on disease stage at presentation. Model delivery options included: 1) surgery first pathway for resectable cases and neoadjuvant pathway for borderline resectable/ locally advanced cases 2) neoadjuvant pathway for resectable and borderline resectable/ locally advanced cases 3) surgery first pathway for resectable cases and palliative pathway for all other stages of disease and 4) subgroup analysis of surgery first versus neoadjuvant pathway for resectable only cases. The model was populated with data from randomised controlled and prospective phase II/III trials. Model uncertainties were tested through probabilistic sensitivity analysis whereby transition probabilities were drawn from the entirety of data distribution for each parameter contained within the model over 10000 iterations. The data distributions were varied reflect the disease stage of each simulated patient entering the model.

Results: Overall the greatest benefit was seen in borderline cases treated in the neoadjuvant pathway (13.92months (10.98 QALMs) or 13.93 months (10.89 QALMs) in options 1 and 2 respectively) compared to the palliative chemotherapy pathway modeled in option 3 (8.50months; 6.41 QALMs). For resectable cases there was a marginal overall survival advantage with option 2, the neoadjuvant pathway (20.02months; 17.16 QALMs versus 17.49months; 13.11 QALMs in options 1 and 3). Subgroup analysis showed that for resectable only cases the neoadjuvant pathway gave a mean survival time of 20.01months (18.45 QLAMs) compared to 16.55months (14.19 QALMs) in the upfront surgery pathway. A minimum significant difference threshold of below 3.5months favoured

neoadjuvant therapy but at this threshold or above the selection frequency was 40.6% for neoadjuvant pathway and 59.4% for indifference between pathways demonstrating minimal difference between pathways. Furthermore within the subgroup analysis unresected cases in the neoadjuvant pathway had an expected incremental value of between 1.5 to 5.5months over a resection probability range of 47-94% had they been treated in the upfront surgery pathway. Threshold analysis showed that for cases that are either resectable or borderline resectable at presentation and for whom upfront surgery is a feasible alternative, the superior pathway selection was determined by the individual patient's probability of receiving multimodal treatment (neoadjuvant therapy and resection or resection and adjuvant therapy) in either pathway. These findings held strong even when results from emerging RCTs that have reported advances in treatments within both pathways were integrated within the simulation model. This also provided predicted thresholds that must be achieved in real life, where complexity is not controlled for, in order that the reported positive findings of such RCTs are applicable at individual patient level.

Conclusion: The methods presented in this paper adds a further dimension to the debate surrounding the treatment of pancreatic cancer and has a potential role in future cost-effectiveness analysis. These findings challenge the main criticism of neoadjuvant therapy for resectable disease, mainly that patients who do not undergo resection have lost their window of resectability, by quantifying the potential survival gains achieved with upfront resection, which are limited. Importantly it moves future research towards supporting

better clinical decision making through personalised predictive medicine and highlights the need to develop ways of statistically engaging with complexity to achieve this.

Introduction

Previous Markov decision-analysis studies have reported a slight survival benefit with neoadjuvant approach (de Geus *et al*., 2016; Sharma *et al*., 2015; Van Houton *et al*, 2012). The results presented in the previous section have taken the application of such a modeling approach even further through Markov cohort analysis, which suggested that superior treatment pathway selection should actually be determined at individual patient level, depending on individual patient and tumour factors such as tumour resectability and probability of the individual patient being able to physiologically cope with interventions inorder that they receive multimodal theapy (Bradley *et al*., 2018; Bradley & Van der Meer, 2019). The problem is that the majority of existing observational and cohort studies comparing neoadjuvant and surgery first approaches combine resectable and borderline resectable cases of pancreatic cancer within the neoadjuvant arm, therefore they fail to offer a true like-for-like comparison for the treatment of resectable disease. Therefore what has not yet been established is how alternative options for treatment pathway delivery affects outcomes at individual patient level depending on the stage of their disease at presentation. Consequently key areas for targeted pathway improvement could be overlooked.

In this section a systems level approach combined with DES modeling is used to model the treatment pathway of people presenting to tertiary level care with resectable, borderline resectable and unresectable pancreatic cancer. The approach combines statistically and clinically meaningful risk groups, determined by the disease stage of their pancreatic cancer at presentation, with a simulation model that captures the natural history of the disease across competing treatment pathways. These individual patients are simulated as they progress through the pancreatic cancer natural history. Transition times and transition probabilities are used to capture the variability between groups with different stages of the disease, and intervention and treatment programmes are modelled through changes to the nature of these parameters. For example, a particular treatment might delay recurrence or progression of pancreatic cancer, modeled by increasing the individual's dwelling times in natural history states, with survival time quality adjusted accordingly, or by changing their transition probabilities. The approach is illustrated through a synthesisation of data from randomised controlled and phase II/III drug trials. This allows the model to capture best quality data across all treatment pathways on a global scale hence avoiding limitations and bias inherent in institutional databases and the lack of necessary details regarding treatment inherent in national databases.

Operational research techniques, many utilising computer simulation including DES, have been widely applied to healthcare primarily in the areas of healthcare systems operations, disease progression modelling, screening modelling and health behaviour modeling

(Zhang *et al*., 2018) with their application in cancer services mainly centring around screening, planning and scheduling care (Saville *et al*., 2019). Few studies however exist that explore the application of operational research techniques to pancreatic cancer as a particularly challenging malignancy. Furthermore techniques of operational research have not yet been fully utilised to support better individualised patient selection across competing treatment pathways in this field. This work differs from published work in this area, not only in its uniqueness of application to pancreatic cancer, but also in incorporating a systems modelling approach, using patient classification techniques based on cancer staging, coupled with DES. The approach lends itself well to disease modelling incorporating different treatment strategies targeted at different risk groups (Harper *et al*., 2003). It also offers an advantage over previous decision analysis studies on the management of pancreatic cancer that have relied on Markov modelling which can have reduced accuracy due to lack of memory within the Markov model, the effect of depletion of susceptibles, and the timing of transitioning within the model being dependent on a per cycle, as opposed to time-to-event, basis (Caro *et al*., 2010). The wider application of this research will not only be in supporting better decision making and treatment pathway selection at individual patient level, but also in identifying priority areas for future research and investment by identifying key components of treatment pathways where making improvements could stand to have the highest beneficial impact.

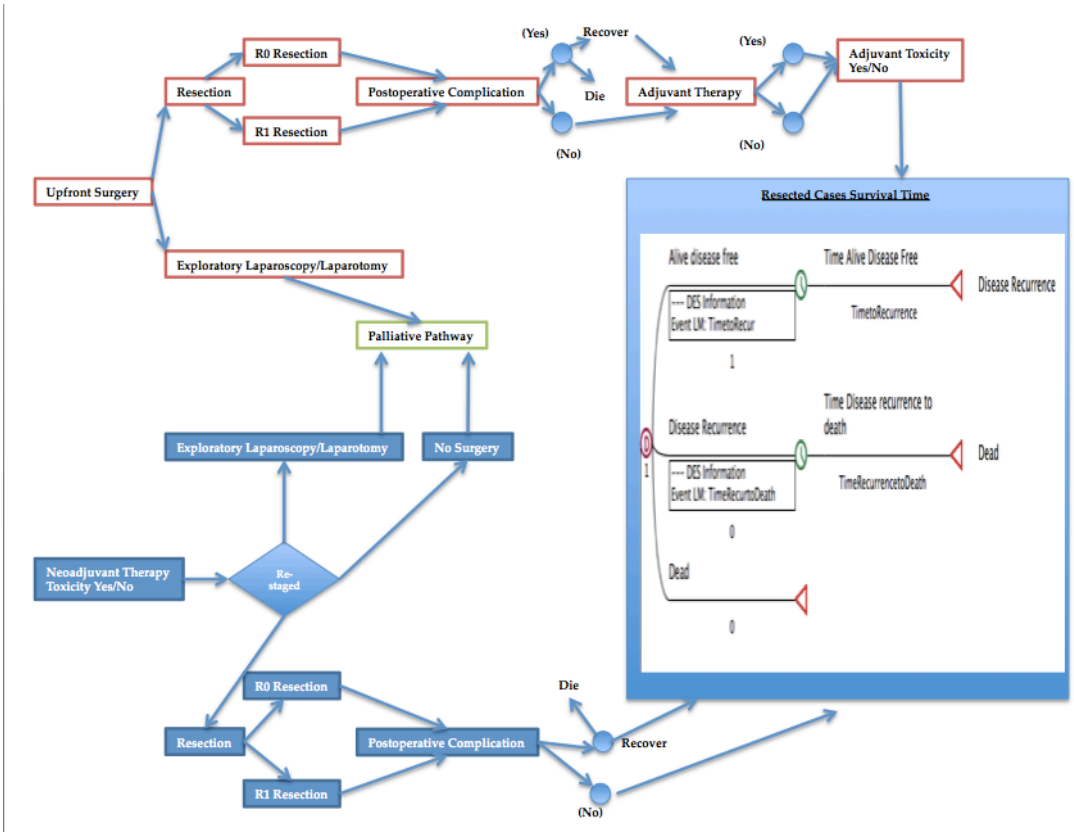
The rest of this section is structured as follows. Section 4.3.1 discusses the model development and section 4.3.2 reports the

results of the model in terms of predicted survival outcomes across alternative treatment pathways depending on disease stage at presentation. Section 4.3.3 reports the results of subgroup analysis of resectable only cases treated in competing treatment pathways. Section 4.3.4 reports the results of scenario testing to assess the potential impact of emerging findings from RCTs to assess the potential impact of such findings when the complexity of real-world application is not controlled for. Section 4.4 then triangulates the accuracy of output from the Markov model and DES model decision-analysis using data from the West of Scotland Pancreatic Unit database for resectable pancreatic cancer.

4.3.1 Model development: transparency of analyses

A combined systems and DES model was designed using TreeAge Pro 2019 (TreeAge Software Ins., Williamstown, MA) to assess all pathway options for treatment delivery for each category of disease stage. Option 1 reflects current guidelines with all resectable cases treated in the surgery first pathway and borderline resectable cases treated in the neoadjuvant pathway. In option 2 both resectable and borderline resectable cases are treated in the neoadjuvant pathway. In option 3 resectable cases are treated in the surgery-first pathway and the borderline resectable cases are treated in the palliative chemotherapy pathway. In all three options unresectable disease is treated with palliative chemotherapy or supportive care (Figure 38).

Figure 38: Discrete Event Simulation Model Structure



The benefit of each option was calculated in terms of life months and QALMs gained. The latter is a generic health utility measure based on duration and quality-of-life. Utilities were based on quality-of-life indicies taken from published literature (Ljungman *et al.*, 2011; Murphy *et al.*, 2012) and also based on World Health Organization and European Quality of Life Survey (Eshuis *et al.*, 2015; Romanus *et al.*, 2012; Tam *et al.*, 2013). This scaled survival from 0 (equivalent to death) to 1 (equivalent to perfect health) (deGeus *et al.*, 2016).

Data Source

To determine transition probabilities within the model for each disease stage category across competing treatment pathway options a thorough search of the literature was undertaken across MEDLINE, Embase, PubMed and Cochrane database and Cochrane database of Clinical Trials and strictly adhered to PRISMA guidelines (Moher *et al*., 2009). The inclusion criteria was RCTs and prospective phase II and III studies of neoadjuvant therapy for the treatment of PDAC, published in English language since 2000, involving chemo/radiotherapy-naive human subjects over 18 years of age with preoperatively staged pancreatic cancer. Retrospective and cohort studies, case series and case reports were excluded. Data on resectable and borderline resectable/ locally advanced stages of disease were separated to populate the respective pathways. As the majority of trials were single arm, to populate the surgery-first pathway, the same electronic databases were searched for RCTs of surgery and adjuvant therapy, with the same inclusion criteria.

By definition these patients have adequate performance status to survive surgery, be randomised to adjuvant therapy and therefore had not developed early metastatic disease. To overcome this bias cohort studies comparing neoadjuvant and surgery first pathways were included in the surgery first pathway to offer comparison across outcomes of resection, R0 resection rates, receipt of adjuvant therapy and corresponding survival times.

A comprehensive literature searches was performed to populate the model (Appendix N). Risk-of-bias of all included studies was assessed using the Cochrane Collaboration's risk of bias tool (Higgins et al., 2011) and ROBINS-I tool (Sterne et al., 2016) and these results are provided along with a summary of included studies in appendix N to help guide judgement on the certainty and reliability of findings.

Transition Probabilities and Time to Event Data

Transition probabilities are displayed in table 45 and were calculated based on weighted pooled estimates of proportions from included studies, calculated using Freeman-Tukey arcsine square root transformation under random effects model to account for heterogeneity across studies (Freeman & Tukey,1950). As evidence has shown that unbiased pooled estimates of median survival times cannot be achieved by weighted averaging of medians (Gillen *et al*., 2010; Rounder & Speckman 2004), we used Gillen *et al*. (2010) approach to calculate weighted time-to-event data. As previously explained this approach is based on averaging parameter estimates of a presumed density function of survival time, or in this case time-to-event, then using the pooled distribution parameter, in this case the exponential distribution which implies a time constant hazard rate corresponding to the sole distribution parameter λ, to recalculate the estimate of the median from the pooled distribution parameter (Gillen *et al*., 2010). The data distribution for each model variable was determined by the mean, standard deviation and variance of the input data and fitted against 55 possible data

distributions with the best fit determined by the Anderson Darling statistic (Table 24).

Table 24: DES model transition probabilities

Variable	Transition Probability (95% Confidence Interval)	Range	Standard Deviation; Variance	Data distribution; Parameters; (Anderson Darling Statistic)
Neoadjuvant Pathway for Resectable Pancreatic Cancer				
Grade 3+ toxicity	0.41 (0.90-0.97)	0.70-1.00	0.09037; 0.00817	Gen. Pareto; k=0.16131 σ=0.06585 μ=-0.00512 (0.37908)
Resection	0.63 (0.57-0.69)	0.32-0.85	0.02102; 4.4190E-4	Gen. Extreme Value; k=0.07104 σ=0.01585 μ=0.03134 (0.30431)
Exploratory Laparoscopy/Laparotomy Only	0.16 (0.09-0.25)	0-1.00	0.00633; 4.0057E-5	Johnson SB; γ=2.0682 δ=1.7897 λ=0.0624 ζ=-0.00855 (0.56039)
R0 resection	0.63 (0.49-0.76)	0.53-0.92)	0.03079; 9.4797E-4	Cauchy; σ=0.013 μ=0.05608; (0.21049)
Grade 3+ post-operative complication	0.19(0.13-0.26)	0.06-0.64	0.00457; 2.0931E-5	Gen. Extreme. Value; k=-0.32622 σ=0.0048 μ=0.01075 (0.27029)
Die from post-operative complication	0.16(0.07-0.27)	0-0.57	0.00217; 4.7206E-6	Pareto 2; α=0.22134 β=4.0418E-13 (-6.8426)

	Time in months	Standard Deviation; Variance	Data distribution; Parameters; (Anderson Darling Statistic)
Time to disease recurrence following R0 resection	16.68 months	7.604; 57.821	Normal; σ=7.604 μ=17.314 (0.21707)

Time from disease recurrence to death after R0 resection	14.55 months	12.594; 158.6	Gamma; α=6.7752 β=4.8382 ϒ=0 (0.46345)
Time to disease recurrence following R1 resection	16.68 months	7.604; 57.821	Normal; σ=7.604 μ=17.314 (0.21707)
Time from disease recurrence to death following R1 resection	5.25 months	6.7752; 45.904	Normal; σ=6.7752 μ=23.525 (0.51294)
Time to death for no surgery, or following Exploratory Laparoscopy/ Laparotomy only	11 months		

Neoadjuvant Pathway for Borderline Resectable Disease				
	Transition Probability (95% Confidence Interval)	**Range**	**Standard Deviation; Variance**	**Data distribution; Parameters; (Anderson Darling Statistic)**
Grade 3+ toxicity	0.47 (0.35-0.0.61)	0.11-1.00	0.39317; 15.458	Frechet (3P); α=1.0622 β=16.652 γ=6.5416 (0.92678)
Resection	0.26 (0.20-0.32)	0-0.83	0.18283; 3.3426	General Extreme Value; k=0.08754 σ=0.13054 μ=19.251 (0.20875)
Exploratory Laparoscopy/Laparotomy Only	0.17 (0.08-0.27)	0-1.00	0.28376; 8.052	Pareto 2; α=0.27763 β=7.1795E-12 (-9.9894)
R0 resection	0.68(0.57-0.79)	0-1.00	0.29291; 8.5795	Johnson SB; γ=0.81833 δ=0.75289 λ=130.76 ζ=-22.571 (0.33527)
Grade 3+ post-operative complication	0.22(0.08-0.40)	0-0.88	0.30851; 9.5178	General Extreme Value; k=0.31817 σ=16.032 μ=9.6341 (0.41599)
Die from post-operative complication	0.12(0.03-0.26)	0-0.50	0.19149; 3.6667	Chi-squared; ν = 32
	Time in months	**Standard Deviation;**	**Data distribution;**	

		Variance	Parameters; (Anderson Darling Statistic)
Time to disease recurrence following R0 resection	13.8 months	6.9555; 48.38	Gamma; α=5.2708 β=3.0296 ϒ=0 (0.29205)
Time from disease recurrence to death following R0 resection	11.73 months	12.594; 158.6	Gamma; α=6.7752 β=4.8382 ϒ=0 (0.46345)
Time to disease recurrence following R1 resection	13.8 months	6.9555; 48.38	Gamma; α=5.2708 β=3.0296 ϒ=0 (0.29205)
Time from disease recurrence to death following R1 resection	9.91 months	9.4044; 88.444	Gamma; α=7.6629 β=3.3973 ϒ=0 (0.43885)
Time to death for no surgery, or following Exploratory Laparoscopy/ Laparotomy only	11 months		

Unresectable Disease		
	Transition Probability (95% Confidence Interval)	**Data distribution; Parameters; (Anderson Darling Statistic)**
Palliative Chemotherapy for stage III disease at presentation	0.50	Normal
Palliative Chemotherapy for stage IV disease at presentation	0.28	Normal
Supportive Care Only	0.72	Normal
Toxicity with palliative chemotherapy	0.52	Normal

Surgery First Pathway for Resectable Disease				
	Transition Probability (95% Confidence Interval)	**Range**	**Standard Deviation; Variance**	**Data distribution; Parameters; (Anderson Darling Statistic)**
Resection	0.94 (0.90-0.96)	0.70-1.0	0.1219; 0.01486	Burr: k=0.0595 α=10.327 β=0.00112 (0.12818)
R0 resection	0.56 (0.51-0.62)	0.16-0.86	0.09869; 0.00974	Pearson 5: α=0.61636

				β=7.0460E-4 (0.18259)
Grade 3+ post-operative complication	0.22 (0.13-0.33)	0.04-0.54	0.01297; 0.0002	Log-Pearson 3: α=66.845 β=-0.09425 γ=2.0838 (0.29235)
Die from post-operative complication	0.07(0.02-0.13)	0-0.36	0.00948; 0.0002	Cauchy: σ=0.00373 μ= 0.00639 (0.38658)
Receiving adjuvant therapy	0.61(0.57-0.66)	0.26-0.94	0.10088; 0.01018	Burr: k=0.26048 α=2.145 β=9.2071E-4 (0.18949)
Adjuvant toxicity grade 3+	0.43(0.25-0.62)	0.09-0.98	0.02753; 0.00076	Log-Pearson 3: α=1916.0 β=-0.02672 γ=47.081 (0.34508)

	Time in months	**Standard Deviation; Variance**	**Data distribution; Parameters; (Anderson Darling Statistic)**
Time to disease recurrence flowing R0 resection and adjuvant therapy	11.4 months	3.0732; 9.4446	Gamma α=18.994 β=0.70515 ϒ=0 (0.8653)
Time to disease recurrence following R0 resection but no adjuvant therapy	5.1 months	3.7186; 13.828	Gamma α=5.4235 β=1.5968 ϒ=0 (0.23772)
Time from disease recurrence to death following R0 resection and adjuvant therapy	9.97 months	7.1312; 50.854	Gamma α=12.033 β=2.0557 ϒ=0 (0.21824)
Time from disease recurrence to death following R0 resection but no adjuvant therapy	14 months	1.6407; 2.692	Normal σ=1.6407 μ=18.32 (0.22305)
Time to disease recurrence following R1 resection and adjuvant therapy	9.5 months	3.0732; 9.4446	Gamma α=18.994 β=0.70515 ϒ=0 (0.8653)
Time to disease recurrence following R1 resection but no adjuvant therapy	3.4 months	3.7186; 13.828	Gamma α=5.4235 β=1.5968 ϒ=0 (0.23772)

Time from disease recurrence to death following R1 resection and adjuvant therapy	6.87 months	3.6622; 13.412	Normal (0.21954)
Time from disease recurrence to death following R1resection but no adjuvant therapy	11.25 months	1.6407; 2.692	Normal (0.22305)

Survival State	**Utility for QALM**
Living with stable pancreatic cancer	0.81
Undergoing chemo/radiotherapy	0.81
Experiencing chemo/radiotherapy complications	0.53
Recovering from pancreatic surgery	0.59
Experiencing surgical complications	0.48
Living with unresectable disease and pre-operative quality-of-life	0.65

Dealing with Uncertainty and Validation: treating data analysis as a partial remnant

A cohort of 10000 patients was modeled through Monte Carlo first order microsimulation. Distributions were applied around all model parameters and the simulation run over 10000 iterations to capture the possible range and frequency of possible values and describe first-order uncertainty. Second order uncertainty was captured through probabilistic sensitivity analysis. Each trial was run 1000 times with the possible mean of each parameter drawn from the data distribution hence capturing uncertainty surrounding the sample mean.

Black box validation was employed to check that inputs and outputs were as expected (Pidd, 2004). Input and output data on survival time, disease-free survival time, and resection rates et cetera was

compared with results from published data. To assess the internal workings of the model white box validation was employed through validation of input parameters to ensure outputs resulting from different distribution inputs provided a reasonable fit to empirical data (Pidd, 2004). Further applications of white box validation included both static logic validation and dynamic logic validation. Face validity was tested by consultation with clinical experts to agree the model structure throughout the process.

4.3.2 Results

Overall the greatest benefit was seen in borderline cases treated in the neoadjuvant pathway (13.92months (10.98 QALMs) or 13.93 months (10.89 QALMs) in options 1 and 2 respectively) compared to the palliative chemotherapy pathway modelled in option 3 (8.50months; 6.41 QALMs) (Table 25). For resectable cases there was a marginal overall survival advantage with option 2, the neoadjuvant pathway (20.02months (17.16 QALMs) *versus* 17.49months (13.11 QALMs) in options 1 and 3), which was explored further in subgroup analysis directly comparing surgery first and neoadjuvant pathways for resectable only cases. Outcomes for resected disease varied across model options with a survival advantage demonstrate for both resectable and borderline resectable disease category treated in the neoadjuvant pathway with conversion to resectability for the latter disease category producing almost equivocal survival outcomes to disease that is resectable at presentation and treated in the same pathway (Table 26).

Table 25: Summary of Overall Pathway Analysis Results

Disease Stage Category	Option 1: Surgery first pathway for resectable disease, neoadjuvant pathway for borderline/ locally advanced disease		Option 2: Neoadjuvant pathway for resectable and borderline/ locally advanced disease		Option 3: Surgery first for resectable disease, palliative chemotherapy pathway for borderline/ locally advanced disease	
Resectable	Mean	17.49months (13.11 QALMs)	Mean	20.02months (17.16 QLAMs)	Mean	17.49months (13.11 QALMs)
	Range	17.24-17.71months (12.93-13.28 QALMs)	Range	19.65-20.47months (16.87-17.52 QLAMs)	Range	17.27-17.69months (12.94-13.29 QALMs)
	Standard deviation	0.07	Standard deviation	0.13	Standard deviation	0.07
	Median	17.50months (13.11 QLAMs)	Median	20.02months (17.16 QLAMs)	Median	17.50months (13.11 QALMs)
Borderline/ Locally Advanced	Mean	13.92months (10.89 QALMs)	Mean	13.93months (10.89 QLAMs)	Mean	8.50months (6.41 QALMs)
	Range	13.60-14.19months (10.62-11.14 QALMs)	Range	13.65-14.625months (10.66-11.17 QALMs)	Range	8.38-8.65months (6.33-6.52 QALMs)
	Standard deviation	0.09	Standard deviation	0.09	Standard deviation	0.02
	Median	13.92months (10.89 QALMs)	Median	13.93months (10.89QALMs)	Median	8.50months (6.41 QALMs)
Unresectable	Mean	3.12months (2.70QALMs)	Mean	3.12months (2.70QLAMs)	Mean	3.12months (2.70 QALMs)
	Range	3.05-3.19months (2.66-2.75QALMs)	Range	3.07-3.23months (2.67-2.78QALMs)	Range	3.03-3.19months (2.64-2.75 QALMs)
	Standard deviation	0.02	Standard deviation	0.02	Standard deviation	0.02
	Median	3.12months (2.70QALMS)	Median	3.12months (2.70QALMs)	Median	3.12months (2.70 QALMs)

Table 26: Pathway results for disease that was resected

Disease stage category	Option 1: Surgery first pathway for resectable disease, neoadjuvant pathway for borderline/ locally advanced disease	Option 2: Neoadjuvant pathway for resectable and borderline/ locally advanced disease	Option 3: Surgery first for resectable disease, palliative chemotherapy pathway for borderline/ locally advanced disease
Resectable	18.07 months (13.94 QALMs)	25.30 months (22.52 QALMs)	18.07 months (13.94 QALMs)
Borderline/ Locally Advanced	22.24 months (19.23 QALMs)	25.31 months (22.62 QALMs)	

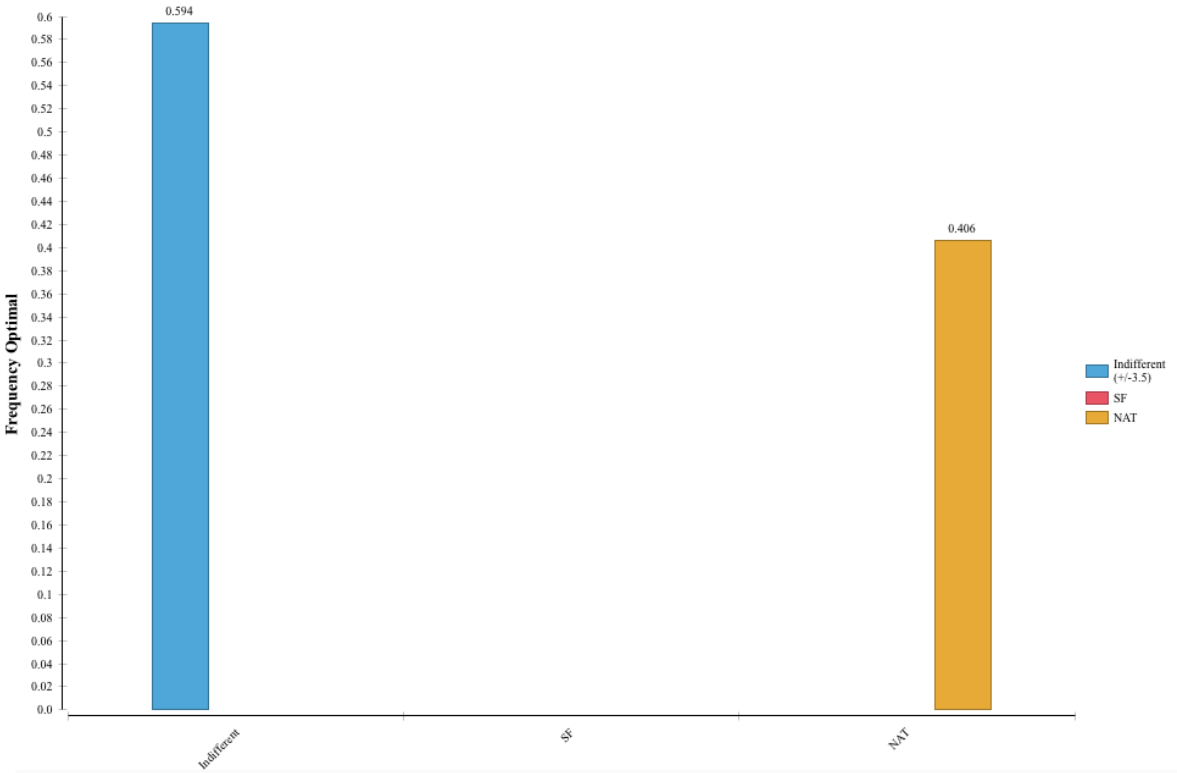
4.3.3 Subgroup Analysis of Resectable Only Cases

The Monte Carlo microsimulation of 10000 patients treated in surgery first and neoadjuvant pathways found that the neoadjuvant pathway gave a mean survival time of 20.01months (18.45 QLAMs) compared to 16.55months (14.19 QALMs) in the surgery first pathway (Table 27). When minimum significant difference threshold for determining the superior pathway was set to 3.5months the selection frequency was 40.6% for neoadjuvant pathway and 59.4% for indifference between pathways (Figure 39).

Table 27: Summary of Subgroup Analysis Results

	Surgery First	**Neoadjuvant Therapy**
Mean	16.55months (14.19 QALMs)	20.01months (18.45 QALMs)
Median	16.65months	20.02months
Minimum	16.26months	19.44months
Maximum	16.80months	20.48months
Standard Deviation	0.07	0.14
Variance	0.01	0.02

Figure 39: Monte Carlo Strategy Selection Frequency. The y-axis depicts the probability that each treatment option along the x-axis is selected as the optimal treatment pathway after 10000patients are simulated through the model. NAT means neoadjuvant pathway and SF means upfront surgery pathway.



In the upfront surgery pathway achieving R0 resection and adjuvant therapy gave 21.27months (17.34QALMs) (range 21.13 to 21.59months; standard deviation 0.08; variance 0.01). R0 resection without adjuvant therapy gave 19.10months (17.84QALMs) (range 18.96 to 19.22months; standard deviation 0.04; variance 0.00). In the neoadjuvant pathway R0 resection gave 29.04months (27.79QALMs) (range 28.55 to 29.57months; standard deviation 0.17; variance 0.03). R1 resection in the upfront surgery pathway gave

16.26months (13.47QALMs) (range 16.10 to 16.43months; standard deviation 0.05; variance 0.00) with adjuvant therapy and 14.65months (13.91QALMs) (range 14.51 to 14.79months; standard deviation 0.04; variance 0.00) without adjuvant therapy. R1 resection in the neoadjuvant pathway gave 20.39months (21.42QALMs) (range 19.92 to 20.76months; standard deviation 0.11; variance 0.01).

Sensitivity Analysis: Threshold analysis

All possible probabilities for variables and time to event data within the simulation model were sampled from the data distribution of each individual variable over 10000 iterations of the microsimulation to establish thresholds that determined pathway superiority. The probability of resection in the neoadjuvant pathway had to be greater than 38% for neoadjuvant pathway to be superior (Figure 40). Furthermore the probability of R0 resection in the neoadjuvant pathway had to be greater than 15.4% (Figure 41).

Figure 40. Threshold Analysis. This figure shows the probability threshold for resection in the neoadjuvant pathway displayed on the x-axis that must be reached for this pathway to be superior. NAT means neoadjuvant therapy. SF means surgery first pathway. Expected value on y-axis is in life months.

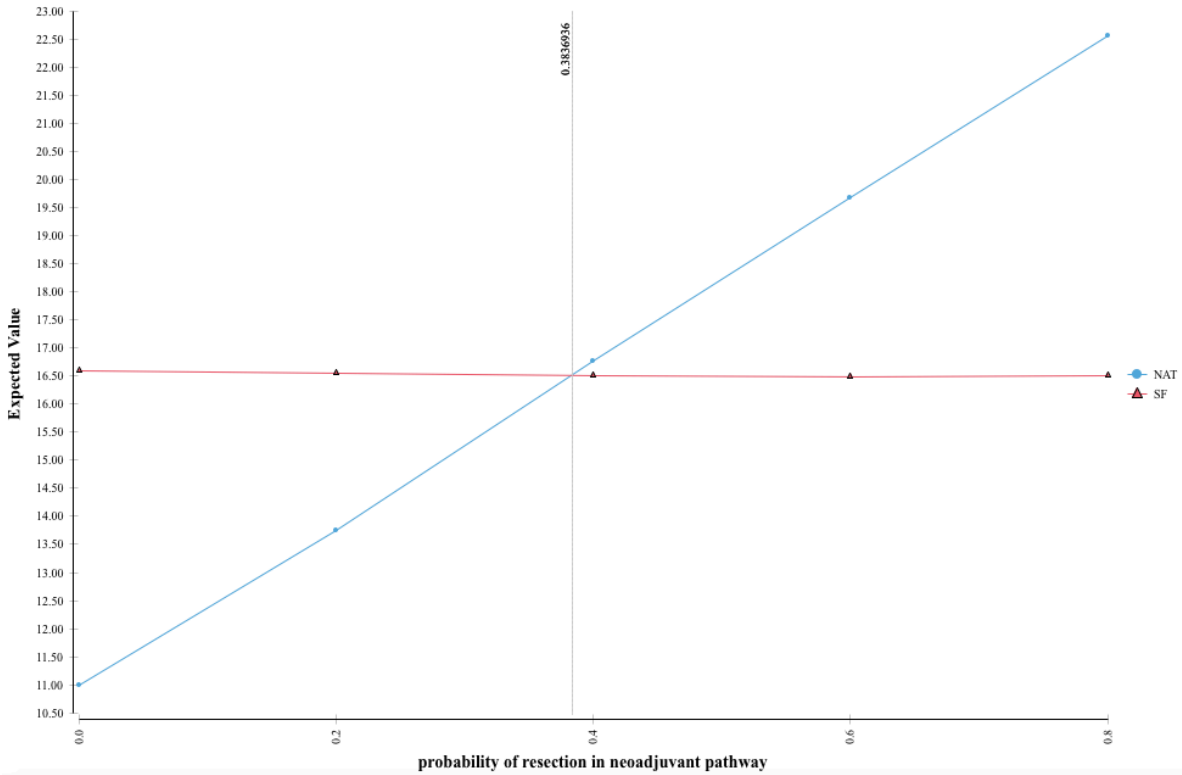
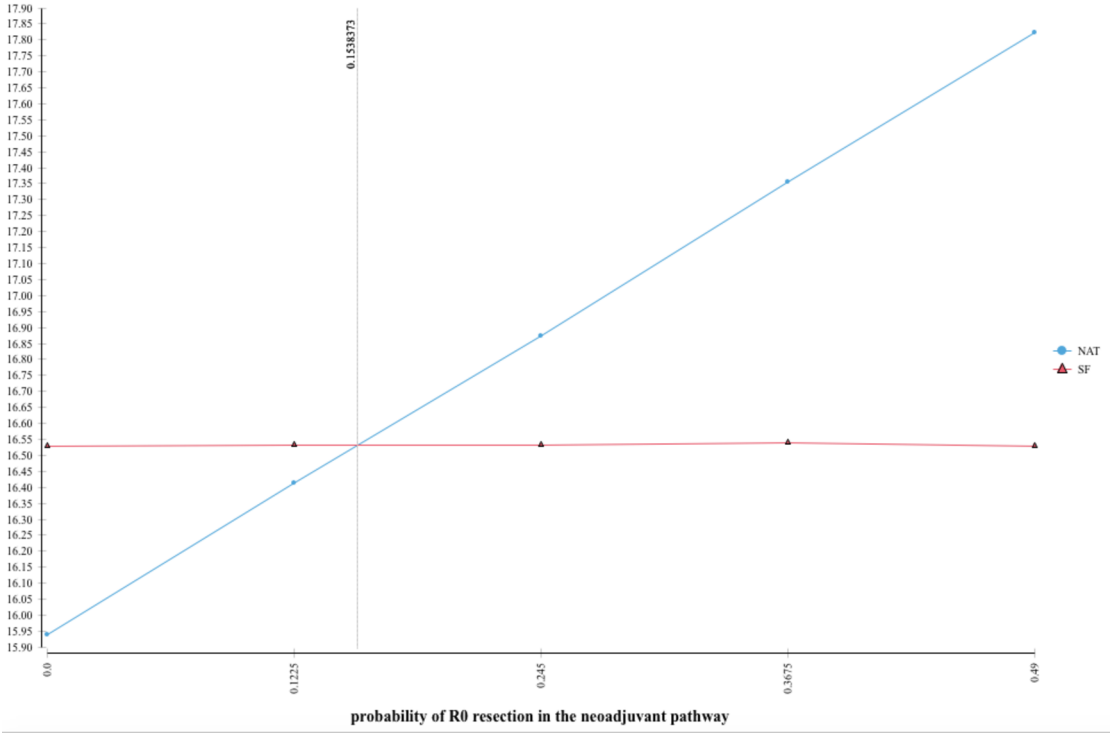


Figure 41: Threshold analysis. Probability of resection in neoadjuvant pathway (x-axis) against expected outcome in months (y-axis). NAT is neoadjuvant pathway. SF is surgery first pathway.



Although the probability of R0 resection in the neoadjuvant pathway did not produce a threshold to alter overall pathway superiority from neoadjuvant to surgery first, the incremental value for patients with a low probability of achieving R0 resection in the neoadjuvant pathway remained below 3months when compared to the expected value in the surgery first pathway (Table 28).

Table 28: Sensitivity analysis of R0 resection in neoadjuvant pathway

Probability of R0 Resection in Neoadjuvant Pathway	Value in Surgery First Pathway	Value in Neoadjuvant Therapy	Incremental Value in Neoadjuvant Pathway
0.0	16.48months	16.72montns	0.24months
0.1225	16.48months	17.39months	0.91months
0.245	16.48months	18.04months	1.56months
0.3675	16.48months	18.65months	2.17months
0.49	16.49months	19.32months	2.84months

Furthermore for patients with the highest chance of R0 resection in the surgery first pathway the expected incremental value with neoadjuvant pathway was only 2months (Table 29). For patients with the highest probability of receiving adjuvant therapy in the surgery first pathway the expected incremental value in the neoadjuvant pathway was 3months (Table 30). This raises the possibility that although neoadjuvant pathway is likely to benefit most patients with resectable pancreatic cancer, those patients with a combination of the highest probabilities of an early R0 resection and receiving adjuvant therapy could potentially benefit from the upfront surgery pathway.

Table 29. Sensitivity analysis of R0 resection in surgery first pathway

Probability of Probability of R0 Resection in Surgery First Pathway	**Value in Surgery First Pathway**	**Value in Neoadjuvant Therapy**	**Incremental Value in Neoadjuvant Pathway**
0.0	14.15months	20.06months	5.91months
0.225	15.09months	20.06months	4.97months
0.45	16.03months	20.06months	4.03months
0.675	16.98months	20.06months	3.08months
0.9	17.97months	20.06months	2.09months

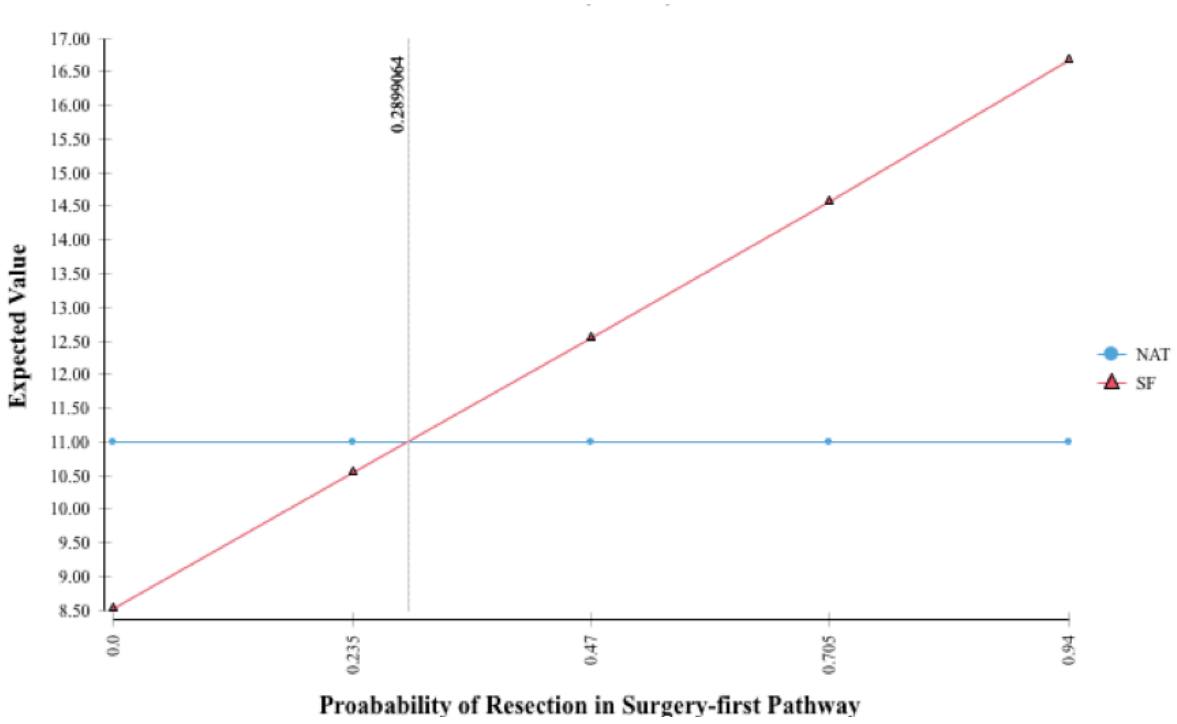
Table 30. Sensitivity analysis of receiving adjuvant therapy in the surgery first pathway

Probability of Adjuvant Therapy in Surgery First Pathway	Value in Surgery First Pathway	Value in Neoadjuvant Therapy	Incremental Value in Neoadjuvant Pathway
0.0	15.53months	20.06months	4.53months
0.225	15.86months	20.06months	4.20months
0.45	16.22months	20.06months	3.84months
0.675	16.54months	20.06months	3.52months
0.9	16.89months	20.06months	3.17months

For patients in the surgery first pathway who were found to have unresectable disease the mean survival time was 8.50months (6.41QALMs). Patients in the neoadjuvant pathway who did not undergo surgery had a mean survival time of 11months (8.04QALMs). These patients are at the centre of the controversy surrounding the use of neoadjuvant therapy for cases of pancreatic caner that are resectable at presentation as critics of this approach highlight the dangers of losing the window of resectability. For this group of patients their probability of resection in the surgery first pathway had to be greater than 29% for the surgery first pathway to be the superior choice (Figure 42). As these patients are presenting

with resectable disease it seems likely that such a threshold would be reached, making surgery first the superior pathway choice for them. However, proponents of the neoadjuvant approach for resectable pancreatic cancer have argued that this particular group of patients represents more aggressive tumour types that are successfully filtered away from futile surgery through neoadjuvant approach. As Figure 42 demonstrates, the expected incremental value in terms of months survival for patients who did not undergo surgery in the neoadjuvant pathway, if treated in the surgery first pathway, were 1.5months (-0.86QALMs), 3.5months (1.14QALMs) and 5.5months (3.14QALMs) corresponding to a probability of resection in the surgery first pathway of 47%, 70.5% and 94% respectively.

Figure 42. Threshold Analysis for Unresectable Disease in the Neoadjuvant Pathway. This figure shows the probability threshold for resection in the surgery first pathway, displayed on the x-axis, that must be reached for this pathway to be superior. NAT means neoadjuvant therapy. SF means upfront surgery pathway. Expected value on y-axis is in life months.



Section 4.3.4 Scenario Testing

As the model was based on synthesised data from pre-existing trials the scenarios outlined in table 31 were introduced to the model to assess the effect of emerging research findings on model outputs. In particular these scenarios account for preliminary findings from RCTs of neoadjuvant therapy for borderline and resectable pancreatic cancer compared to surgery first approach as well as

emerging evidence of improved outcomes with more effective adjuvant therapy in the surgery first treatment pathway.

Table 31: Scenario Testing

Scenario	Description and implementation within the model	Results
Preliminary results from PREOPANC-1 trial comparing upfront surgery and neoadjuvant therapy for borderline resectable cases have reported improved outcomes with neoadjuvant therapy (Van Tienhoven *et al.*, 2018).	Model option 1 adapted to reflect findings from PREOPANC-1 trial. Resection rate set to 72% in surgery-first pathway and 60% in neoadjuvant pathway. R0 resection rate set to 31% in surgery first pathway and 63% in neoadjuvant pathway. Overall and disease-free survival set at 13.5months and 7.9 months respectively in surgery-first pathway. Overall and disease free survival set at 17.1 and 11.2 months respectively in neoadjuvant pathway for borderline cases. For resected cases overall survival set at 18.8months and 42.2months in surgery-first and neoadjuvant pathways.	Based on intention-to-treat overall and disease-free survival for cohort: <u>Neoadjuvant Pathway:</u> 13.81months (11.79 QALMs); range 13.48-14.17months (11.49-12.08 QALMs); standard deviation 0.11. <u>Surgery First Pathway:</u> 11.75months (7.92 QALMs); range 11.57-11.93 months (7.69-8.11 QALMs); standard deviation 0.05 Based on overall survival time for resected cases: <u>Neoadjuvant Pathway:</u> 25.97 months (19.69 QALMs); range 25.40-26.58 months (19.28-20.14 QALMs); standard deviation 0.19. <u>Surgery First Pathway:</u> 15.71months (10.54 QALMs); range 15.51-15.94 months (10.29-10.81 QALMs); standard deviation 0.07
Meta-analysis by Janssen *et al*. (2018) report improved resection and R0 resection rates with neoadjuvant FOLFIRINOX for borderline cases.	Model option 1, neoadjuvant pathway for borderline resectable cases adapted to implement: resection rates 78.4%; R0 resection rates 87.4%; grade 3-4 toxicity set at 49%	<u>Borderline cases</u>: 19.80 months (16.80 QALMs); range 19.45-20.20 months (16.58-17.27 QALMs); standard deviation: 0.13
Randomised controlled trail comparing adjuvant modified FOLFIRINOX with adjuvant gemcitabine reports improved survival outcomes with modified FOLFIRINOX but an	Model option 4 (subgroup analysis) adapted implement: disease-free and overall survival in the upfront surgery pathway 21.6 months and 54.4 months respectively for	<u>Surgery First Pathway:</u> 38.43months (31.19QALMs)

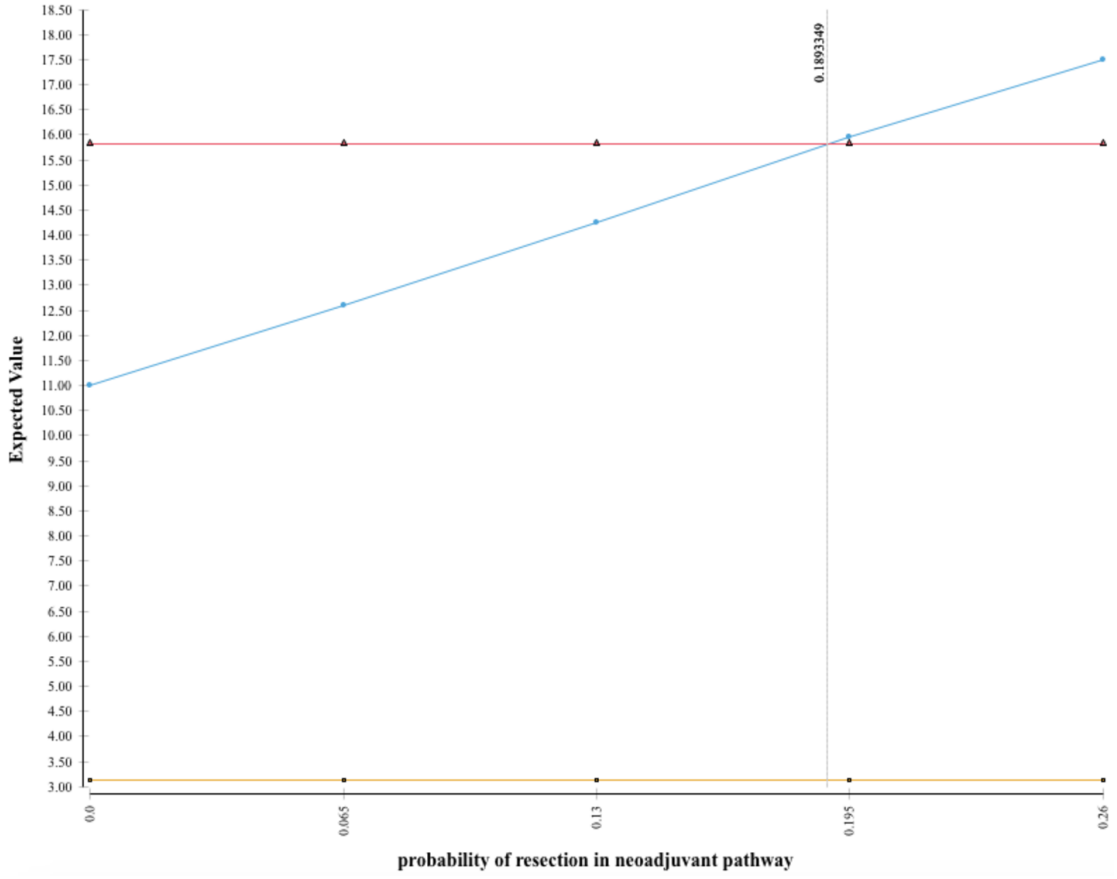
increased toxicity profile (Conroy *et al*., 2018)	resected cases that received adjuvant therapy. Adjuvant therapy toxicity rate set to 75.9%.	
Preliminary findings from Prep-02/JSAP-05 trial comparing upfront surgery and neoadjuvant therapy for resectable pancreatic cancer reports improved survival outcomes with neoadjuvant therapy (Unno *et al*., 2019).	Model option 4 (subgroup analysis) adapted to implement Prep-02/JSAP-05 trial: no statistically significant difference in resection, R0 resection and postoperative complications rates; neoadjuvant grade 3-4 toxicity rate: 72.8%; overall survival 36.72months in neoadjuvant pathway and 26.65months in surgery-first pathway.	Surgery First Pathway: 23.80months (20.06QALMs) Neoadjuvant Pathway: 28.69months (25.92QALMs)

Preliminary results from the PREOPANC-1 trial (Van Tienhoven *et al*., 2018) were incorporated into the structure of model option 1 to compare outcomes for borderline resectable cases treated in both the surgery first and neoadjuvant pathways. For those treated in the neoadjuvant pathway the expected survival time altered very little when compared the original output from model option 1 for borderline resectable cases treated in the neoadjuvant pathway (13.92months to 13.81months). Neoadjuvant pathway was superior compared to surgery first pathway, which gave 11.75months. This however was dependent on the probability of resection in the neoadjuvant pathway being greater than 18.9% (Figure 43). When survival time was altered to reflect that reported for resected cases this produced similar findings (25.97months) to those reported in the original model option 2 for borderline resectable cases that underwent resection (25.31months) (Table 31).

As the majority of trials meeting the inclusion criteria for our study were gemcitabine based scenario testing was used to explore the impact of the growing use of neoadjuvant FOLFIRINOX. This showed that the use of neoadjuvant FOLFIRINOX for borderline resectable

disease improved the expected outcome from 13.81months to 19.80months.

Figure 43: Threshold analysis for borderline resectable cases treated in neoadjuvant and surgery-first pathways. Probability of resection in neoadjuvant pathway (x-axis) against expected outcome in months (y-axis). Red line represents surgery first pathway, blue line represents neoadjuvant pathway, and yellow line represents unresectable disease.



The improved outcomes reported with mFOLFIRINOX in the adjuvant setting of the surgery first pathway increased the overall outcome from the surgery first pathway from 16.56months to

38.43months. However, the apparent benefit of adjuvant therapy only applied to those who underwent resection and receive adjuvant therapy therefore threshold analysis showed that the probability of resection in the surgery first pathway had to be greater than 54% and that the probability of receiving adjuvant therapy also had to be greater than 8% for the surgery first pathway to be the superior option for cases of resectable pancreatic cancer. Inclusion of the preliminary results from the Prep-02/JSAP-05 randomised control trials increased the expected survival time for both pathways but neoadjuvant pathway demonstrated an increased survival advantage from 2.06months to 4.89months.

Discussion

Data as a partial remnant and triangulation

The analysis presented here combines systems modeling with DES to combine mathematical techniques with clinically meaningful data. This provides a useful tool that complements more traditional forms of treatment analysis, such as randomised controlled and cohort studies, to model anticipated outcomes across competing treatment strategies at a more individualised patient level. This gives a useful tool to assist both clinical decision making and future cost-effectiveness analysis. Importantly this study demonstrates how the novel application of operational research methods to the analysis of treatment pathways for pancreatic cancer, an area surrounded by ambiguity and debate, not only adds a further dimension to the on-going debate but drives future research in the direction of more personalised treatment selection strategies.

This analysis found that for cases that were borderline resectable or locally advanced the overall anticipated survival time if treated in the neoadjuvant pathway was 13.92months (10.98 QALMs) or 13.93 months (10.89 QALMs) in options 1 and 2 respectively. The Markov analysis that included all potentially resectable cases gave an overall survival outcome of 23.72 months (18.51 QALMs) in the surgery first pathway *versus* 20.22 months (16.26 QALMs) in the neoadjuvant pathway. The increased survival times in both arms of the Markov model can be explained by the fact that resectable cases were also included in this cohort whereas DES modeling allowed the fitting of different data distributions for resectable and borderline resectable cases. However as with the Markov cohort analysis the DES analysis also showed that where all treatment modalities were received neoadjuvant pathway gave superior survival outcomes.

The results of our analysis corroborate a growing body of evidence reporting survival advantage with neoadjuvant approach for borderline resectable and locally advanced cases of pancreatic cancer (Janssen *et al*., 2018; Versteijne *et al*., 2018). Conversion to resectability produced anticipated outcomes similar to those cases that are resectable at presentation and undergo resection. However, this analysis goes further by integrating preliminary results from the PREOPANC-1 randomised controlled trial, which compared surgery first and neoadjuvant treatment for borderline resectable cases, into the simulation model (Van Tienhoven *et al*., 2018). By placing these preliminary results in a real-world context, where complexity is not controlled, we demonstrated that for borderline resectable cases where surgery first pathway is a viable alternative, the probability of resection after receiving neoadjuvant therapy would have to be

greater than 18.9% for neoadjuvant pathway to be the superior choice.

Both Markov and DES analysis for resectable cases highlighted the importance of more personlised treatment selection strategies. Although overall neither pathway was found to be conclusively superior, there appeared to be a marginal advantage with neoadjuvant pathway. This raises the possibility that for a subgroup of patients with resectable disease who have the highest probability of undergoing an early R0 resection and receiving adjuvant therapy, surgery first could the superior pathway. In the DES analysis there was a marginal overall survival advantage with option 2, the neoadjuvant pathway (20.02months; 17.16 QALMs *versus* 17.49months; 13.11 QALMs in options 1 and 3). Subgroup analysis showed that for resectable only cases the neoadjuvant pathway gave a mean survival time of 20.01months (18.45 QLAMs) compared to 16.55months (14.19 QALMs) in the surgery first pathway. When minimum significant difference threshold was set to 3.5months the selection frequency was 40.6% for neoadjuvant pathway and 59.4% for indifference between pathways. The Markov analysis of resectable only cases also reported a survival advantage with neoadjuvant therapy (26.41 months; 22.54 QALMs) compared to surgery first approach (23.72 months; 18.51 QALMs). Sensitivity analysis of both Markov and DES models supported the importance of individual patient and tumour factors determining superior pathway selection. Within the DES model this was this was found to depend on the individual's probability of resection being greater than 38% and the probability of R0 resection being greater than 15.4%

within the neoadjuvant pathway. In the Markov analysis the probability of resection in the neoadjuvant pathway had to be greater than 47.48% to maintain superiority.

This analysis however went further to provide new insights and add a further dimension to the ongoing debate regarding the treatment of resectable pancreatic cancer, which remains controversial. For patients who do not receive surgery in the neoadjuvant pathway it is either assumed that they missed their window of resectability, and therefore would have had vastly better survival outcomes in the surgery first pathway, or that they had more aggressive disease and were successfully filtered away from futile surgery. This analysis demonstrated that the maximum expected value for these patients in the surgery first pathway was only 5.5months (3.14QALMs).

Preliminary results from the Prep-02/JSAP-05 trial favour neoadjuvant treatment for resectable pancreatic cancer (Unno *et al*., 2019) yet these results are challenged by RCTs comparing adjuvant regimes that report survival outcomes exceeding those reported in the neoadjuvant arms of the former (Conroy *et al.*, 2018). Both RCTs focus on outcomes for cases where all treatment modalities are received. Our study places these findings in a real world context where complexity is no longer controlled and demonstrates that rather than one pathway being conclusively superior for all, superior pathway selection actually could depend on individual patient and tumour factors. When the results of a RCT that reported improved survival time with adjuvant mFOLFIRINOX (Conroy *et al*., 2018) that

exceeded the survival time reported in the neoadjuvant arm of the Prep-02/JSAP-05 trial (Unno *et al*., 2019) was implemented within the model and showed surgery first pathway to be superior, the importance of personalised selection was again demonstrated. When this scenario was tested within the simulation, surgery first pathway was only the superior pathway provided the individual patient had a probability of resection greater than 54% and a probability of receiving adjuvant therapy greater than 8% within the surgery first pathway.

These findings have important implications for the future direction of research. It is widely assumed that through further RCTs one superior pathway for the treatment of all cases of resectable pancreatic cancer will be conclusively established. Theses results show that the future direction of research must not only be on seeking to address these issues through large multicentre RCTs offering a true head-to-head comparison of upfront surgery *versus* neoadjuvant approach for resectable pancreatic cancer, but also on embracing the advances in computational statistics to engage with the reality of the complex systems in which clinical decision making takes place and where complexity is not controlled for. Only in this way can we hope to use emerging data to support better patient selection through personalised predictions of outcomes across competing treatment strategies. Not only will this facilitate better shared decision making but also the more effective allocation of resources. Therefore to further explore the potential advantages of the modeling approach presented here over the more established

Markov modeling technique, the following section will triangulate these findings with those produced when the DES model is populated with data from the West of Scotland Pancreatic Unit database to compare the accuracy of both Markov and DES model outputs against the survival times actually observed.

4.4 Triangulation of Markov and DES Modeling Techniques Using West of Scotland Pancreatic Unit Database

Abstract

Background: The role of neoadjuvant therapy as an alternative treatment pathway to upfront surgery followed by adjuvant therapy for resectable pancreatic cancer remains controversial and there is a lack of RCTs offering conclusive superiority of either pathway. With trials commencing to develop targeted treatment sequencing and earlier disease detection, data is set to become more complex. It is imperative that accurate modeling frameworks are developed now to facilitate decision making. Markov modeling has previously been utilised but concerns have been raised about the accuracy of this method and outputs have not yet been directly compared to patient level data. This section aims to introduce and assess the accuracy of Markov and DES models for predicting likely survival outcomes for patients diagnosed with resectable pancreatic cancer treated in either neoadjuvant or upfront surgery pathways.

Methods: A disease model mapping upfront surgery and neoadjuvant pathways was created. Survival time was then modeled as transitions between health states within a Markov model and alternatively as time-to-event within a DES model. Both models were populated with data from a prospectively maintained tertiary referral centre database and expected value outputs compared against actual survival times outcomes across all treatment sequences received.

Results: There was no statistically significant difference between observed and expected value survival times produced by the DES model (P value 0.122, 95% CI -0.09- -0.58). The difference was statistically significant with the Markov model (P value 0.007; 95% CI l -13.49 - -3.46).

Conclusion: A DES model with appropriate risk equations capturing patient characteristics and treatment history presents improvements over Markov modeling.

Introduction

There are several ongoing key challenges in the management of resectable pancreatic cancer as previously discussed. Firstly the superior treatment pathway for resectable pancreatic caner has not been conclusively established. Secondly superior treatment regime combinations within competing pathways have not been conclusively established. Furthermore the narrative regarding future research has increasingly turned to developing more individualised selection of treatment strategies either determined through biomarkers or gene targeted therapies (Amanam & Chung, 2018).

Novel treatment developments typically come with cost implications. Cost-effectiveness analysis of neoadjuvant *versus* upfront surgery is lacking and, more widely, oncology economic models rarely capture the entire treatment pathway. Both from the clinician and payer's perspective it is essential that an accurate framework for modeling pathway outcomes for resectable pancreatic cancer be developed now so that accurate economic evaluation of emerging treatment

outcomes can be assessed. Such a framework must also have the flexibility to incorporate the anticipated move within cancer research towards more individualised targeted treatments. Although the majority of recent appraisals for oncology treatments use a survival partition model (Woods *et al*., 2019), this is limited by the fundamental structural assumption that disease-free endpoints and overall survival endpoints are independent which creates uncertainty when extrapolating survival time beyond the trial period when overall survival data is immature (Pan *et al*., 2018). Such an approach would also be too simplistic to capture the potential emerging treatment paradigms of more personalised targeted treatments that could have significant impact on clinical outcomes, costs and resource utilisation.

An alternative approach in the form of Markov modeling has been employed for decision analysis of upfront surgery *versus* neoadjuvant therapy for resectable pancreatic cancer (de Geus *et al*., 2016; Sharma *et al*., 2015). These studies were based on synthesised data from published trials and their output was not validated against patient level data. Whilst this approach adds flexibility in sensitivity analysis by incorporating explicit links between end points, it also carries methodological limitations that could inhibit its future application (Caro *et al*., 2010; Miettinen & Caro, 1989). Given the anticipated move towards future personalised targeted treatments the memory-less property of the Markov cohort model makes it less well equipped to handle individual patient data, which can result in reduced accuracy due to depletion of susceptibles and an over simplification of assumptions (Caro *et al*., 2010; Miettinen & Caro,

1989). Furthermore in light of the afore mentioned current challenges in pathway assessment for resectable pancreatic cancer, the implementation of time-dependent transition probabilities when multiple health states and treatment sequences are considered, would make programming and utilising such a model complex (Caro *et al*., 2010; Miettinen & Caro, 1989).

A better framework for modeling treatment pathways for resectable pancreatic cancer could be offered through DES approach as it captures a patient's experience in terms of events and also has the ability to track changes in patient characteristics, health status and treatment history in relation to their impact on outcomes (Pan *et al*., 2018; Caro *et al*., 2010). This potentially makes this approach a more accurate and efficient framework with the flexibility to incorporate future anticipated breakthroughs in personalised targeted treatments. However, the proportion of disease that is resectable at presentation is small considering the data requirements for such a modeling framework and the accuracy of DES approach has not yet been applied to treatment pathway analysis for resectable pancreatic cancer to assess its level of accuracy.

The objective of this study is to compare the accuracy of implementation of Markov modeling and DES modeling within a disease model reflecting the treatment pathways of upfront surgery and neoadjuvant therapy for resectable pancreatic cancer to establish whether either modeling framework displays an advantage in terms of accuracy of output that could be applied to future economic and decision analysis.

Methods

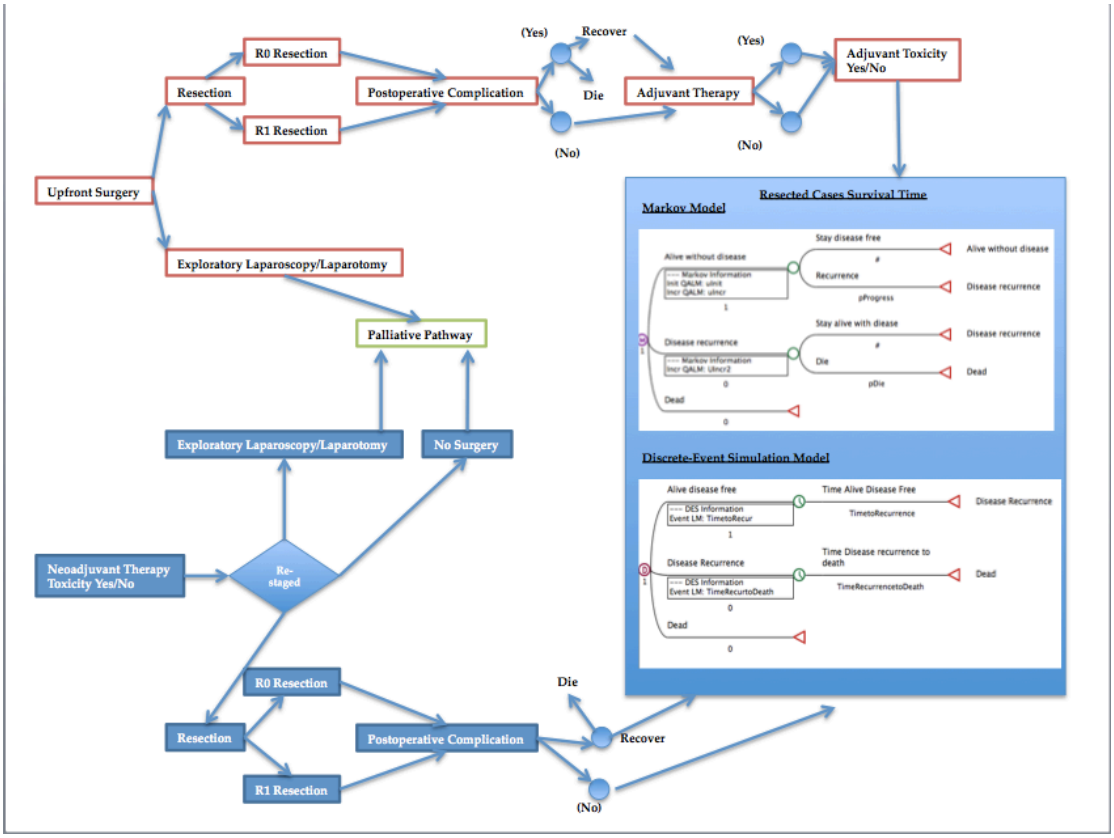
Data Source

A prospectively maintained database for a pancreatic cancer tertiary referral centre contained a cohort of 200 sequential patients diagnosed with non-metastatic pancreatic ductal adenocarcinoma (PDAC). Upfront surgery pathway was performed for all patients from January 2008 to July 2012. From 1st August 2012 to 30th December 2015 100 patients with non-metastatic PDAC were treated in the neoadjuvant pathway, provided multi-disciplinary-team consensus was that R1 resection was likely (tumour extending to any pancreatic margin on CT or EUS evaluation). For this model only those patients with resectable PDAC on initial staging, prior to commencing neoadjuvant therapy, and who were deemed fit for surgery by multidisciplinary team consensus based on performance status score and cardio pulmonary exercise test (CPET), were included (n = 59). Borderline and locally advanced PDAC, as defined by AHPBA/SSO/SSAT guidelines (Callery *et al*., 2009) were excluded. From August 2012 working backwards, 100 sequential patients in the upfront surgery pathway who had resectable PDAC at presentation, and were deemed fit for surgery based on performance status score and CPET, populated the upfront surgery arm of the model. The neoadjuvant regimen, previously described by Grose *et al*. (2017), was mFOLFIRINOX. Adjuvant therapy regime in the upfront surgery pathway was Gemcitabine monotherapy. All patients underwent surgery in the same unit and no patients were lost to follow-up.

Model Development

A disease model was created using TreeAge Pro 2019 software to reflect the natural history of the disease process across the entirety of both upfront surgery and neoadjuvant treatment pathways including the probabilities of treatment toxicities and operative complications (Figure 44). The model structure was agreed with a panel of experts from a tertiary pancreatic surgical center.

Figure 44: Overview of disease model



The probabilities at chance nodes within the disease model were based on the mean probability and the data distribution of the input data for each chance node was fitted against 55 possible data distributions with best fit of empirical data determined by Anderson

Darling statistic (Table 32). A Markov cycle was then implemented within the disease model to model time to transition between health states of: alive without disease, alive with disease and dead. For comparison a DES model was implemented within the same disease model to model time-to-event with events defined as disease recurrence and death (Figure 44).

Table 32: Transition probabilities within disease model.

Variable	Transition Probability	Variance	Standard Deviation	Data Distribution: parameters (Anderson Darling Statistic)
Grade 3+ toxicity with NAT	0.22	0.21607	0.0046483	D. Uniform: a=0 b=1 (34.26)
Resection in NAT pathway	0.66	0.29876	0.0054659	Poisson: λ=0.84746 (23.333)
Exploratory Laparoscopy/Laparotomy	0.10	0.29876	0.0054659	Poisson: λ=0.84746 (23.333)
No surgery	0.24	0.29876	0.0054659	Poisson: λ=0.84746 (23.333)
R0 resection NAT pathway	0.49	0.24984	0.0049984	Poisson: λ=0. 51282 (19.177)
Grade 3-4 post-operative complication NAT pathway	0.18	0.25895	0.0050887	D. Uniform: a=0 b=1 (19.928)
Grade 5 post-operative complication NAT pathway	0.03	0.25895	0.0050887	D. Uniform: a=0 b=1 (19.928)
Resection SF pathway	0.78	0.1716	0.0041425	Poisson: λ=1.22 (42.381)
R0 resection SF pathway	0.21	0.16305	0.004038	Bernoulli: p=0.79487 (16.14)
Grade 3-4 post-operative complication SF pathway	0.27	0.29692	0.005449	Poisson: λ= 0.34884 (50.791)
Grade 5 post-operative complication SF pathway	0.04	0.29692	0.005449	Poisson: λ= 0.34884 (50.791)
Receiving adjuvant therapy	0.50	0.25	0.005	Poisson: λ=0.5 (38.757)
Adjuvant toxicity grade 3+	0.36	0.24377	0.0049373	Bernoulli: p=0.57895 (33.585)

Statistical Analysis

Bayesian statistical package in SPSS version 25.0.0 was used to conduct Bayesian ANOVA analysis of each variable contained within the database against the dependent variable of survival time. Bayesian inference about Pearson correlation coefficient was then performed to assess the linear relation between each variable and survival time to draw Bayesian inference by estimating Bayes factors and characterising posterior distributions. Linear regression analysis within the context of Bayesian inference was undertaken with variables assessed for their ability to explain and predict values of survival time as a scaled outcome. Log-linear regression was then performed to test the independence of each variable against the outcome of survival time. A default setting of least informed prior was used. Bayesian statistical approach was employed as Bayesian estimation can not only obtain otherwise impossible parameters estimates but produce more accurate parameter estimates (Kim *et al*., 3013c; Depaoli, 2013), even in situations of small sample sizes (Zhang *et al*., 2007).

Receiving multimodal treatment *(P value* < 0.01) and R0 resection (*P value* 0.025) were found to be statistically significant in determining survival time. Different equations were therefore used to determine survival time for different treatment sequences to capture the full patient experience: R0 resection with and without adjuvant therapy, R1 resection with and without adjuvant therapy, neoadjuvant therapy and R0 resection and neoadjuvant therapy and R1 resection.

Within the Markov model survival time was estimated based on transition probability from one health state to another:

- Probability of transitioning from alive without disease to alive with disease (varies by treatment received and resection status)
- Probability of transitioning from alive with disease to death (varies by treatment received, resection status and time in alive without disease state)

Within the DES model survival time was estimated based on time-to-event data distribution:

- Time from alive without disease to event disease recurrence (varies by treatment received and resection status)
- Time from disease recurrence to event death (varies by treatment received, resection status, and time to event recurrence)

In both the Markov and DES models the patient is subject to the competing risk of death defined in the transition to death state probability equation and the time to death equation respectively. Where death is the next state within the Markov model or the next event within the DES model the patient exits the model. The probability of death occurring from postoperative complication was included in the transition probability and time to death equations based on the probability distribution of experiencing a grade 5 complication as defined by the Clavien dindo classification. As all patients within the database were already assessed to have European Cooperative Oncology Group (ECOG) score between 0-1, hence deeming them fit to undergo surgery, ECOG score was not included in this equation.

Simulation cohort

Analysis of all PDAC patients contained within the institution database, regardless of disease stage, performance status or treatment pathway (n=418), revealed that tumour size (*P value* 0.005), ECOG score (*P value* 0.002) and American Joint Committee Cancer (AJCC) stage (*P value* <0.001) were statistically significant factors in determining survival time. As this study was comparing outcomes for treatment pathways for resectable only disease the models were based on data where patients in both cohorts were matched for these factors. Individual patient profiles from this study population of resectable disease were cloned and run through each arm of the models to simulate perfect patient randomisation. Attributes or profiles updated during the simulation included treatment actually received, resection status and impact of treatment complications.

Markov Model

The Markov model was set to 60 cycles with each cycle representing 1 month. The number of cycles ensured that all patients were followed-up within the model until time of death. Transition probabilities within the model were based on the cohort mean. To assess the degree of uncertainty surrounding input data for each variable contained within the model Monte Carlo probabilistic sensitivity analysis was performed and set to 10000 iterations with data for each variable sampled from the entirety of the data

distribution for each variable. Markov states included: alive without disease, alive with disease and dead.

Discrete Event Simulation Model

The DES model simulated 10000 patients with resectable pancreatic cancer treated in upfront surgery or neoadjuvant pathways that were followed-up until time of death. Events in time-to-event analysis included time to disease recurrence and time from disease recurrence to death. To capture first-order uncertainty distributions were applied around all model parameters. The simulation was ran over 10000 iterations. Second order uncertainty was captured through probabilistic sensitivity analysis by running each trial 1000 times with the possible mean of each parameter drawn from the data distribution hence capturing uncertainty surrounding the sample mean.

Black box validation was applied by comparing input and output data with data in the literature (Pidd, 2004). White box validation was carried out through validation of input parameters to ensure outputs resulting from different distribution inputs provided a reasonable fit to empirical data and through both static logic validation and dynamic logic validation (Pidd, 2004).

Validation and analysis of model outcomes

Both Markov and DES model outcomes were survival time in months for each treatment sequence within both upfront surgery and

neoadjuvant pathways. These outcomes were compared to the survival time taken from the Kaplan-Meir survival curve of the study population for each treatment sequence within each pathway (Figure 21).

Results

Overall the Markov analysis showed neoadjuvant pathway gave 32.90 months (28.51 QALMs) compared to 24.68 months (19.23 QALMs) for upfront surgery pathway. The DES analysis showed that the neoadjuvant pathway gave 23.74 months (22.69 QALMs) compared to 16.91 months (14.39 QALMs) for upfront surgery pathway. The results of model outputs across all treatment sequences with both treatment pathways are summarised in table 33 with the observed survival time taken from the Kaplan-Meir survival graph.

Table 33: Summary of results. Comparison of expected outcomes from Markov and discrete-event simulation models with observed survival times from study population.

Treatment Sequence	Observed survival time in months	Markov model expected survival time in months	Discrete-event simulation model expected survival time in months
R0 resection + adjuvant therapy	52	52.59	52.05
R0 resection + no adjuvant therapy	15	22.31	14.84
Neoadjuvant Therapy + R0 Resection	38	45.36	37.14
R1 resection + adjuvant therapy	21	33.37	20.75
R1 resection + no adjuvant therapy	12	20.93	11.91
Neoadjuvant Therapy + R1 Resection	28	28	27.86

The upfront surgery pathway in the Markov model had a standard deviation of 2.75 and variance of 7.51 and the neoadjuvant pathway had a standard deviation of 0.21 and a variance of 0.04. The DES model had a standard deviation of 0.08 and variance of 0.01 for the upfront surgery pathway and the neoadjuvant pathway had a standard deviation of 0.12 and a variance of 0.01.

Sensitivity analysis for both models showed that the probability of receipt of multimodal treatment determined superior pathway selection. In the Markov model neoadjuvant pathway was expected to be superior if the probability of resection in the neoadjuvant pathway was greater than 34% and in the DES model this threshold was 30.35%. In the DES model where the probability of upfront R0 resection was greater than 74.55% the upfront surgery pathway was

superior. Two-way deterministic sensitivity analysis in the Markov model corroborated this finding by demonstrating that superior pathway selection depended on the probability of receiving resection in the neoadjuvant pathway and R0 resection and adjuvant therapy in the upfront surgery pathway (Figure 45).

Figure 45a: Markov Two-way sensitivity analysis. Y-axis shows probability of receiving adjuvant therapy in upfront surgery (SF) pathway and x-axis shows probability of receiving resection in neoadjuvant (NAT) pathway. The red area depicts the range whereby upfront surgery pathway would be the superior pathway. The blue area depicts the range over which neoadjuvant pathway would be the superior pathway.

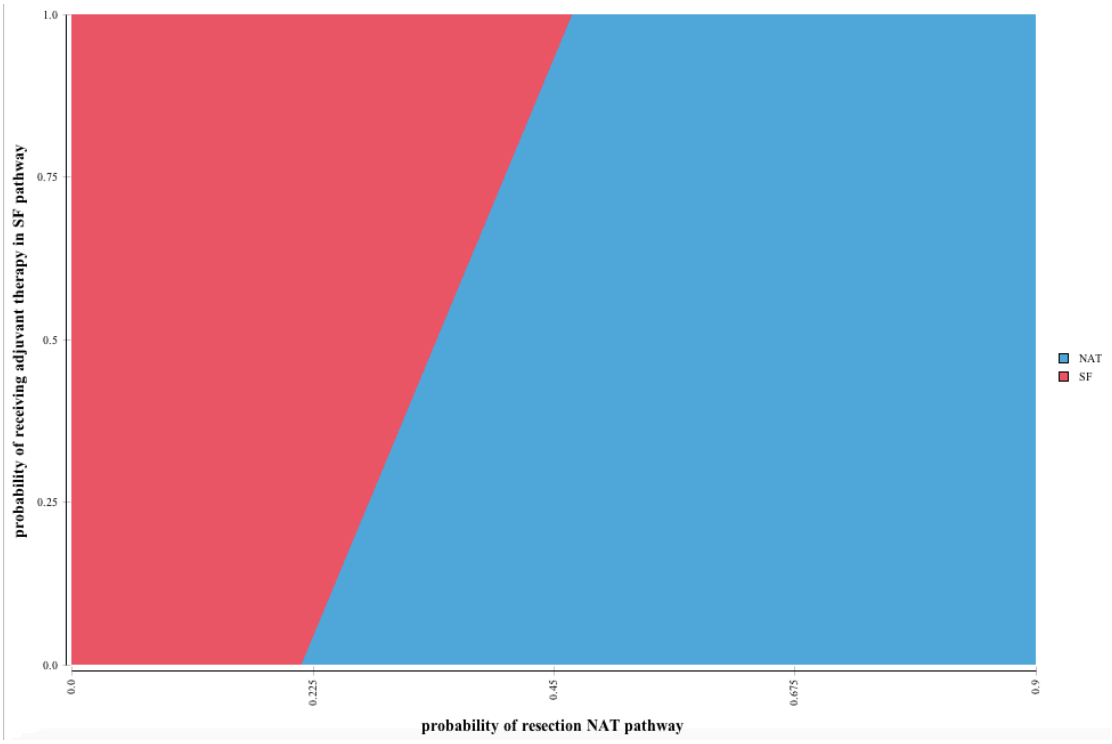
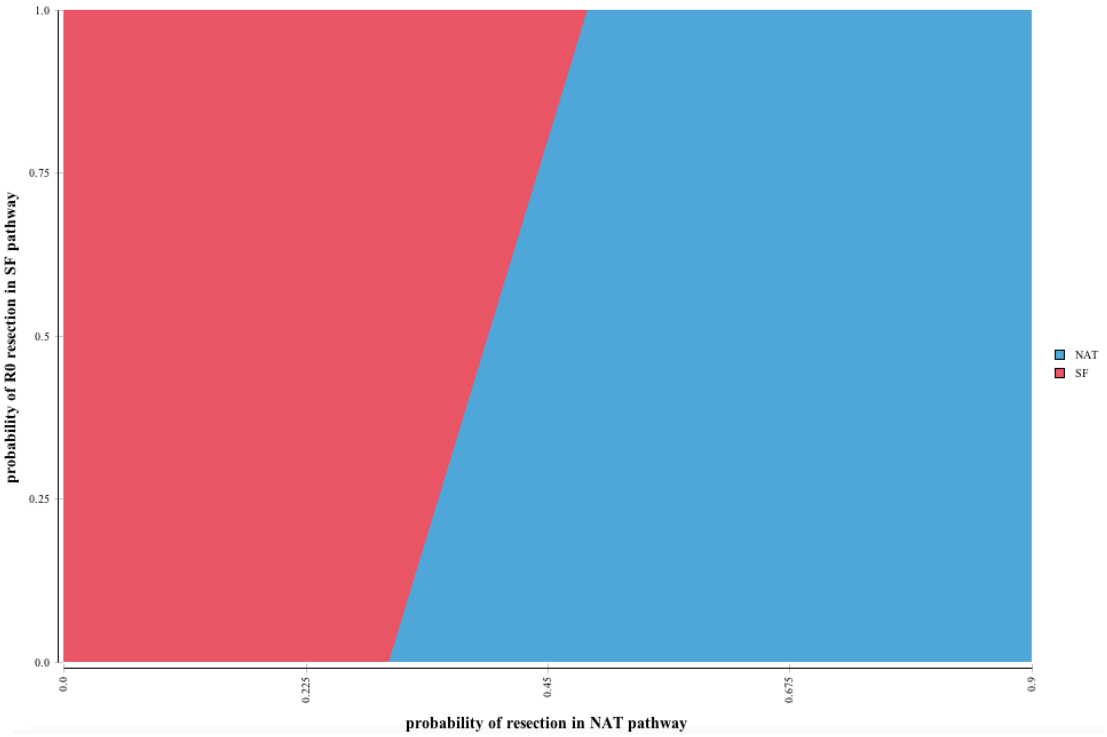


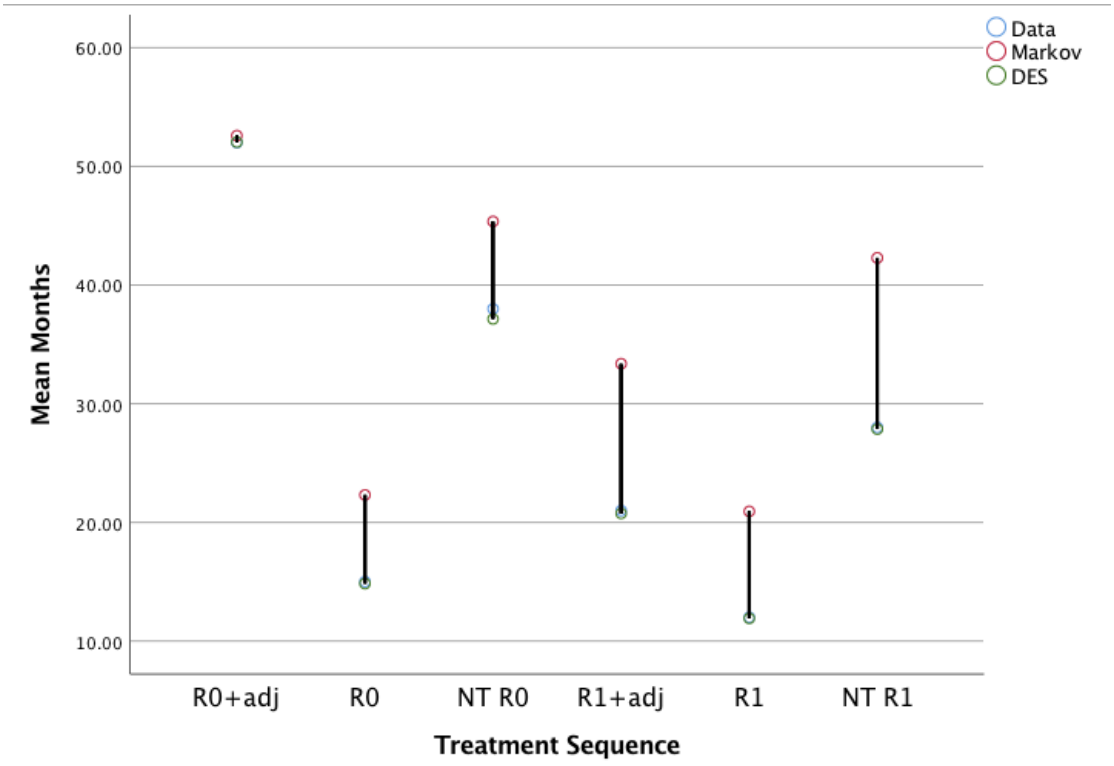
Figure 45b: Markov Two-way sensitivity analysis. Y-axis shows probability of receiving R0 resection in upfront surgery (SF) pathway and x-axis shows probability of receiving resection in neoadjuvant (NAT) pathway. The red area depicts the range whereby upfront surgery pathway would be the superior pathway. The blue area depicts the range over which neoadjuvant (NAT) pathway would be the superior pathway.



The DES model produced expected values closer to the study population survival data across all treatment sequences (Table 33; Figure 46). Paired T-Test showed that there was a mean difference of -8.48 months in the study population survival times when compared to the expected value of the Markov model with a standard deviation of 4.77 and standard error of 1.95. The difference between expected values from the Markov model and the study population was statistically significant (P value 0.007; 95% confidence interval -13.49 - -3.46). When the study population data was compared to the

expected values from the DES model the mean difference was 0.24, standard deviation 0.32, standard error 0.13. There was no statistically significant difference found between the study population data and the expected values from the DES model (P value 0.122, 95% CI -0.09- -0.58).

Figure 46: Results of Paired T-test. Comparison of expected outcomes from Markov and discrete-event simulation models with observed survival times from study population. NT= neoadjuvant therapy. Adj = adjuvant therapy.



Discussion

DES modeling has previously been applied to chronic conditions (*Pan et al*., 2018; Barton *et al*., 2004; Chen *et al*., 2006; Jobanputra *et al*., 2002; Kobelt *et al*., 2009; Lindgren *et al*., 2009; Malottki *et al*., 2011; Tran-Duy *et al*., 2011; Wu *et al*., 2015) but its application to oncology has largely been limited to screening, planning and scheduling care (Saville *et al*., 2019). Markov modeling is growing in popularity and is a widely used method within cost-effectiveness analysis but concerns have been raised regarding its level of accuracy (Caro et al., 2010; Miettinen & Caro, 1989). This study compared the feasibility, validity and potential benefits of Markov and DES modeling frameworks for modeling outcomes for resectable pancreatic cancer.

Expected outcomes from DES modeling were found to be more accurate than those from Markov modeling when compared to study population data across all treatment sequences within both upfront surgery and neoadjuvant pathways. The Markov model overestimated expected values with a mean difference of 8.48months when compared to observed outcomes within the study population. This finding corroborates concerns regarding the assumptions made within Markov models resulting in overestimations as they are based on modeling the cohort population and therefore become less accurate due to depletion of susceptibles and lack of memory making it difficult for the model to characterise how complications or events may be determinants of future events and outcomes (Caro *et al*., 2010; Miettinen & Caro, 1989).

DES and Markov modeling have the disadvantage of being potentially complex with substantial data requirements and are subject to structural and parameter uncertainty (Pan *et al*., 2018). This study is based on a small, non-randomised cohort from a single centre database and therefore the model outputs have limited generalisability. However, DES modeling displayed a high degree of accuracy over Markov modeling in modeling the study data. Overall both models suggested a marginal advantage with the neoadjuvant pathway but sensitivity analysis from both Markov and DES models showed that superior pathway selection depended on individual probability of receiving multimodal treatment within either pathway with the small subgroup of patients who received early R0 resection and adjuvant therapy having the greatest survival time. These findings demonstrate the insights gained from both modeling approaches having the flexibility to test parameter uncertainty through sensitivity analysis. Furthermore these findings support a growing move within research towards more personalised targeting of treatments. Due to its increased flexibility the DES modeling framework stands to become an increasingly important tool. Large multicenter RCTs comparing upfront surgery and neoadjuvant approach for resectable pancreatic cancer are currently lacking but are anticipated to increase in number and explore different treatment regimes within both pathways. By accounting for individual patient characteristics and treatment history at individual patient level the DES modeling framework presented here displayed a high level of accuracy. Using DES modeling to model individual patient data could be used to accommodate deviation from trial protocols when, for example, patients have to switch to second line

treatment regimes or where more individualised targeted treatments are being trialed (Pan *et al*., 2018). Furthermore trails are beginning within pancreatic cancer research that will aim to discover methods of earlier disease detection (UCL, 2019). As results from these trials begin to emerge the initial follow-up time could be shorter than the clinical course of the disease, particularly if earlier disease detection comes to fruition. In such cases DES modeling can make projections of survival time to provide valuable insight (Pan *et al*., 2018).

Conclusion

Considering the emergence of neoadjuvant therapy as an alternative treatment pathway to upfront surgery for resectable pancreatic cancer and the ever increasing complexity of competing treatment pathways, with current trials underway to explore earlier disease detection and targeted therapies, this study set out to explore the feasibility, validity and benefit of Markov and DES modeling frameworks. Our study showed that a DES model with properly developed risk equations to capture individual patient characteristics and treatment history could accurately simulate real-world outcomes for resectable pancreatic cancer treated in both upfront surgery and neoadjuvant pathways. As data begins to emerge from large multicenter RCTs and more complex data emerges from developments in early detection and biomarker and gene targeted therapies, DES modeling stands to become an important tool in the development and assessment of more personalised pancreatic cancer treatment with the associated impact on survival outcomes, resource planning and utilisation.

4.5 Cost-Effectiveness Analysis

Abstract

Background: The aim of this section is to analyse the cost-effectiveness of neoadjuvant therapy (NAT) versus surgery-first (SF) treatment pathways for resectable pancreatic cancer.

Methods: The study was conducted from a National Health Service (UK) perspective with discounting of costs and benefits set at 3.5% and willingness-to-pay set (WtP) at £2,500 per quality-adjusted-life-month (QALM) (£30,000 per quality-adjusted-life-year). A Markov model with 1-month cycle length set to a maximum follow-up time of 60-cycles was created to estimate incremental lifetime costs and benefits. Deterministic and probabilistic sensitivity analysis were undertaken to test model uncertainties including alternative discounting rates. A DES model was also designed to perform Monte Carlo first order microsimulation over 10000 iterations. Second order uncertainty was captured through probabilistic sensitivity analysis. Each trial was run 1000 times with the possible mean of each parameter drawn from the data distribution hence capturing uncertainty surrounding the sample mean. Costs across tertiary and primary level care, including end-of-life care, were included.

The models were populated from synthesised data form RCTs and phase II/III trials. Populating the models with data from the West of Scotland Pancreatic Unit database then triangulated these results. Benefits were measured as QALMs with cost-effectiveness presented as incremental costs, incremental effectiveness, incremental cost-

effectiveness ratio (ICER) and incremental net monetary benefit (NMB).

Markov Model Results: Using synthesised data NAT gave 21.27 QALMs at a cost of £109879.65. SF gave 17.59 QALMs at a cost of £101251.75. NAT therefore had an incremental cost of £8627.90 more than SF for an incremental effectiveness of 3.68 QALMs and an ICER of £2344.16. Using West of Scotland Pancreatic Unit data NAT gave 26.71 QALMs at a cost of £117426.89. SF gave 21.27 QALMs at a cost of £109879.65. NAT therefore had an incremental cost of £29126.08 more than SF for an incremental effectiveness of 8.48 QALMs and an ICER of £3433.07.

DES Model Results: Using synthesised data NAT gave 16.45 QALMs at a cost of £81934.19. SF gave 13.84 QALMs at a cost of £69630.42. NAT therefore had an incremental cost of £12303.77 more than SF for an incremental effectiveness of 2.61 QALMs and an ICER of £4708.51. Based on West of Scotland Pancreatic Unit data NAT gave 21.60 QALMs at a cost of £72083.26. SF gave 13.87 QALMs at a cost of £45813.65. NAT therefore had an incremental cost of £26219.61 more than SF for an incremental effectiveness of 7.73 QALMs and an ICER of £3390.51.

In both Markov and DES models using synthesised and institutional data the main driver of the ICER was receipt of multimodal treatment in the NAT pathway.

Conclusions: When end-of-life care was included NAT pathway was found to cost more than SF pathway but with greater effectiveness. NAT could be considered a cost-effective alternative for the

management of resectable pancreatic cancer when WtP was altered to account for the inclusion of end-of-life care. This analysis also suggests that individualised treatment pathway selection could result in a more cost-effective delivery of care.

Introduction

The significant morbidity associated with pancreatic cancer carries substantial costs to society (Tingstedt *et al*., 2011). Contemporary financial constraints and subsequent limitations on healthcare resources mandate the cost-effectiveness evaluation of competing treatment choices. This, juxtaposed with the lack of level I evidence guiding treatment sequencing, presents an opportunity to evaluate treatment in alternative ways (Drummond *et al*., 2015). Currently only one cost-effectiveness analysis of NAT *versus* SF approach for resectable pancreatic cancer exists and is limited by drawing data for competing treatments from different databases which increases bias and limits generalisability of findings (Abbott *et al*., 2013).

This aims of this study is to synthesise best available international published data to perform cost-effectiveness analysis of SF *versus* NAT approach to the management of resectable pancreatic cancer. First this will be performed using Markov modeling before triangulating these findings with DES modeling for cost-effectiveness analysis. Using both modeling approaches the findings based on synthesised data will also be triangulated with findings from analysis of a tertiary referral centres prospectively maintained database. Specifically this study aims to explore the existing hypothesis that

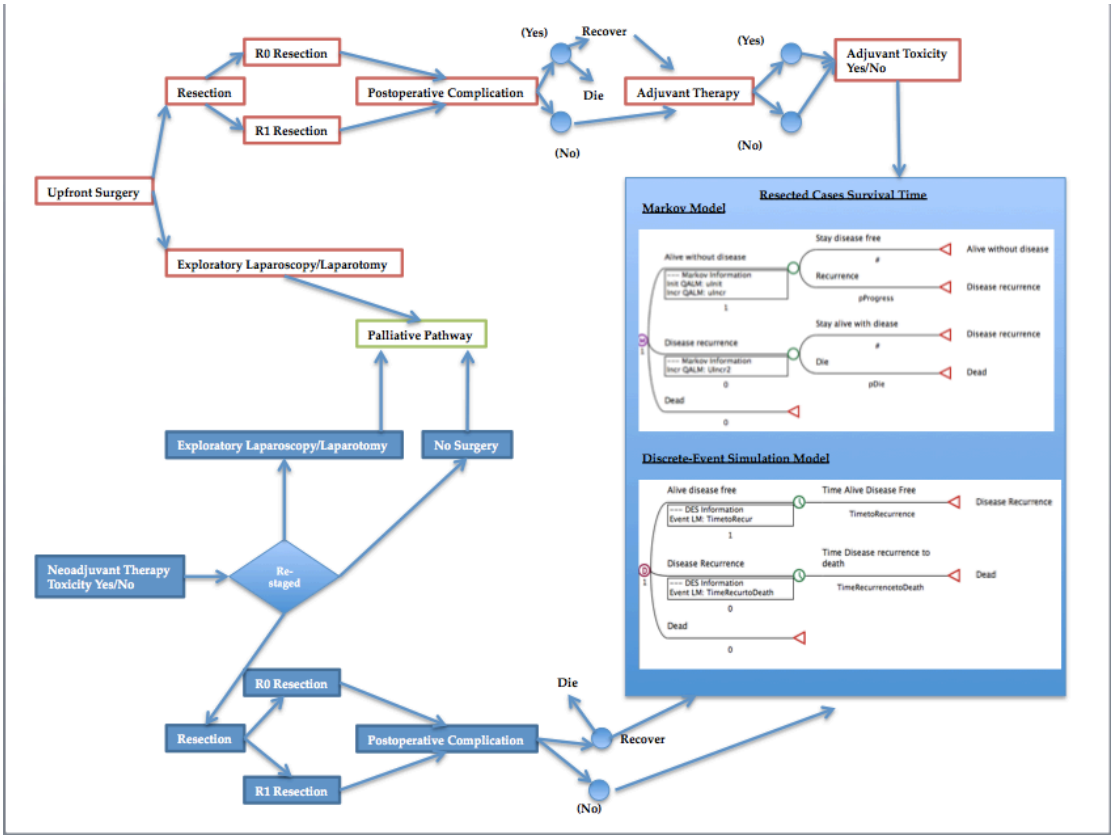
NAT is more cost-effective by filtering patients with more aggressive tumours away from futile yet costly surgery (Abbott *et al*., 2013). The alternative hypothesis that the Markov models developed in section 4.2 and the DES model developed in section 4.3 will more effectively engage with the complexity of the healthcare system being modeled to reveal new insights supporting a cost-effectiveness argument for a move towards better individualised treatment selection strategies will also be tested.

Methods

Analytical Overview

Following guidelines from Consolidated Health Economic Evaluation Reporting Standards (CHEERS) (Husereau et al., 2013) and NICE (NICE, 2011; NICE 2013), the cost-effectiveness of SF and NAT for the management of resectable pancreatic cancer was compared using a Markov model and then triangulated with results from analysis using a DES modeling technique (Figure 47). Both models were designed in an advanced decision-tree format constructed using TreeAge Pro 2017 (TreeAge Software Ins., Williamstown, MA). The development of the Markov and DES models has been previously described in section 4.2 and 4.3 respectively.

Figure 47: Overview of models



Study Population

The base case, surgery first followed by adjuvant therapy (which included chemotherapy, chemoradiotherapy, or both), was compared to NAT (which included chemotherapy and/or chemoradiotherapy) followed by re-staging and, if appropriate, by surgical resection on an intention-to-treat basis. The transition probabilities from syntnesised data within both Markov and DES models were derived from weighted pooled estimates of proportions calculated using Freeman-Tukey arcsine square root transformation under random effects model with corresponding 95% Confidence Intervals (Freeman & Tukey, 1950). The selection process for included trials

has previously been outlined in section 4.2. The ranges in reported literature, standard deviation and variance were used to test uncertainty in model output through both deterministic and probabilistic sensitivity analysis. For further triangulation the Markov and DES models were also populated with data from the West of Scotland Pancreatic Unit database. This study population has been previously described in sections 4.2 and 4.4. Transition probabilities are summarised again in table 34 for the Markov model using synthesised data. Table 35 summarises transition probabilities for the disease based on the West of Scotland Pancreatic Unit database used in both Markov and DES modeling with the time to event data for the latter taken from the Kaplan Meir survival curve (figure 34).

Table 34: Transition Probabilities and Payoff Utility for Quality-Adjusted-Life-Months (QALMs)

Variable	**Transition Probability (95% Confidence Interval)**	**Range**	**Standard Deviation; Variance**	**Data distribution; Parameters; (Anderson Darling Statistic)**
Grade 3+ toxicity with NAT	0.41 (0.90-0.97)	0.70-1.00	0.09037; 0.00817	Gen. Pareto; k=0.16131 σ=0.06585 μ=-0.00512 (0.37908)
Resection in NAT pathway	0.63 (0.57-0.69)	0.32-0.85	0.02102; 4.4190E-4	Gen. Extreme Value; k=0.07104 σ=0.01585 μ=0.03134 (0.30431)
Exploratory Laparoscopy/Laparotomy	0.12 (0.08-0.17)	0-0.36	0.00633; 4.0057E-5	Johnson SB; γ=2.0682 δ=1.7897 λ=0.0624 ζ=-0.00855 (0.56039)
R0 resection NAT pathway	0.49 (0.36-0.62)	0.06-0.71	0.03079; 9.4797E-4	Cauchy; σ=0.013 μ=0.05608; (0.21049)
Grade 3-4 post-operative complication NAT pathway	0.19(0.13-0.26)	0.06-0.64	0.00457; 2.0931E-5	Gen. Extreme. Value; k=-0.32622 σ=0.0048 μ=0.01075 (0.27029)
Grade 5 post-operative complication NAT pathway	0.02(0.01-0.04)	0-0.12	0.00217; 4.7206E-6	Pareto 2; α=0.22134 β=4.0418E-13 (-6.8426)
Resection SF pathway	0.94 (0.90-0.96)	0.70-1.0	0.1219; 0.01486	Burr: k=0.0595 α=10.327 β=0.00112 (0.12818)
R0 resection SF pathway	0.56 (0.51-0.62)	0.16-0.86	0.09869; 0.00974	Pearson 5: α=0.61636 β=7.0460E-4 (0.18259)
Grade 3-4 post-operative complication SF pathway	0.22 (0.13-0.33)	0.04-0.54	0.01297; 0.0002	Log-Pearson 3: α=66.845 β=-0.09425 γ=2.0838 (0.29235)

Grade 5 post-operative complication SF pathway	0.07(0.02-0.13)	0-0.36	0.00948; 0.0002	Cauchy: σ=0.00373 μ= 0.00639 (0.38658)
Receiving adjuvant therapy	0.61(0.57-0.66)	0.26-0.94	0.10088; 0.01018	Burr: k=0.26048 α=2.145 β=9.2071E-4 (0.18949)
Adjuvant toxicity grade 3+	0.43(0.25-0.62)	0.09-0.98	0.02753; 0.00076	Log-Pearson 3: α=1916.0 β=-0.02672 γ=47.081 (0.34508)

Survival State	**Utility for QALM**
Living with stable pancreatic cancer	0.81
Undergoing chemo/radiotherapy	0.81
Experiencing chemo/radiotherapy complications	0.53
Recovering from pancreatic surgery	0.59
Experiencing surgical complications	0.48
Living with unresectable disease and pre-operative quality-of-life	0.65

Table 35: Transition probabilities based on West of Scotland Pancreatic Unit Database

Variable	Transition Probability	Variance	Standard Deviation	Data Distribution: parameters (Anderson Darling Statistic)
Grade 3+ toxicity with NAT	0.22	0.21607	0.0046483	D. Uniform: a=0 b=1 (34.26)
Resection in NAT pathway	0.66	0.29876	0.0054659	Poisson: λ=0.84746 (23.333)
Exploratory Laparoscopy/Laparotomy	0.10	0.29876	0.0054659	Poisson: λ=0.84746 (23.333)
No surgery	0.24	0.29876	0.0054659	Poisson: λ=0.84746 (23.333)
R0 resection NAT pathway	0.49	0.24984	0.0049984	Poisson: λ=0. 51282 (19.177)
Grade 3-4 post-operative complication NAT pathway	0.18	0.25895	0.0050887	D. Uniform: a=0 b=1 (19.928)
Grade 5 post-operative complication NAT pathway	0.03	0.25895	0.0050887	D. Uniform: a=0 b=1 (19.928)
Resection SF pathway	0.78	0.1716	0.0041425	Poisson: λ=1.22 (42.381)
R0 resection SF pathway	0.21	0.16305	0.004038	Bernoulli: p=0.79487 (16.14)
Grade 3-4 post-operative complication SF pathway	0.27	0.29692	0.005449	Poisson: λ= 0.34884 (50.791)
Grade 5 post-operative complication SF pathway	0.04	0.29692	0.005449	Poisson: λ= 0.34884 (50.791)
Receiving adjuvant therapy	0.50	0.25	0.005	Poisson: λ=0.5 (38.757)
Adjuvant toxicity grade 3+	0.36	0.24377	0.0049373	Bernoulli: p=0.57895 (33.585)

As DES modeling allows microsimulation of patient level data and more specific time-to-event data, further details could be included in this analysis such as the effect of palliative chemotherapy (table 36).

Table 36: DES Model transition probabilities

Variable	Transition Probability (95% Confidence Interval)	Range	Standard Deviation; Variance	Data distribution; Parameters; (Anderson Darling Statistic)
Neoadjuvant Pathway for Resectable Pancreatic Cancer				
Grade 3+ toxicity	0.41 (0.90-0.97)	0.70-1.00	0.09037; 0.00817	Gen. Pareto; k=0.16131 σ=0.06585 μ=-0.00512 (0.37908)
Resection	0.63 (0.57-0.69)	0.32-0.85	0.02102; 4.4190E-4	Gen. Extreme Value; k=0.07104 σ=0.01585 μ=0.03134 (0.30431)
Exploratory Laparoscopy/Laparotomy Only	0.16 (0.09-0.25)	0-1.00	0.00633; 4.0057E-5	Johnson SB; γ=2.0682 δ=1.7897 λ=0.0624 ζ=-0.00855 (0.56039)
R0 resection	0.63 (0.49-0.76)	0.53-0.92)	0.03079; 9.4797E-4	Cauchy; σ=0.013 μ=0.05608; (0.21049)
Grade 3+ post-operative complication	0.19(0.13-0.26)	0.06-0.64	0.00457; 2.0931E-5	Gen. Extreme. Value; k=-0.32622 σ=0.0048 μ=0.01075 (0.27029)
Die from post-operative complication	0.16(0.07-0.27)	0-0.57	0.00217; 4.7206E-6	Pareto 2; α=0.22134 β=4.0418E-13 (-6.8426)
Time to disease recurrence following R0 resection	16.68 months		7.604; 57.821	Normal; σ=7.604 μ=17.314 (0.21707)
Time from disease recurrence to death after R0 resection	14.55 months		12.594; 158.6	Gamma; α=6.7752 β=4.8382 ϒ=0 (0.46345)
Time to disease recurrence following R1 resection	16.68 months		7.604; 57.821	Normal; σ=7.604 μ=17.314 (0.21707)
Time from disease	5.25 months		6.7752;	Normal;

recurrence to death following R1 resection			45.904	σ=6.7752 μ=23.525 (0.51294)
Time to death for no surgery, or following Exploratory Laparoscopy/ Laparotomy only	11 months			
Unresectable Disease				
Palliative Chemotherapy (Source: Cancer Research UK, 2019)	0.50			Normal
Supportive Care Only (Source: Cancer Research UK, 2019)	0.72			Normal
Toxicity with palliative chemotherapy (Source: Cancer Research UK, 2019)	0.52			Normal
Surgery First Pathway for Resectable Disease				
Resection	0.94 (0.90-0.96)	0.70-1.0	0.1219; 0.01486	Burr: k=0.0595 α=10.327 β=0.00112 (0.12818)
R0 resection	0.56 (0.51-0.62)	0.16-0.86	0.09869; 0.00974	Pearson 5: α=0.61636 β=7.0460E-4 (0.18259)
Grade 3+ post-operative complication	0.22 (0.13-0.33)	0.04-0.54	0.01297; 0.0002	Log-Pearson 3: α=66.845 β=-0.09425 γ=2.0838 (0.29235)
Die from post-operative complication	0.07(0.02-0.13)	0-0.36	0.00948; 0.0002	Cauchy: σ=0.00373 μ= 0.00639 (0.38658)
Receiving adjuvant therapy	0.61(0.57-0.66)	0.26-0.94	0.10088; 0.01018	Burr: k=0.26048 α=2.145 β=9.2071E-4 (0.18949)
Adjuvant toxicity grade 3+	0.43(0.25-0.62)	0.09-0.98	0.02753; 0.00076	Log-Pearson 3: α=1916.0 β=-0.02672 γ=47.081 (0.34508)
Time to disease recurrence flowing R0 resection and adjuvant therapy	11.4 months		3.0732; 9.4446	Gamma α=18.994 β=0.70515 ϒ=0 (0.8653)
Time to disease recurrence following R0 resection but no adjuvant therapy	5.1 months		3.7186; 13.828	Gamma α=5.4235 β=1.5968 ϒ=0

				(0.23772)
Time from disease recurrence to death following R0 resection and adjuvant therapy	9.97 months		7.1312; 50.854	Gamma α=12.033 β=2.0557 ϒ=0 (0.21824)
Time from disease recurrence to death following R0 resection but no adjuvant therapy	14 months		1.6407; 2.692	Normal σ=1.6407 μ=18.32 (0.22305)
Time to disease recurrence following R1 resection and adjuvant therapy	9.5 months		3.0732; 9.4446	Gamma α=18.994 β=0.70515 ϒ=0 (0.8653)
Time to disease recurrence following R1 resection but no adjuvant therapy	3.4 months		3.7186; 13.828	Gamma α=5.4235 β=1.5968 ϒ=0 (0.23772)
Time from disease recurrence to death following R1 resection and adjuvant therapy	6.87 months		3.6622; 13.412	Normal (0.21954)
Time from disease recurrence to death following R1resection but no adjuvant therapy	11.25 months		1.6407; 2.692	Normal (0.22305)
Survival State	**Utility for QALM**			
Living with stable pancreatic cancer	0.81			
Undergoing chemo/radiotherapy	0.81			
Experiencing chemo/radiotherapy complications	0.53			
Recovering from pancreatic surgery	0.59			
Experiencing surgical complications	0.48			
Living with unresectable disease and pre-operative quality-of-life	0.65			

Costs and Effectiveness

This study was undertaken from the UK NHS payer's perspective.

Costs included in the SF arm were: initial consultant surgeon led

clinic appointment, multidisciplinary pre-operative assessment clinic, cost of surgery with occurrence of significant post-operative complications counted as extra bed days, histopathology assessment of resected specimen, cancer multidisciplinary team meetings, surgical and oncology follow-up clinic appointments, adjuvant therapy with occurrence of significant toxicity counted as extra bed days, and cost of palliative care input from time of disease reoccurrence. In the NAT arm costs included: initial consultant oncologist led clinic, NAT with significant toxicity counted as extra bed days, re-staging outpatient CT scan, cancer multidisciplinary team meetings, initial consultant surgeon led clinic appointment, multidisciplinary pre-operative assessment clinic, cost of surgery with occurrence of significant post-operative complications counted as extra bed days, histopathology assessment of resected specimen, surgical and oncology follow-up clinic appointments and cost of palliative care input from time of disease reoccurrence. Productivity losses were excluded and it is assumed that these costs, including inactivity in terms of work, are comparable across NAT and SF arms.

Cost data was taken from NHS Reference Costs 2017/2018 (NHS, 2017), which provides patient level costs as an average unit cost to the NHS. Discount for both cost and benefit were set at 3.5% with WtP set at £30,000 per quality-adjusted-life-year (QALY), or £2,500 per QALM as per NICE guidelines (NICE, 2011; NICE 2013).

The model's follow-up time was set to 60 Markov cycles (equivalent to 60 months) or until death. The DES model was set to 10000 iterations with all patients followed up until time of death.

Incremental pay-offs were calculated in terms of life months and QALMs. Within the Markov model cumulative payoffs were calculated as a sum of weighted median survival time in months adjusted to quality of survival time spent in each Markov health state which included: alive without disease, alive with disease, and dead. For the DES model the cumulative payoffs were calculated according to time to event, with model events corresponding to time to disease recurrence and time from disease recurrence to death. For the analysis using synthesised data median survival time was used as a better measure of centrality than mean survival time for the following reasons (Dudley *et al*., 2016). Firstly the majority of studies report median survival time due to the fact that survival data is often skewed (Dudley *et al*., 2016; Jager *et al*., 2008). Secondly the mean survival time cannot be truly calculated where patients have not been censored by the end of study follow-up time period therefore it remains unknown if or when they will experience an event (Dudley *et al*., 2016; Jager *et al*., 2008). Furthermore, as evidence has shown that unbiased pooled estimates of median survival times cannot be achieved by weighted averaging of medians, Gillen *et al*. (2010) approach to calculate weighted median survival times was adopted to minimise bias and increase accuracy (Rouder & Speckman, 2004).

For triangulation a cost-effectiveness analysis was then performed by calculating Markov and DES model transition probabilities of interventions, clinical outcomes, and survival in both SF and NAT pathways from the West of Scotland Pancreatic Unit database as described in section 4.2 and 4.3 with overall and disease-free-

survival calculated from median survival time taken from Kaplan-Meir survival analysis based on treatment received (Figure 21).

As previously described utilities were based on quality-of-life indicies which scaled survival from 0 (equivalent to death) to 1 (in complete health) (Sharma *et al*., 2015; deGeus *et al*., 2016). These indicies were taken from published literature (Ljungman *et al*., 2011; Murphy *et al*., 2012) and also analysed based on World Health Organization and European Quality of Life Survey (Eshuis *et al*., 2015; Romanus *et al*., 2012; Tam *et al*., 2013).

Sensitivity Analysis

Markov Model uncertainties were extensively tested through one and two-way deterministic analysis with baseline transition probabilities and costs altered between highest and lowest reported values. The level of confidence in the model output in relation to uncertainty in model input was assessed through Monte Carlo probabilistic sensitivity analysis set to 10000 iterations. Data distributions of the input data for each variable was determined from the median, standard deviation and variance of the input data and fitted against 55 possible data distributions with the best fit determined by the Anderson Darling statistic.

For the DES model first-order uncertainty was captured by the data distributions being applied around all model parameters and simulation being ran over 10000 iterations. Second order uncertainty was captured through probabilistic sensitivity analysis by running

each trial 1000 times with the possible mean of each parameter drawn from the data distribution hence capturing uncertainty surrounding the sample mean. Black box validation was applied by comparing input and output data with data in the literature (Pidd, 2004). White box validation was carried out through validation of input parameters to ensure outputs resulting from different distribution inputs provided a reasonable fit to empirical data and through both static logic validation and dynamic logic validation (Pidd, 2004).

As there is ongoing debate about whether discount should be applied at the same rate for both cost and benefits, model outcomes for both the Markov and DES models were also reported at a discount rate of 6% for costs and 1.5% for benefits (NICE, 2011).

4.5.1 Results: Markov Model

Using synthesised data NAT gave 21.27 QALMs at a cost of £109879.65. SF gave 17.59 QALMs at a cost of £101251.75. NAT therefore had an incremental cost of £8627.90 more than SF for an incremental effectiveness of 3.68 QALMs and an ICER of £2344.16. NAT was therefore found not to breach the WtP threshold set at £2,500 per QALM and could be considered a cost-effective alternative to the traditional SF pathway (Figure 48a and Figure 49a). When these results were triangulated with the cost-effectiveness analysis of the West of Scotland Pancreatic Unit, NAT gave 26.71 QALMs at a cost of £117426.89. SF gave 21.27 QALMs at a cost of £109879.65. NAT therefore had an incremental cost of £29126.08 more than SF for an

incremental effectiveness of 8.48 QALMs and an ICER of £3433.07 (table 37). In this analysis NAT was therefore found to breach the WtP threshold set at £2,500 per QALM and a WtP threshold of £3,500 per QALM was required for NAT to be cost-effective (Figure 48b; Figure 49b).

Table 37: Summary of Results of Cost-Effectiveness Analysis

a) Using Synthesised Data

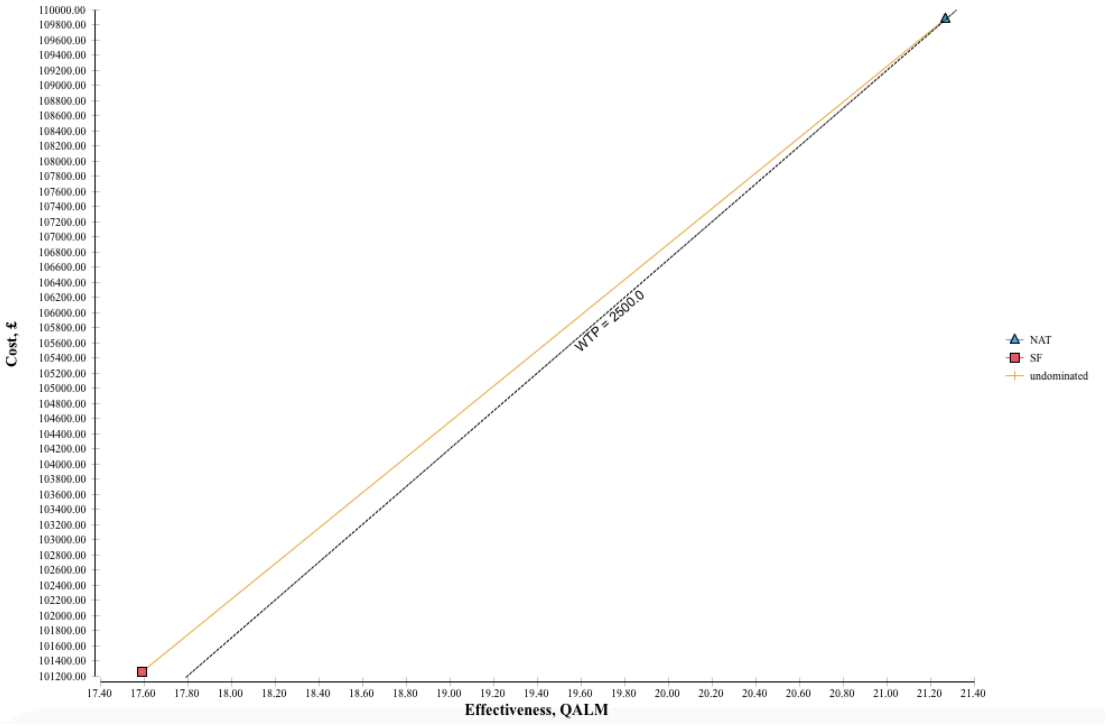
Strategy	**Cost**	**Incremental cost**	**Effectiveness (QALMs)**	**Incremental Effectiveness**	**Incremental cost effectiveness ratio**	**Net Monetary Benefit**	**Cost/ effectiveness**
NAT	£109879.65	£8627.90	21.27	3.68	£2344.16	-£56709.34	£5166.40
SF	£101251.75		17.59			-£57282.91	£5757.02

b) Using West of Scotland Pancreatic Unit Data

Strategy	**Cost**	**Incremental cost**	**Effectiveness (QALMs)**	**Incremental Effectiveness**	**Incremental cost effectiveness ratio**	**Net Monetary Benefit**	**Cost/ effectiveness**
NAT	£117426.89	£29126.08	26.71	8.48	£3433.07	-50652.80	£4396.42
SF	£88300.81		18.23			-42736.66	£4844.86

Figure 48: Results of cost-effectiveness analysis of NAT *versus* SF for resectable pancreatic cancer based on Willingness-to-Pay (WtP) set at £2500 per QALM.

a) Using Synthesised Data



b) Using West of Scotland Pancreatic Unit Data

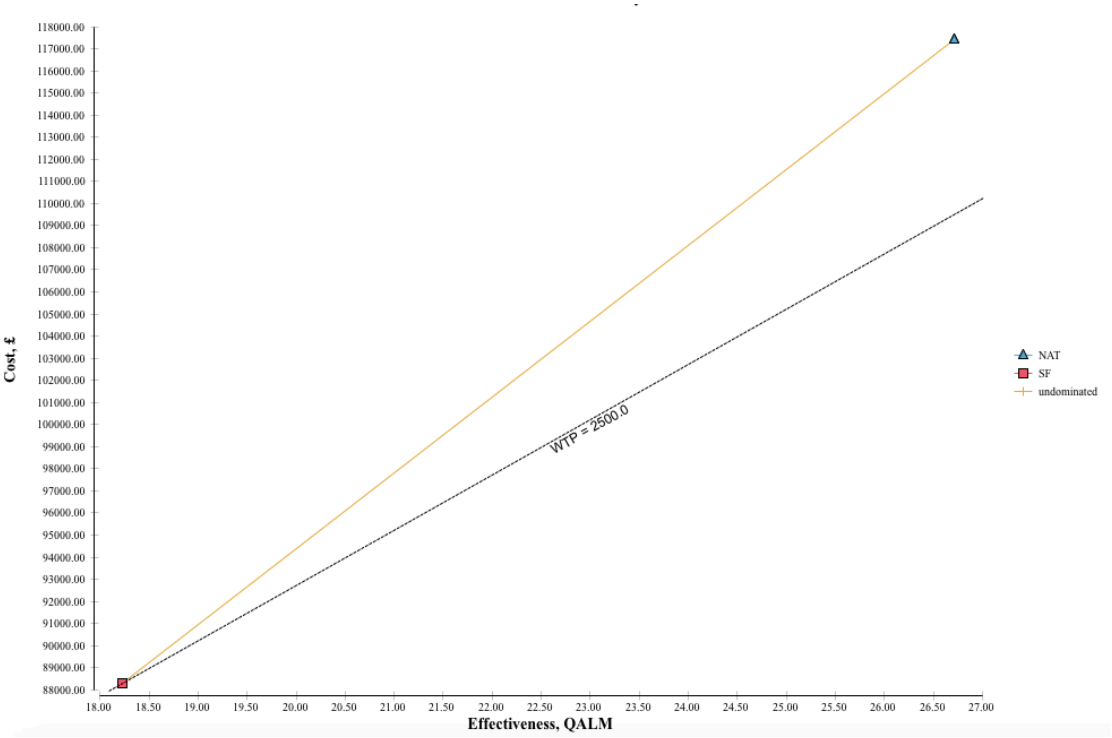
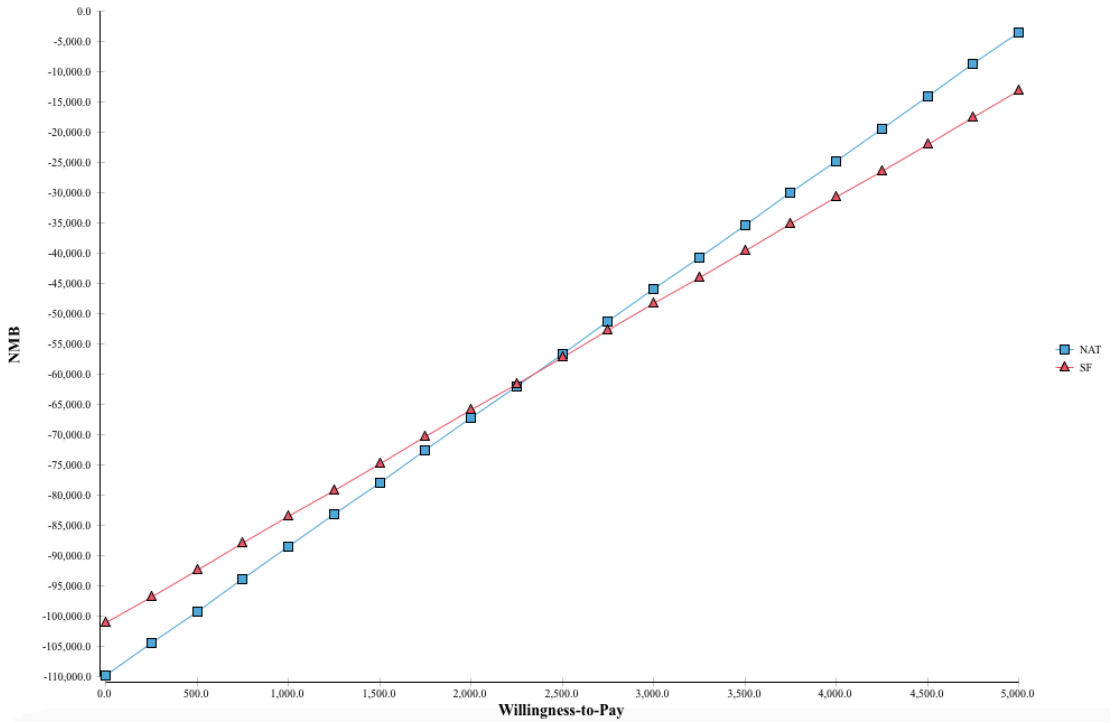
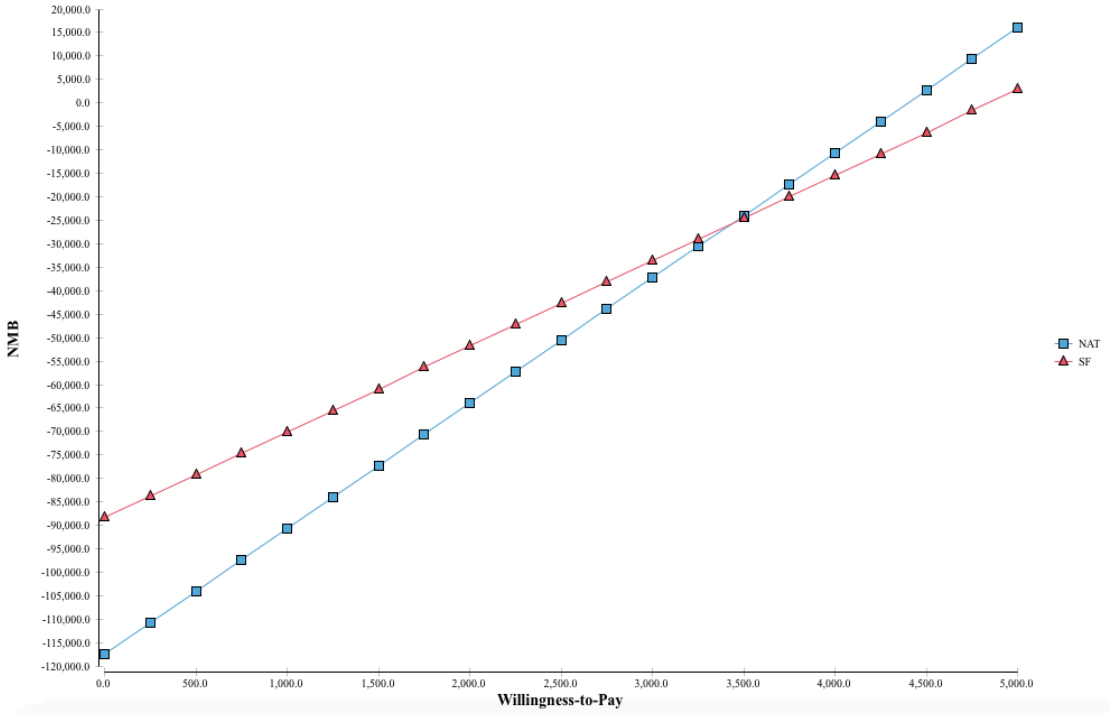


Figure 49: Results of cost-effectiveness analysis of NAT *versus* SF for resectable pancreatic cancer across range of Net Monetary Benefit (NMB) and Willingness-to-Pay (WtP)

a) Using Synthesised Data



b) Using West of Scotland Pancreatic Unit Data



Sensitivity Analysis

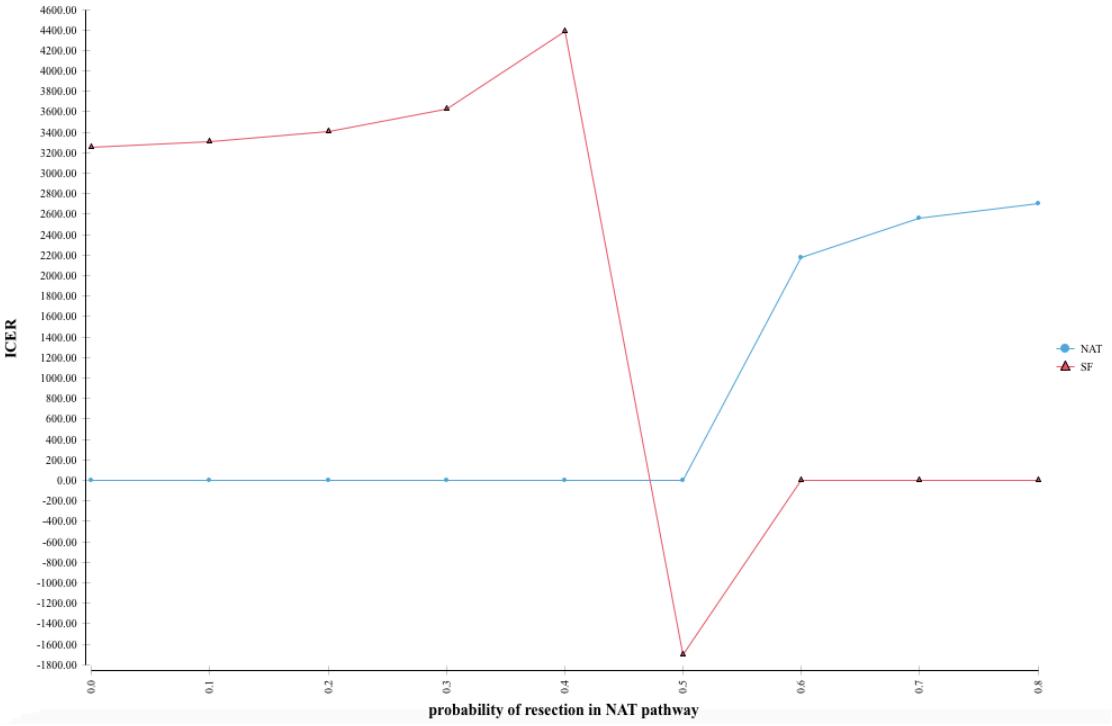
Altering discounting to 6% for costs and 1.5% for benefits did not alter the overall findings summarised in figure 108. At these rates of discounting using synthesised data NAT gave 21.98 QALMs at a cost of £106059.25 with a cost-effectiveness ratio of £4825.88 per QALM. SF gave 18.10 QALMs at a cost of £97746.32 and a cost-effectiveness ratio of £5399.69 per QALM. NAT therefore had an incremental cost of £8312.93 more than SF for an incremental effectiveness of 3.87 QALMs and an ICER of £2145.29. Triangulating these results with the West of Scotland Pancreatic Unit database analysis, NAT gave 27.72 QALMs at a cost of £112907.32 with a cost-effectiveness ratio of £4074.42 per QALM. SF gave 18.78 QALMs at a cost of £85226.37 and a cost-effectiveness ratio of £4537.45 per QALM. NAT therefore had

an incremental cost of £27680.96 more than SF for an incremental effectiveness of 8.93 QALMs and an ICER of £3100.34.

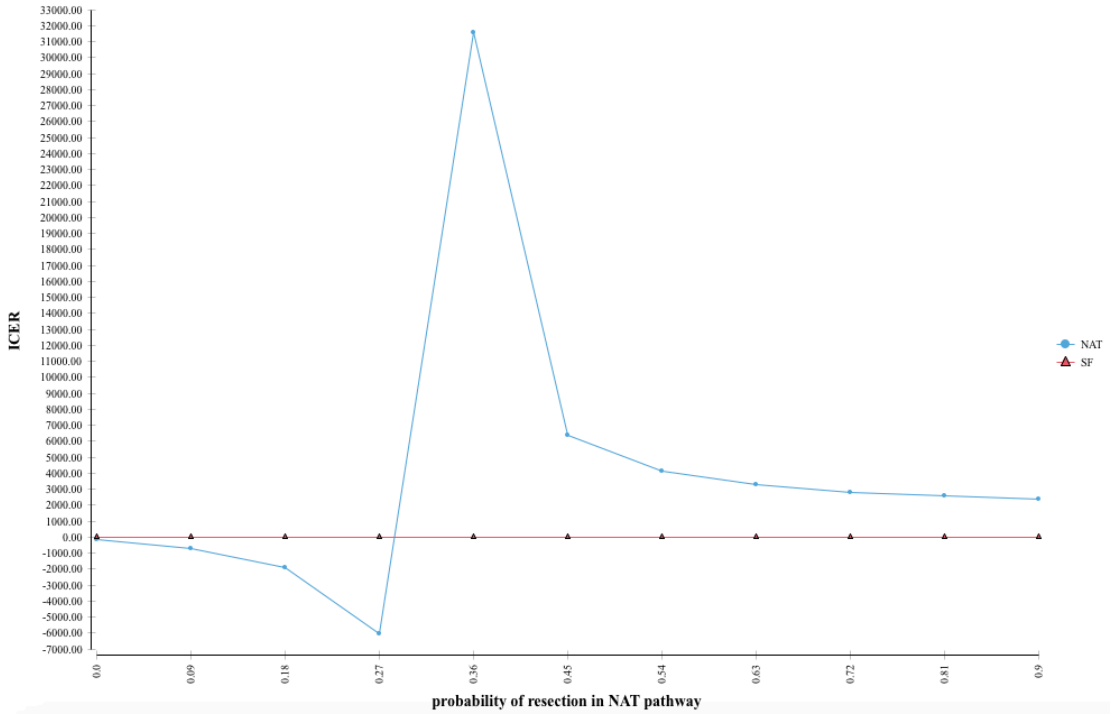
Deterministic sensitivity analysis of both synthesised and patient data showed that the main driver of ICER was the receipt of multimodal treatment. In particular deterministic sensitivity threshold analysis of synthesised data demonstrated a probability of resection greater than 48% in the NAT pathway was required for this treatment pathway to be the more effective option. For West of Scotland Pancreatic Unit database this threshold was 34%. The main driver of the ICER was found to be the probability of undergoing resection in the NAT pathway (figure 50).

Figure 50: Incremental cost-effectiveness ratio (ICER) for NAT and SF pathways across range of probabilities for receiving resection in NAT pathway.

a) Using Synthesised Data



b) Using West of Scotland Pancreatic Unit Data



The results of probabilistic Monte Carlo sensitivity analysis set to 10000 iterations are summarised in table 38 and Figure 51. Optimal pathway selection with WtP set at £2,500 per QALM showed a selection frequency of 77% for NAT with synthesised data and 72% with West of Scotland Pancreatic Unit data. Furthermore during probabilistic sensitivity analyses of the West of Scotland Pancreatic Unit database NAT pathway was found not to breech the WtP threshold and therefore, as with synthesised data analysis, could be considered a cost-effective alternative to traditional SF pathway (Figure 52).

Table 38: Summary of Monte Carlo Probabilistic Sensitivity Analysis

a) Using Synthesised Data

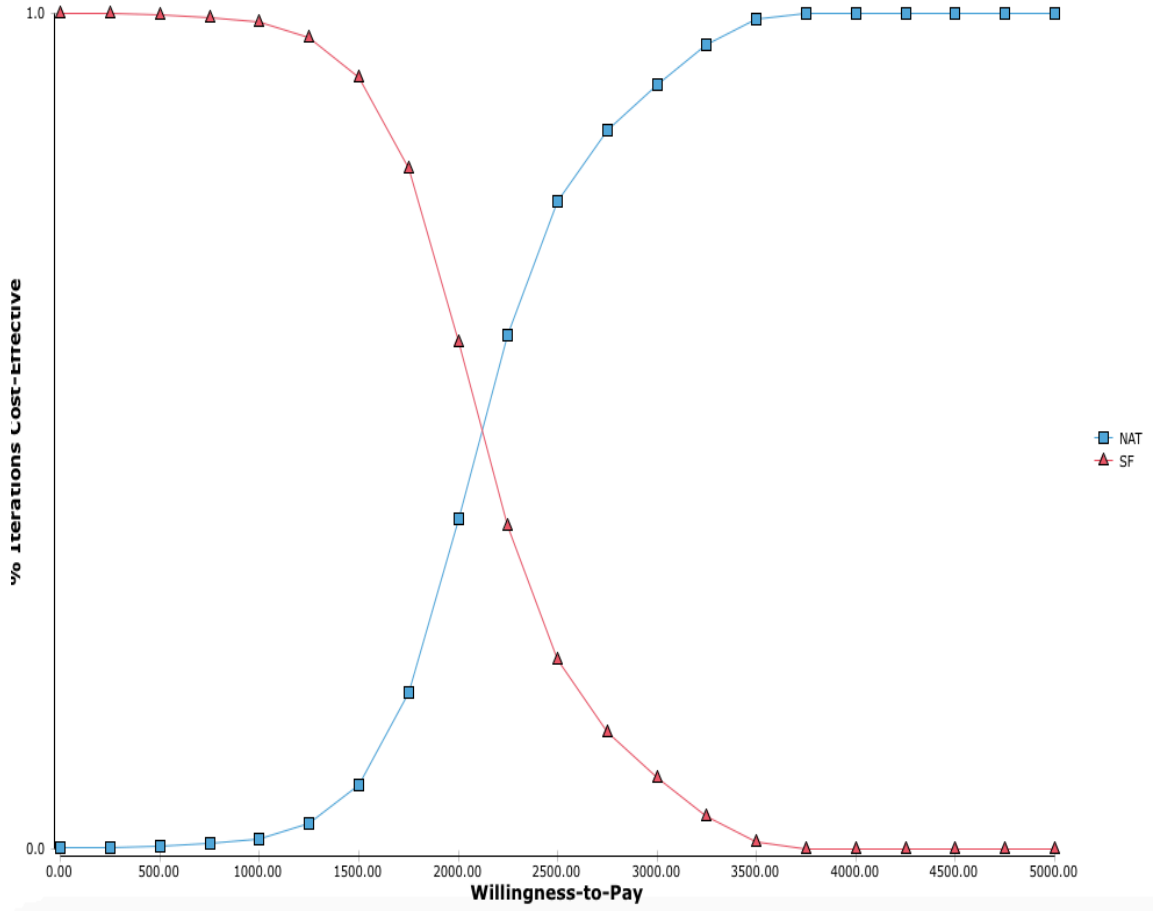
	Mean		Range		Standard Deviation	
Pathway	SF	NAT	SF	NAT	SF	NAT
Cost	£101256.83	£109882.63	£60142.62 to £107681.66	£103899.41 to £116363.85	5814.60	1599.27
Effectiveness (QALMs)	17.59	21.27	7.70 to 19.66	19.32 to 23.36	1.42	0.51

b) Using West of Scotland Pancreatic Unit Data

	Mean		Range		Standard Deviation	
Pathway	SF	NAT	SF	NAT	SF	NAT
Cost	£92713.52	£112809.03	£74229.24to £108324.92	£107483.79 to £116207.52	4960.33	1426.96
Effectiveness (QALMs)	20.38	27.21	11.16 to 30.07	26.44 to 27.71	2.74	0.21

Figure 51: Cost Effectiveness acceptability curve

a) Using Synthesised Data



b) Using West of Scotland Pancreatic Unit Data

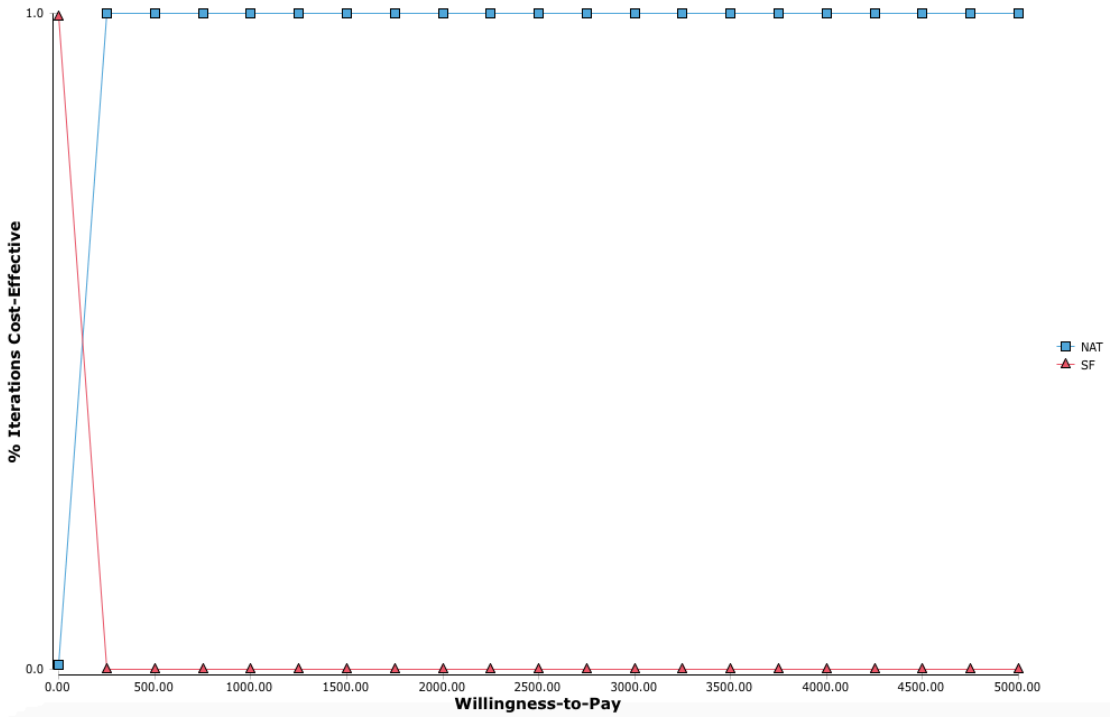
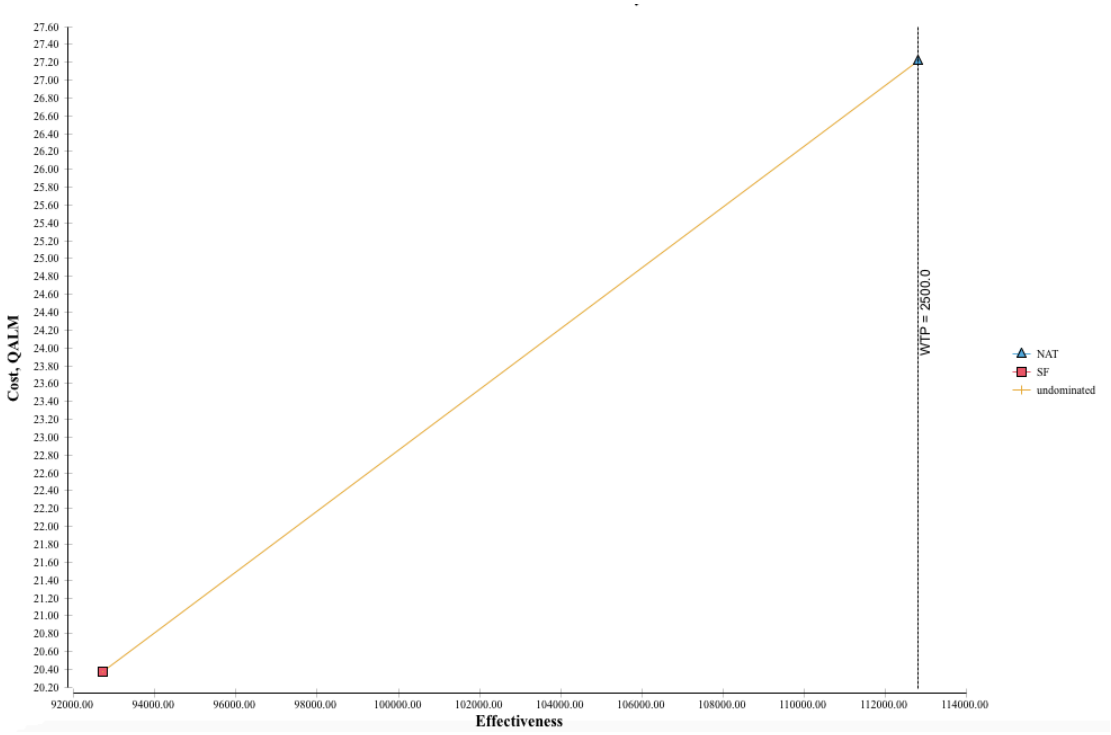


Figure 52: Results of cost-effectiveness analysis of NAT versus SF for resectable pancreatic cancer based on Willingness-to-Pay (WtP) set at £2500 per QALM.



4.5.2 Results: DES Model

Using synthesised data NAT gave 16.45 QALMs at a cost of £81934.19. SF gave 13.84 QALMs at a cost of £69630.42. NAT therefore had an incremental cost of £12303.77 more than SF for an incremental effectiveness of 2.61 QALMs and an ICER of £4708.51 (table 39). NAT was therefore found to exceed WtP threshold set at £2,500 per QALM (Figure 53). However, as this model included palliative care costs including palliative chemotherapy, a WtP threshold of up to £5833.33 per QALM (£70,000 per QALY) can be acceptable and would deem NAT a cost-effective alternative to traditional SF pathway (Figure 54 and Figure 55).

Figure 53: Results of cost-effectiveness analysis of NAT versus SF for resectable pancreatic cancer based on Willingness-to-Pay (WtP) set at £2500 per QALM.

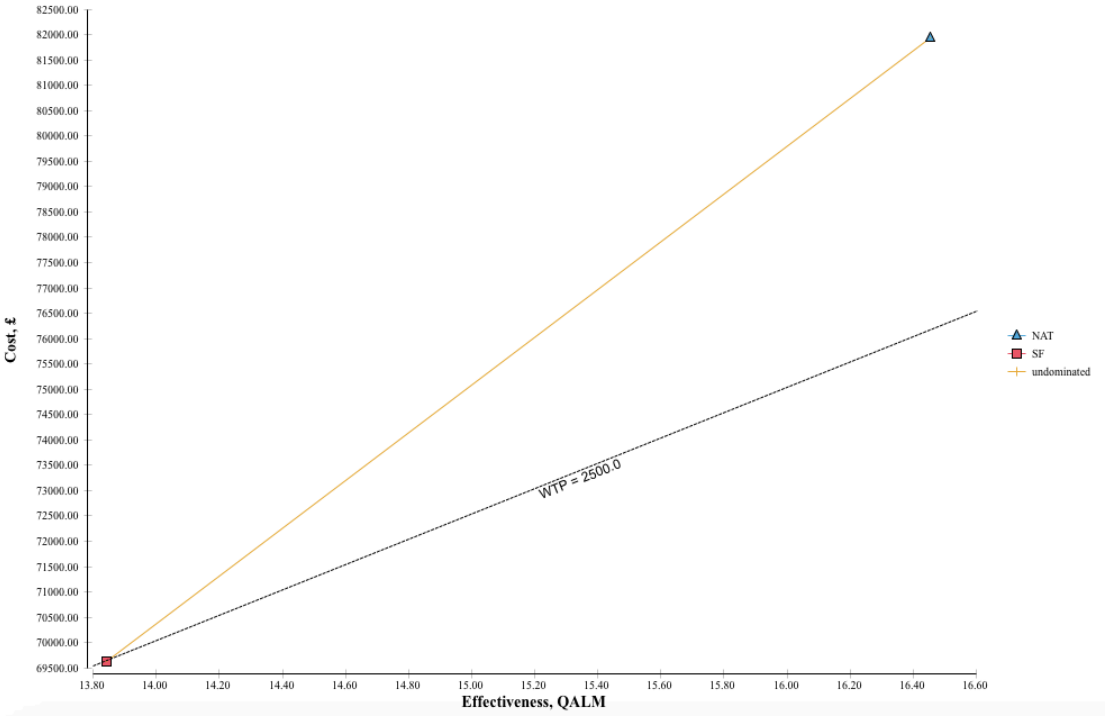


Figure 54: Results of cost-effectiveness analysis of NAT versus SF for resectable pancreatic cancer based on Willingness-to-Pay (WtP) set at 5833.33 per QALM.

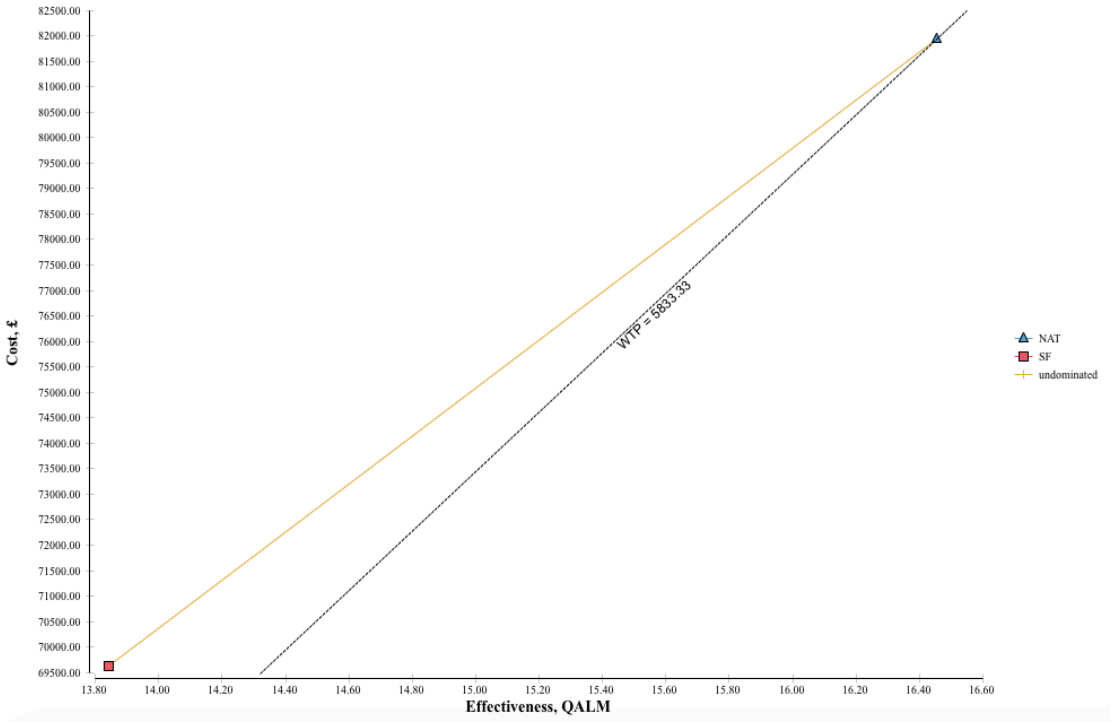
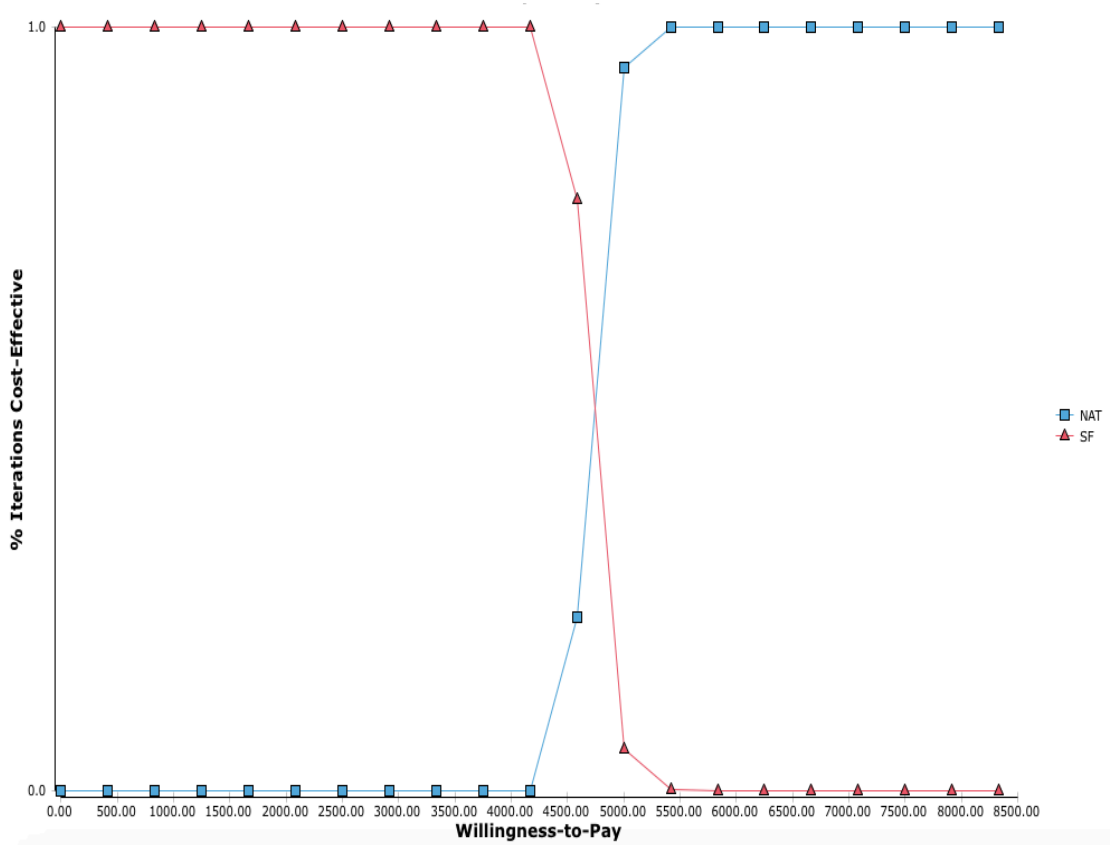
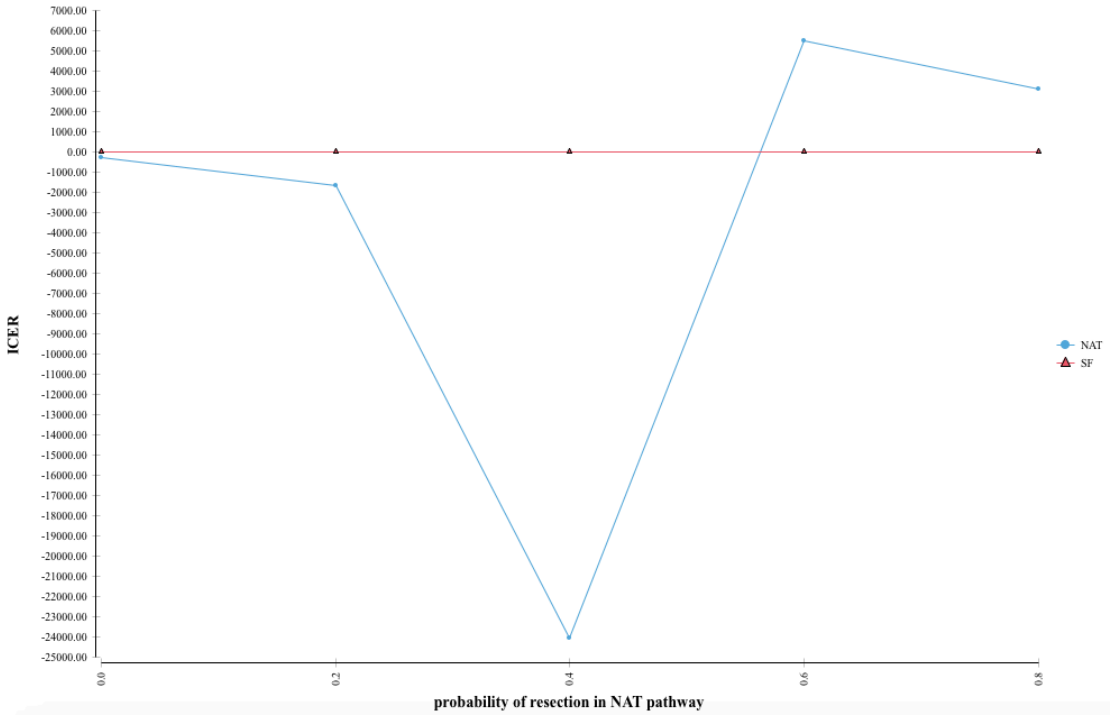


Figure 55: Cost Effectiveness acceptability curve



As with the cost-effectiveness analysis using Markov modeling the main driver of the ICER was found to be the probability of receiving resection in the NAT pathway (Figure 56). Altering the discount rate for cost to 6% and 1.5% for benefits did not alter these overall findings. At these rates of discounting NAT gave 16.82 QALMs at a cost of £81941.17 with a cost-effectiveness ratio of £4870.61 per QALM and an ICER £4434.25. SF gave 14.05QALMs at a cost of £69656.57 and a cost-effectiveness ratio of £4956.63 per QALM. NAT therefore had an incremental cost of £12284.60 more than SF for an incremental effectiveness of 2.77QALMs.

Figure 56: Incremental cost-effectiveness ratio (ICER) for NAT and SF pathways across range of probabilities for receiving resection in NAT pathway.



When these results were triangulated with the cost-effectiveness analysis of the West of Scotland Pancreatic Unit NAT gave 21.60 QALMs at a cost of £72083.26. SF gave 13.87 QALMs at a cost of £45813.65. NAT therefore had an incremental cost of £26219.61 more than SF for an incremental effectiveness of 7.73 QALMs and an incremental cost-effectiveness ratio of £3390.51 (table 39; table 40). In this analysis NAT was therefore found to exceed the WtP threshold set at £2,500 per QALM but when the WtP was increased to reflect the inclusion of palliative interventions NAT pathway could be considered a cost-effective alternative to SF pathway (Figure 57; Figure 58).

Table 39: Summary of Results of Cost-Effectiveness Analysis

a) Using Synthesised Data

Strategy	**Cost**	**Incremental cost**	**Effectiveness (QALMs)**	**Incremental Effectiveness**	**Incremental cost effectiveness ratio**	**Net Monetary Benefit**	**Cost/ effectiveness**
NAT	£81934.19	£12303.77	16.45	2.61	£4708.51	-£13377.86	£4979.74
SF	£69630.42		13.84			-£11961.97	£5030.95

b) Using West of Scotland Pancreatic Unit Data

Strategy	**Cost**	**Incremental cost**	**Effectiveness (QALMs)**	**Incremental Effectiveness**	**Incremental cost effectiveness ratio**	**Net Monetary Benefit**	**Cost/ effectiveness**
NAT	£72083.26	£26219.61	21.60	7.73	£3390.51	-£17936.73	£3336.45
SF	£45813.65		13.87			-11934.50	£3306.31

Table 40: Summary of Results of Probabilistic Sensitivity Analysis

a) Using Synthesised Data

	Mean		Range		Standard Deviation	
Pathway	SF	NAT	SF	NAT	SF	NAT
Cost	£69614.34	£81913.62	£68672.42 to £70729.69	£80174.64 to £83586.57	297.98	436.35
Effectiveness (QALMs)	13.84	16.44	13.65 to 14.07	15.96 to 16.88	0.06	0.13

b) Using West of Scotland Pancreatic Unit Data

	Mean		Range		Standard Deviation	
Pathway	SF	NAT	SF	NAT	SF	NAT
Cost	£45240.89	£70512.02	£44722.85 to £45664.73	£69824.18 to £71130.15	133.28	191.01
Effectiveness (QALMs)	13.88	21.60	13.64 to 14.17	21.10 to 22.10	0.08	0.13

Figure 57: Results of cost-effectiveness analysis of NAT versus SF for resectable pancreatic cancer based on Willingness-to-Pay (WtP) set at £2500 per QALM.

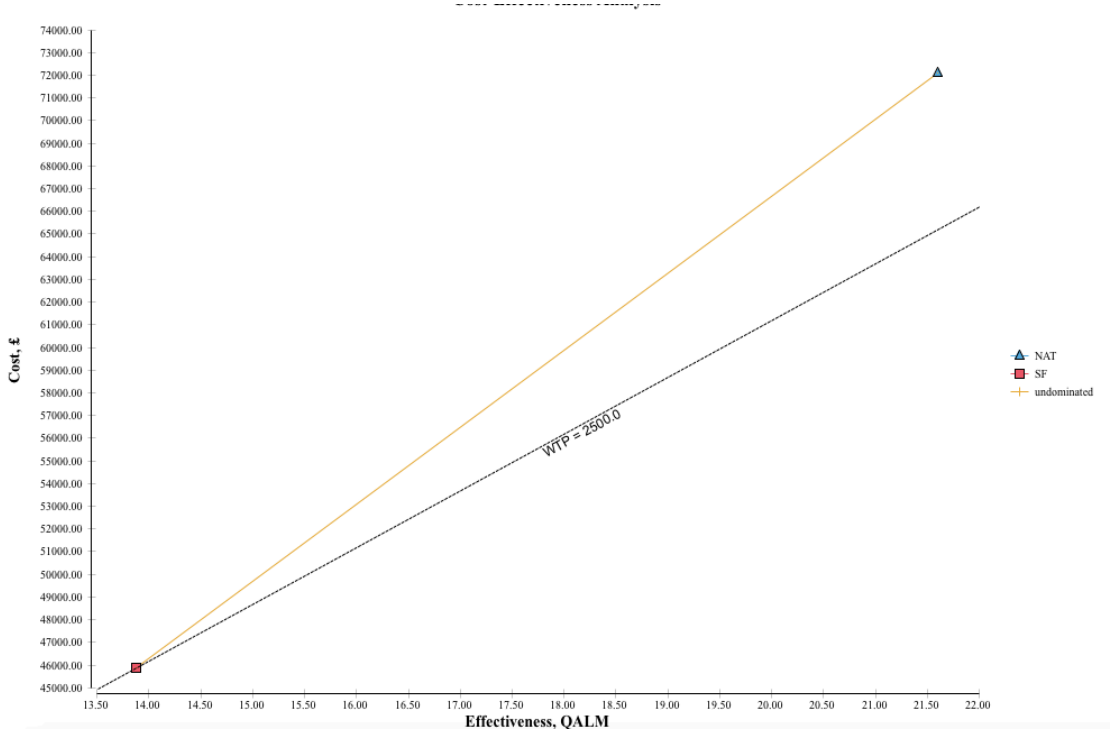


Figure 58: Results of cost-effectiveness analysis of NAT versus SF for resectable pancreatic cancer based on Willingness-to-Pay (WtP) set at 5833.33 per QALM.

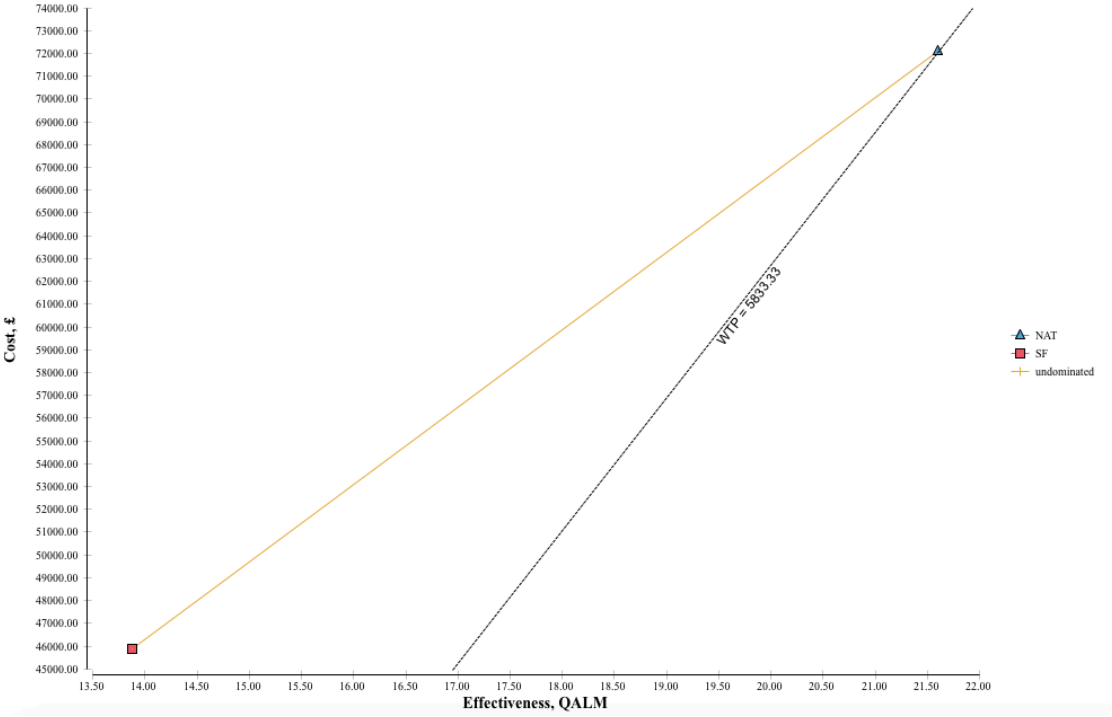
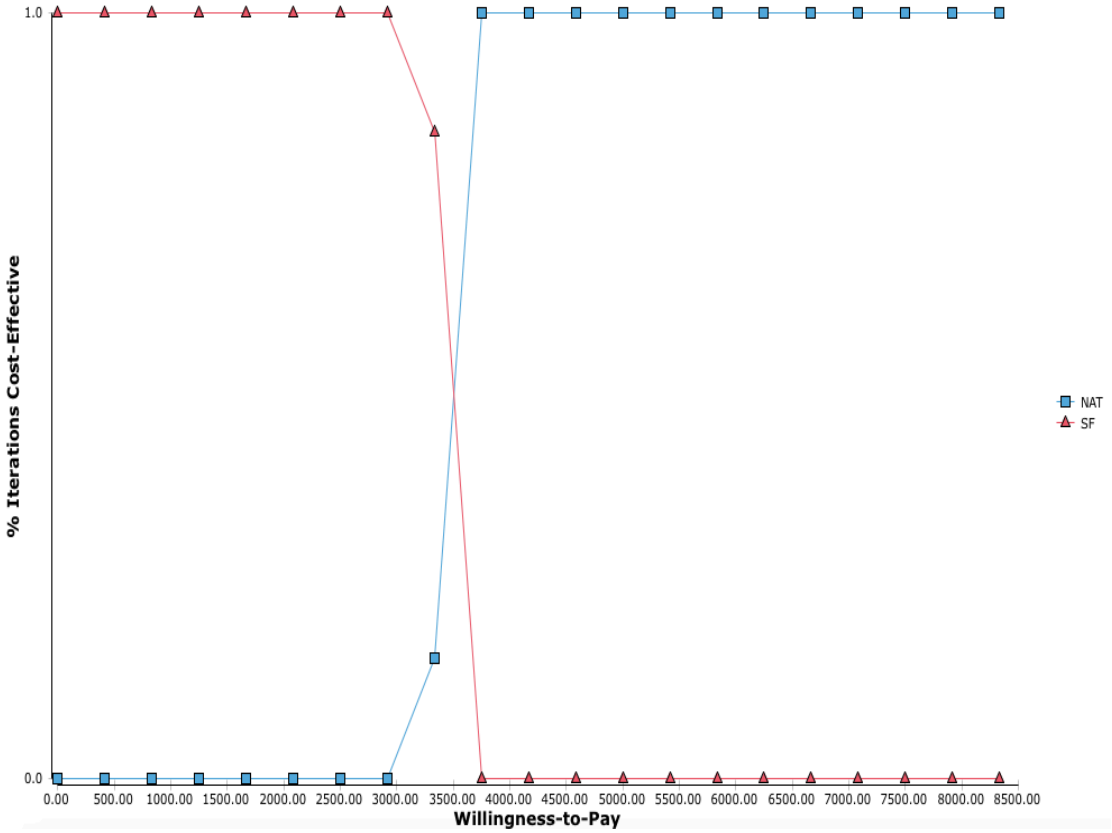
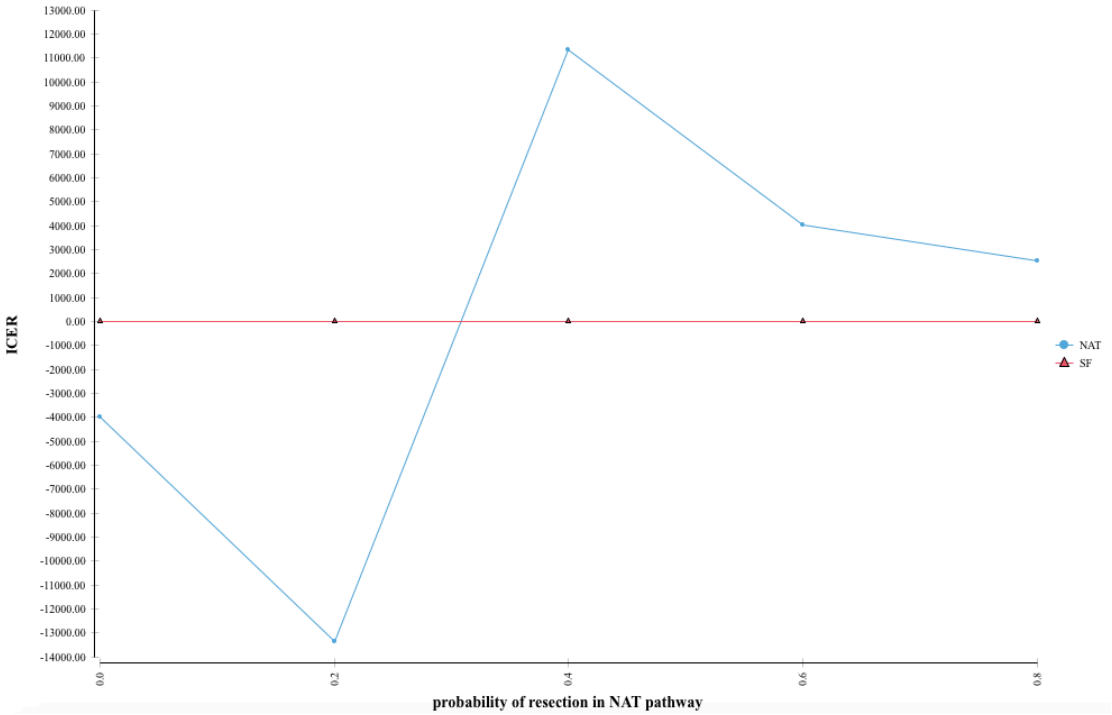


Figure 59: Cost Effectiveness acceptability curve



Once again in sensitivity analysis the main driver of the ICER was found to be the probability of resection in the NAT pathway (Figure 60). Altering the discount rates to 6% for costs and 1.5% for benefits did not alter these overall findings. At these rates of discounting NAT gave 22.09 QALMs at a cost of £70510.14 with a cost-effectiveness ratio of £3191.53 per QALM and an ICER £3169.15. SF gave 14.12 QALMs at a cost of £45238.94 and a cost-effectiveness ratio of £3204.17 per QALM. NAT therefore had an incremental cost of £25271.20 more than SF for an incremental effectiveness of 7.97 QALMs.

Figure 60: Incremental cost-effectiveness ratio (ICER) for NAT and SF pathways across range of probabilities for receiving resection in NAT pathway.



Discussion

The limitations stemming from the FUPS characteristics of the data used within the Markov and DES models have been discussed in sections 4.2 to 4.4 and once again the cost-effectiveness analysis presented here is presented as a partial remnant. Our Markov cost-effectiveness analysis study found that NAT gave 21.27 QALMs compared to 17.59 QALMs in the SF pathway with an ICER of £2344.16. The robustness of these findings was substantiated by extensive sensitivity analysis whereby every variable, treatment probability, cost and outcome was altered between highest and lowest observed value and all model probabilities were sampled

from the entire data distribution through probabilistic sensitivity analysis. Repeating the cost-effectiveness analysis using data from the West of Scotland Pancreatic Unit Database then provided triangulation of these results. This analysis also showed that NAT pathway was more effective (26.71 QALMs *versus* 18.23 QALMs) with an ICER of £3433.07. Monte Carlo probabilistic sensitivity analysis of the Markov models modeling both synthesised and patient data showed that NAT pathway was cost-effective 77% and 72% of the time respectively over 10000 iterations.

The DES cost-effectiveness analysis corroborated the findings from the Markov analysis that NAT pathway was more effective giving but cost more. Using synthesised data NAT gave 16.45 QALMs at a cost of £81934.19 whereas SF gave 13.84 QALMs at a cost of £69630.42. NAT therefore had an ICER of £4708.51. Using the West of Scotland Pancreatic Unit database NAT gave 21.60 QALMs at a cost of £72083.26 whereas SF gave 13.87 QALMs at a cost of £45813.65. NAT therefore had an ICER of £3390.51. The variation between outcomes of effectiveness between models can be explained by the fact that the memory-less property of the Markov model makes it less well equipped to handle individual patient data, which can result in reduced accuracy due to depletion of susceptibles and an over simplification of assumptions (Caro *et al*., 2010; Miettinen & Caro, 1989). DES modeling allows the model to capture a patient's experience in terms of events and also to track changes in patient characteristics, health status and treatment history in relation to their impact on outcomes (Pan *et al*., 2018; Caro *et al*., 2010). This flexibility has led some to conclude that DES modeling is a more

efficient method for engaging with complexity by modeling patient level data to perform micro simulation. The results from section 4.4 go some way to supporting this assumption. For the purposes of this cost-effectiveness analysis study DES modeling allowed a more detailed analysis of palliative interventions including the impact of second line palliative chemotherapy when disease reoccurred, or progressed despite NAT, or was found to have progressed at the time of attempted resection in the SF pathway.

Deterministic sensitivity analysis showed that the main driver of the ICER when modeling both synthesised data and the West of Scotland Pancreatic Unit database in both the Markov and DES models was the receipt of multimodal treatment, with resection in the NAT pathway being the most significant driver of ICER.

As previously discussed the survival findings from our study are in keeping with the few existing RCTs, meta-analysis and Markov decision analysis studies comparing NAT and SF which have report some survival benefit with NAT. Comparison of our findings with the only other existing cost-effectiveness analysis of NAT *versus* SF for the treatment of resectable pancreatic cancer is unfeasible due to the lack of reporting of ICER in this study (Abbott *et al*., 2013). Our study findings also suggested a benefit in terms of effectiveness with NAT and highlight the importance of patient selection and embracing multidisciplinary approach in aiming to deliver multimodal treatment (Abbott *et al*., 2013). However the conclusions drawn from our study are more cautious. Whilst Abbott *et al*. (2013) suggested that the benefit of NAT approach is in filtering out patients with more

aggressive tumours, in whom surgery would be ultimately futile, hence avoiding unnecessary associated risks and costs of surgery, opponents of NAT would argue that this demonstrates losing the window of resectability (Asare *et al*., 2016; Lee *et al*., 2016). Furthermore unlike the existing cost-effectiveness study by Abbott *et al.* (2013) our study also included costs associated with follow-up and palliative care input from the time of disease recurrence which is significant considering the survival benefit reported with NAT, which was more significant in their study therefore could have affected overall results (Abbott *et al*., 2013) particularly as it has been reported that costs of care escalate during end-of-life care.

The inclusion of palliative care input also raises an important issue regarding boundary setting when modeling a complex system for cost-effectiveness analysis that echo the discussion from previous work by Cilliers (2005a; 2008). As highlighted in these works this also has ethical implications concerning the drawing of conclusions from this analysis particularly considering the lack of complete knowledge surrounding the management of pancreatic cancer. In the UK NICE have set the WtP threshold for curative treatments at between £20,000 and £30,000 per QALY. However, for palliative cases a WtP of up to £70,000 (£5833.33 per QALM) is accepted. For many who have their pancreatic cancer resected survival outcomes remain poor and this raises the question as to whether the lower WtP threshold set for curative procedures is acceptable for resectable pancreatic cancer. The analysis presented here used the WtP threshold of £30,000 per QLAY or £2,500 per QALM as baseline analysis. When the Markov model was populated with West of

Scotland Pancreatic Unit data it was found that a WtP of £42,000 per QALY (£3,500 per QLAM) was required for NAT pathway to be cost effective. Within the DES model the WtP had to be £57000 per QALY (£4750 per QALM) in the model using synthesised data and £42,000 per QALY (£3500 per QALM) in the model using West of Scotland Pancreatic Unit data. This highlights not only the implications of WtP threshold for pancreatic cancer, but also the impact of boundary setting when modeling complex systems for the purposes of cost effectiveness analysis and the potentially detrimental impact of drawing overly simplistic conclusions from analysis of FUPS data. The West of Scotland Pancreatic Unit database was small and did not include details of follow-up or administration of palliative chemotherapy at the time of disease recurrence for those with resected disease or those who failed to proceed to resection in either pathway. It therefore had the potential for bias, which limited the generalisability of findings.

In conclusion the management of resectable pancreatic cancer is challenging and the role of NAT remains controversial. Rather than seeking to prove that one pathway is more cost-effective than the other, by engaging with the FUPS characteristics of the available data and the complexity of the system being modeled, this analysis shows that NAT could be considered a cost effective alternative to the traditional SF pathway for resectable pancreatic cancer when the entirety of the patient pathway including palliative treatments were modeled and the WtP threshold was adapted to reflect this. However, the significant finding that was beginning to emerge is that the key driver of the ICER was receipt of multimodal treatment therefore

better personalised patient selection could result in more cost-effective care delivery. The next section therefore explores how statistical modeling, viewed through the lens of complexity theory, can be taken even further to facilitate better patient selection through individualised patient prediction of outcomes across competing treatment strategies.

Section 4.6 Bayesian belief network (BBN) to select optimal treatment pathway for resectable pancreatic cancer

Publications resulting from this analysis:

Bradley, A., Van der Meer, R. and McKay, C.J. (2019). 'A prognostic Bayesian network that makes personalized predictions of poor prognostic outcome post resection of pancreatic ductal adenocarcinoma'. *PLoS One,* 14(9): e0222270. doi.org/10.1371/journal.pone.0222270

Bradley, A. *et al.* (2019). 'Personalized prognostic Bayesian network for pancreatic cancer: delivering personalized pancreatic cancer management throughout the patient journey'*. Pancreatology*, 19 (S1). pp. S31-S32.

Bradley, A. *et al.* (2019). 'Making personalized predictions of poor outcome post resection of pancreatic ductal adenocarcinoma (PDAC): a prognostic Bayesian network with pre and post-operative application'. *Pancreatology*, 19 (S1), pp. S122. P6-13

Bradley, A., Van Der Meer, R., McKay, C.J. (2020) 'Optimising outcomes for resectable pancreatic cancer by learning lessons from military strategy and the stock market: creation of a prognostic Bayesian belief network that makes personalised pre and post-operative predictions of outcomes across competing treatment strategies.' *British Journal of Surgery*: accepted

Abstract

Background: Survival outcomes for pancreatic cancer remain poor. Surgical resection is the only potentially curative treatment but approximately 10% present with resectable disease with 5year survival for resected cases between only 7% and 25%. Adjuvant therapy after surgery is required to prolong survival but up to 50% of patients fail to receive adjuvant therapy. Neoadjuvant therapy has emerged as an alternative treatment strategy but carries the risk of losing the window of resectability. Risks of failure therefore exist throughout both treatment pathways. Thus far in this thesis decision-analysis modeling has suggested that the selection of the optimal treatment pathway depends on individual patient and tumour factors.

Methods: This section presents a Bayesian Belief Network (BBN) model that evaluates the risk of failure and consequence factors across surgery-first and neoadjuvant treatment pathways for potentially resectable PDAC.

Results: To demonstrate the application of BBN it was applied to the database of a tertiary referral pancreatic unit. Area Under the Curve (AUC) of the Receiver Operating Characteristic Curve (ROC) for pre-operative prognostic predictions ranged from 60% to 70% and 74% to 94% for poor and good prognosis respectively with ranges reflecting the models' ability to cope with missing data. AUC for prognostic updating ranged from 70% to 80% and 75% to 97% for poor and good prognosis respectively again with ranges reflecting

the impact of missing data points.
Conclusion: BBN is capable of providing personalised predictions of poor and good prognosis post resection of PDAC. This can support clinical decision action and patient counseling by identifying superior pathway selection at individualised level, and justify resource allocation by identifying patients at higher risk of failure who require additional actions to reduce risk respectively.

Introduction

The decision-analysis presented in this thesis have broadened the current debate surrounding the treatment of potentially resectable pancreatic cancer by suggesting that superior pathway selection is more complex and depends on the interactions of multiple individual patient and tumour factors rather than simply whether the tumour is technically resectable (Bradley *et al*., 2018; Bradley & Van Der Meer, 2019a).

Cancer is estimated to cost the European Union economy €126 billion with 40% spent on healthcare alone (Luengo-Fernandez *et al*., 2013). It is therefore imperative that lessons drawn from the application of operational research be applied to the management of pancreatic cancer. In congruence with Lawrence's (1976) definition, failure risk has been defined as the combination of probability and impact severity of a particular situation that negatively impacts the ability of infrastructure to obtain objectives. Such outcomes result from multiple complex interactions between a plethora of variables. The challenges involved in the management of pancreatic cancer can

be viewed as the evaluation of risk and prediction of outcome within a complex and dynamic system. Lessons can therefore be drawn from the innovative application of operational research in other fields. To illustrate, when assessing the risk of water mains failure, failure could manifest as structural integrity, hydraulic capacity and water quality (Kabir *et al*., 2015). With pancreatic cancer failure can manifest as non-completion of multimodal therapy, complications, or a short post-operative survival time. What both examples have in common is the importance of determining the cause and effect of probability of failure. Furthermore, a successful risk assessment programme provides predictive tools to assess the causes and consequences of failure, recommend prioritisation and enable the development of long and short term management plans (Moustafa, 2010). This was achieved with water mains failure by using a Bayesian Belief Network (BBN) to identify vulnerabilities and justify decision action (Kabir *et al*., 2015). Multiple points of vulnerability exist across pancreatic management pathways concerning patient, disease and treatment factors. A similar approach could help to better guide decision actions to reduce risks prior to undergoing interventions. Parallels can also be drawn with areas of military operational planning and banking. With the former Bayesian network analysis modeled existing planning process concepts, with uncertainties and subjective judgments clearly represented, to perform impact analysis and determine which course of action is most likely to achieve a desired outcome (Falzon, 2004). For pancreatic cancer management the 'plan-of-attack' is either a surgery first or neoadjuvant pathway. Modeling these existing planning processes, including uncertainty and subjective judgments, could

determine the best course of action through impact analysis on survival outcome. Bayesian networks to predict bankruptcy could offer another important parallel. Bankruptcy samples are usually small and bankrupt firms tend to have missing data (Sun & Shenoy, 2007) therefore share many of the FUPS characteristics. Only a small percentage of pancreatic cancer cases are resectable and clinicians often make decisions despite missing data and uncertainty permeating the existing body of evidence pertaining to conclusive pathway superiority.

The objective of this study is to develop a new and effective BBN model to evaluate the risk of poor prognosis (1year or less) and predict good prognosis (3-years or more) post resection of PDAC. In this research risk factors that lead to failure events, or factors that lead to risk reduction or positive events, and the consequence factors that result from either, are studied. It is hoped that the BBN will aid health professionals to better assess and address vulnerability factors proactively, plan service resource allocation, and achieve best possible survival outcomes for each individual patient. The remainder of this section is organised as follows. Section 4.6.1 discusses the application of probabilistic concepts underlying the BBN. In Section 4.6.2 the structure of the BBN is explained and results of the BBN models' performances when validated against the West of Scotland Pancreatic Unit database is presented in Section 4.6.3. Section 4.6.4 then concludes by summarising and discussing the findings from the assessment and validation of the BBNs' performances.

4.6.1 Application of Probabilistic Concepts Underlying the BBN

Background

Several model development studies have attempted to predict post resection survival for PDAC using a variety of variables (Appendix P). As previously discussed in chapter 2 their clinical application is limited, relying primarily on post-operative data, failing to encompass the emerging role of neoadjuvant therapy and lacking the ability to determine patients who would, or would not, benefit from competing treatment pathways. The majority lack validation and are based on single institution databases, which limits generalisability and potentiates bias. Mainly based on logistic regression techniques they fundamentally regard prognosis as an isolated event at a pre-determined time hence neglecting the dynamic nature of the care processes reflected in the unfolding relationships between variables with expected patient outcomes evolving as more information becomes available (Verduijn *et al*., 2007). A review of machine-learning to support decision-making in the management of pancreatic cancer in chapter 2 revealed limitations including: use of small single centre databases limiting generalisability, lack of transparency, and lack of external validation (Bradley *et al*., 2019b).

BBNs have been proven to perform well across a plethora of domains as both classification and predictions models (Sun & Shenoy, 2007). They have the ability to model complex relationships between variables, update predictions when new information is learned, and

incorporate subjective human knowledge therefore are easily interpreted by clinicians (Verduijn *et al*., 2007; Lewis & Vollmer, 2012; School *et al*., 2013;Velikova *et al.,* 2014). Unlike most regression techniques they do not depend on the underlying distribution of variables (Pearl, 1988; Sun & Shenoy, 2007). Furthermore they represent the relationship between variables by the direct acyclic graph, which makes them more transparent and intuitive than other machine-learning techniques (Pearl, 1988; Kabir *et al*., 2015).

To allow BBN to become a useful decision aid for PDAC management challenges of variable selection, generalisability, missing data and pre-operative application must be addressed. Within a BBN missing data is handled through probabilistic inference, with predictions being made based on global averages of the patient population. Previous BBNs in healthcare have relied on expert judgment to quantify the required probability relationships. However, for complex problems it is difficult to establish mutual relationships among nodes considering that the number of conditional probabilities increase exponentially to the number states of the parent and child nodes (Nadkarni & Shenoy, 2001; Tang & McCabe, 2007). Under such complex and large conditional probability tables (CPT) expert derived conditional probabilities can become inconsistent and it has been found to be more reliable to construct CPTs from data (Tang & McCabe, 2007; Hager & Andersen, 2010; Kabir *et al*., 2015). To address these challenges we propose a novel two stage weighting process, adapted from Zhao & Weng (2011) to synthesise PubMed survival analysis data.

Bayesian Network Construction

BBNs, also referred to as acyclic directed graphs, are based on probability theory and models relationships between variables, know as nodes, with arcs depicting informational or causal relationships from parent to child nodes (Pearl, 1988; Hager & Andersen, 2010). Each node has a defined and exclusive set of states and the dependencies between nodes are quantified through a set of CPTs whereby the conditional probability of a child node is defined by the state of each of its parent nodes (Kabir *et al*., 2015). Nodes that do not have parent nodes are reduced to the unconditional probability (UP) structure. Where the UPs of a basic input parameter is not known a priori, equal weight are applied to states through the principle of insufficient reasoning (Kabir *et al.*, 2015).

Through Bayes' theorem, BBNs can explicitly represent the conditional probability dependencies between variables, which has been proven to be an effective way of handling uncertainty (Ismail *et al*., 2011; Sun & Shenoy, 2007). In a BBN, the updated probability for n number of mutually exclusive parameters X_i, where $(i= 1,2,3...n)$, and given observed data, Y, can be computed as:

$$p(X_i|Y)= \frac{p(Y|Xi)\text{x } p(Xi)}{\sum j p(Y|Xj)p(Xj)} \quad \text{(i)}$$

where the posterior probability of X given the condition that Y occurs is represented by $p(X|Y)$, the posterior probability of Y given the condition that X occurs is represented by $p(Y|X)$, the prior occurrence probability of X is denoted by $p(X)$ and the marginal occurrence of Y is denoted as $p(Y)$ (Pearl, 1988). In this sense this is often viewed as the likelihood distribution (Pearl, 1988).

This holds several advantages when modeling treatment pathways for potentially resectable PDAC. Through Bayes theorem the prior distribution and observed data are combined to update knowledge in the form of the posterior distribution (Pearl, 1988; Fenton & Neil, 2019). Therefore BBNs allow the modeling of relationships between variables at various stages of the healthcare process, with predictions of outcomes evolving throughout the process by utilising all available patient data at that time (School *et al*., 2013). This means that the model could not only make predictions of outcome pre-operatively but also perform prognostic updating at the post-operative stage of the patient journey. Where patient information is limited probabilistic inference can still make predictions based on global averages of the patient population (Verduijn *et al.,* 2007; Lucas *et al.*, 2004). As more information becomes available the predictions become more patient specific (Verduijn *et al.,* 2007). Furthermore, BBN have the flexibility to perform bottom-up inference (inferring the state of the parent node from the observed state of the child node) and top-down inference (inferring the state of the child node given the observed state of the parent node) (Cockburn & Tesfamariam, 2012; Ismail *et al*., 2012). This is also known as

diagnostic and decision-analysis respectively and is called marginalisation, which is employed to compute the reliability of networks based on statistical data (Cai *et al*., 2013; Nadkarni & Shenoy, 2001; Poropudas & Virtanen, 2011). For the pancreatic cancer model this process will also allow the testing of the model to perform scenario, or "what if" testing to anticipate cause and impact of failure events such as developing side effects from therapy or complications from surgery.

BBN for potentially resectable PDAC

The conditional probabilities used in equation (i) can be obtained from expert opinion. However, as previously explained, such an approach can loss reliability in a larger, more complex BBN (Tang & McCabe, 2007; Hager & Andersen, 2010; Kabir *et al.*, 2015). The other traditional approach is to acquire conditional probabilities from training data (Kabir *et al.*, 2015). However, the percentage of patients presenting with resectable disease is low, the acquisition of sufficiently large and detailed databases is therefore difficult. This has resulted in previous prediction models being limited by small, biased datasets and lacking generalisability. Larger databases such as cancer registry do not contain sufficient enough patient level detail.

Evidence Synthesis for Conditional Probabilities

PubMed is an online database containing over 29million citations for biomedical literature. It was searched for all papers published since 2000 that reported survival analysis of patients with poor (1-year or

less) and good (3years or more) post PDAC resection survival time. This yielded 77 papers (n = 31,214) and 67 papers (n=48691) from which the BBNs predicting poor and good prognosis respectively were constructed. Survival analysis studies mainly study factors associated with either good or poor prognosis, and the variables and their categorisation were therefore independent. BBNs were therefore created separately to predict poor and good prognosis.

Adapting methods from Zhao and Weng (2011), information was extracted on all variables analysed for their association with post PDAC resection prognostic outcome. The original weight for each variable (w^o_i) was calculated as:

$$w^o_i = P_i/N_i$$

where P_i represents the number of studies where the variable was found through multivariate analysis to have a statistically significant association with prognostic outcome. N_i represents the total number of studies in the body of evidence where the variable underwent statistical analysis for its association with prognostic outcome (Zhao & Weng, 2011). This placed each variable on a scale from 0 to 1. In this way P_i and N_i ratio summarise the collective evidence, including conflicting findings, for the variables association with post PDAC resection prognosis (Zhao & Weng, 2011). The original weight, w^o_i, then underwent a secondary process of normalisation to more accurately reflect its weighting within the existing body of evidence. Normalised weights, w_i, were defined as:

$$w_{i=} w^o_i \ (\max(pw^o_1, pw^o_2, \ldots pw^o_n / \max(ps_1, ps_2, \ldots ps_n)$$

whereby $\max(pw^o_1, pw^o_2, \ldots pw^o_n)$ is the sum of the study populations (pw^0) reporting on the variable, and $\max(ps_1, ps_2, \ldots ps_n)$ is the sum of

the study populations of all included studies (Zhao & Weng, 2011). The original weights and the normalised weight, w_i is in the range of 0 to 1 and therefore reflects both conflicting findings of the significance of each variable in relation to the prognostic outcome in question, and the weight of significance of the body of evidence pertaining to each individual variable in relation to the entire existing body of evidence relating to the prognostic outcome in question (Zhao & Weng, 2011) (Table 41).

Table 41: Weighted ranked variables from PubMed studies

Ranked Order	Variable/ Node associated with good prognosis	Variable/ Node associated with poor prognosis
1	Lymph Node Negative for cancer cells	Lymph Node Positive for cancer cells
2	Clear resection margins on removed tumour (R0 resection)	Lymph node ratio of positive to negative lymph nodes for cancer cells
3	Albumin blood test level	Tumour Grade
4	Adjuvant Treatment Completed	Tumour Size
5	American Joint Committee on Cancer (AJCC) Stage: staging system for describing extent of disease progression	Evidence of disease at margins of resected tumour
6	Neoadjuvant treatment response	Adjuvant Therapy not received
7	Tumour Size	T stage: disease staging determined by size and extent of the tumor
8	Tumour Grade: grade determined by degree of abnormality identified within cells of the tumour	Pre treatment tumour marker blood test Ca 19-9
9	Lymph node ratio of positive to negative lymph nodes for cancer cells	American Joint Committee on Cancer (AJCC) Stage: staging system for describing extent of disease progression
10	T stage: disease staging determined by size and extent of the tumor	Vascular Involvement: tumour has invaded vessels
11	Pre treatment tumour marker blood test Ca 19-9	Peri Neural Involvement (PNI): tumour has invaded peri neural structures
12	Tumour Location on the pancreas: head of pancreas, body of pancreas, tail of pancreas	Age
13	Age	Modified Glasgow prognostic score: scoring system based on blood results giving levels of inflammation
14	Peri Neural Involvement (PNI): tumour has invaded peri neural structures	Tumour marker blood test CEA>5
15	Vascular Involvement: tumour has invaded vessels	Performance Status: assessment of how well the patient is such as amount of pre-existing co-morbidities
16	Post-operative blood test of tumour level Ca19-9	Tumour Location on the pancreas: head of pancreas, body of pancreas, tail of pancreas
17	Tumour marker blood test CEA>5	Post-operative blood test of tumour level Ca19-9
18	Blood transfusion required during operation	Blood Transfusion required during operation
19	Modified Glasgow prognostic score: scoring system based on blood results giving levels of inflammation	Albumin blood test level
20	Jaundice: whether the patient's bilirubin level is raised on blood test	Neutrophil Lymphocyte Ratio (NLR) >2: ratio of blood test results
21	Neutrophil lymphocyte ratio (NLR): ratio of blood test results	Jaundice: Bilirubin > 40
22	Performance Status: assessment of how well the patient is such as amount of pre-existing co-morbidities	Diabetes
23	Smoking History	Smoking History
24	Lymphocytes raised on blood test	Response to Neoadjuvant Treatment
25	Diabetes	Body Mass Index (BMI)
26	White Cell Count (WCC) raised	CRP: inflammatory marker blood test

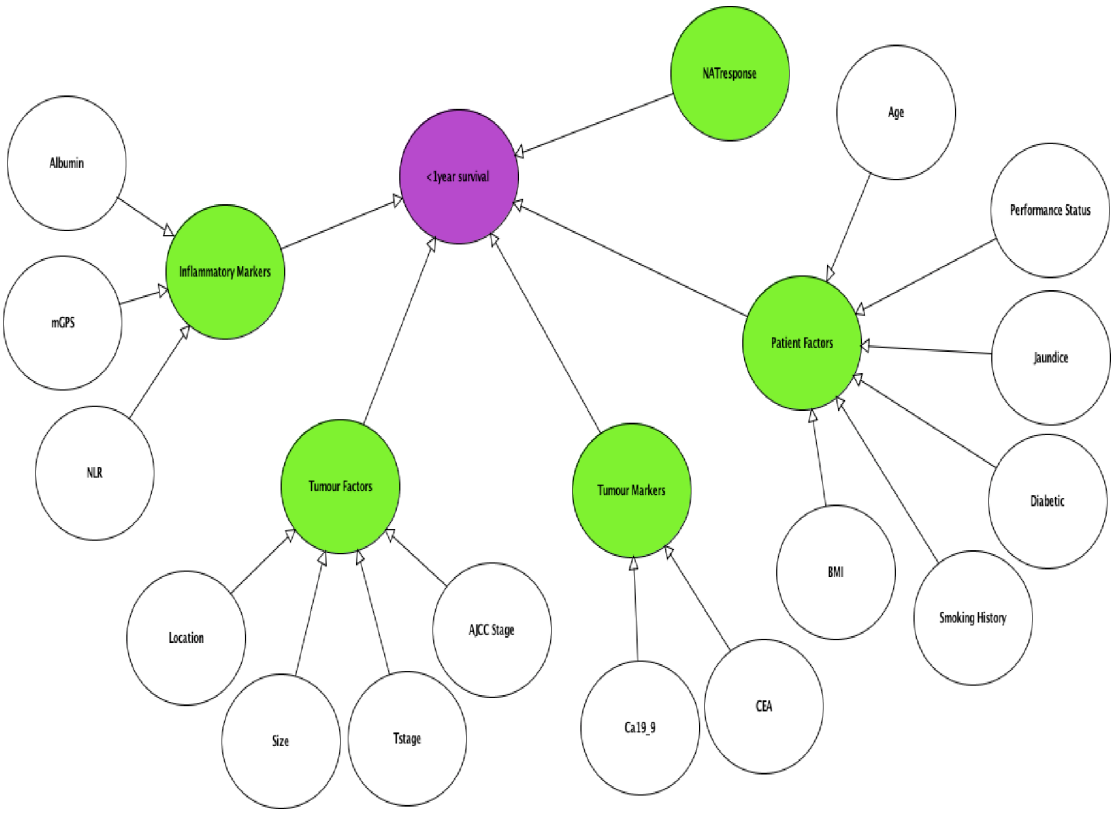
27	Pre-operative biliary stenting performed	Post-operative complication
28	Post-operative complication	LDH: blood test
29	Body Mass Index (BMI)	Hypercalcaemia: high calcium level on blood test
30	CRP: inflammatory marker blood test	Other hematological abnormalities: Elevated Leucocytes, neutrophils, lymphocytes. Thrombocytopenia. Low hemoglobin,
31	LDH: blood test	
32	Neutrophils raised	
33	Anaemia	

3.6.2 BBN Structure

Variables that were ranked within the top 25 were extracted from the list displayed in Table 41 and used to structure the BBN using AgenaRisk version 7.0 software. Each variable was treated as a ranked parent node and linked through causative arcs to their respective child nodes. Variables known pre-operatively were used to construct the pre-operative models (Figure 61) and post-operatively known variables were added to construct the BBN for prognostic updating (Figure 62).

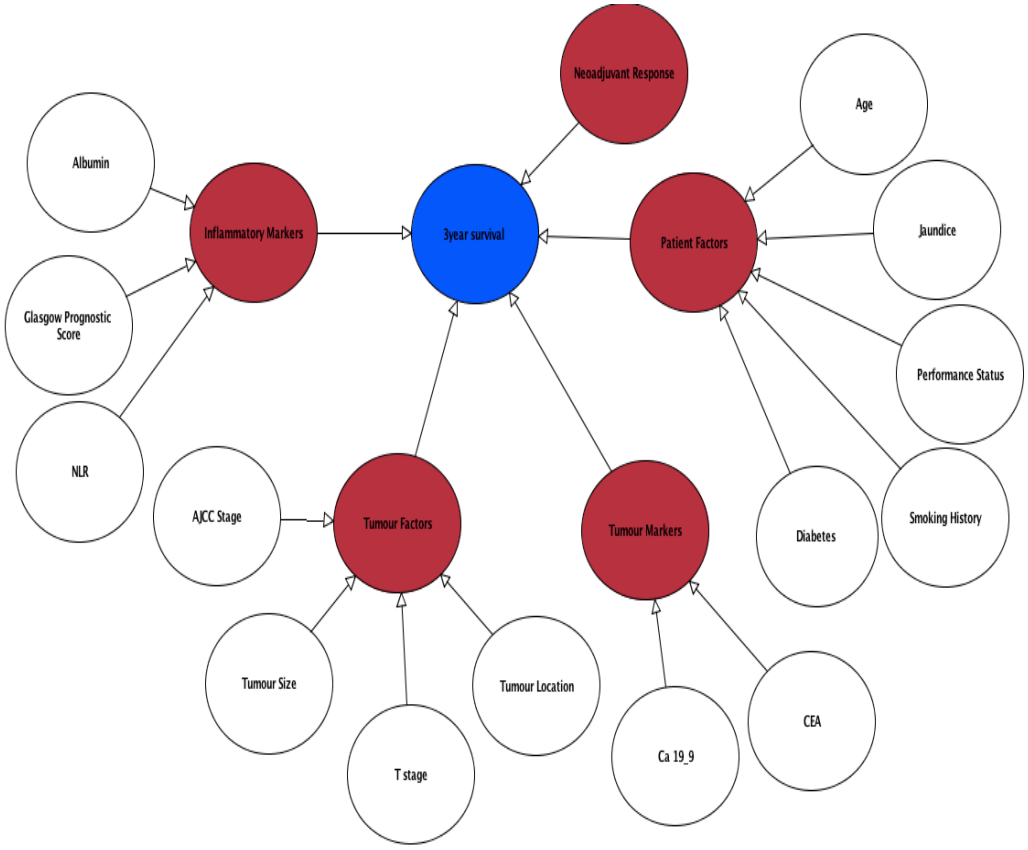
Figure 61: Structure of pre-operative BBNs.

a) Predicting poor prognosis



Parent nodes in white. Child nodes in green are ranked based on weighted mean of weighted parent nodes. Output node in purple.

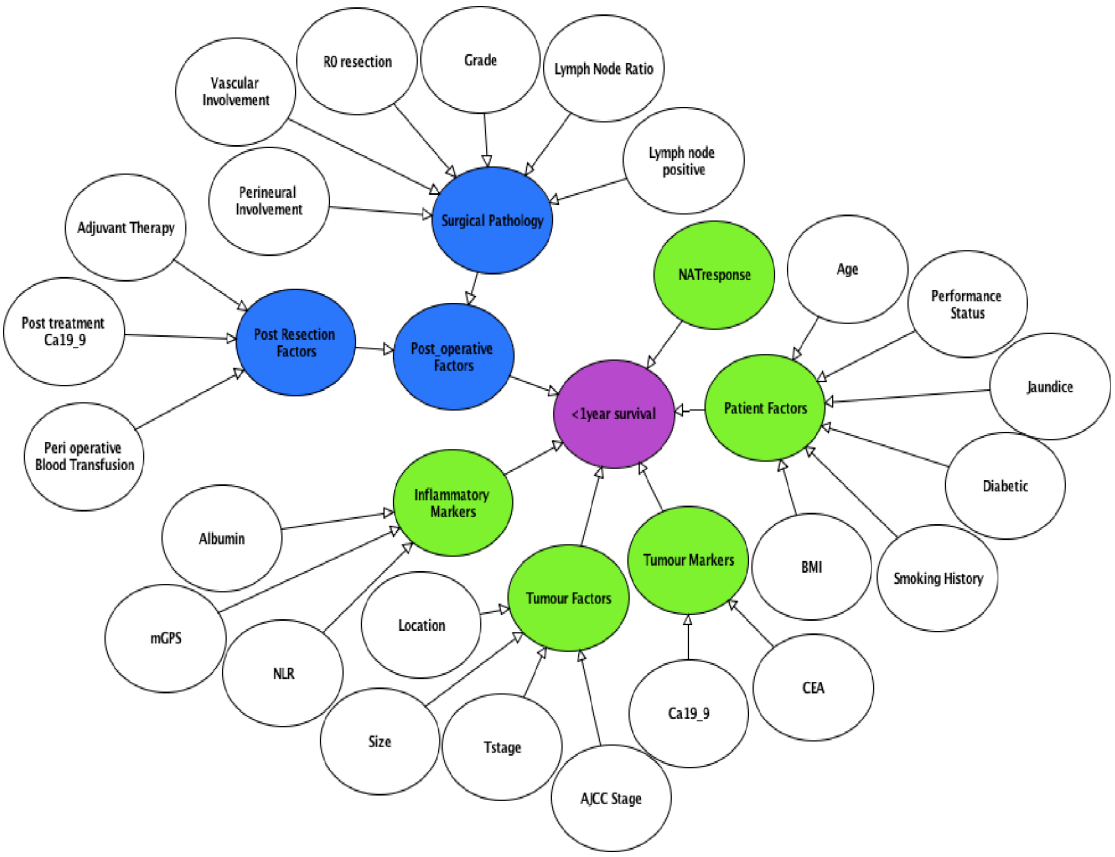
b) Predicting good prognosis



Parent nodes in white. Child nodes in red are ranked based on weighted mean of weighted parent nodes. Output node in blue.

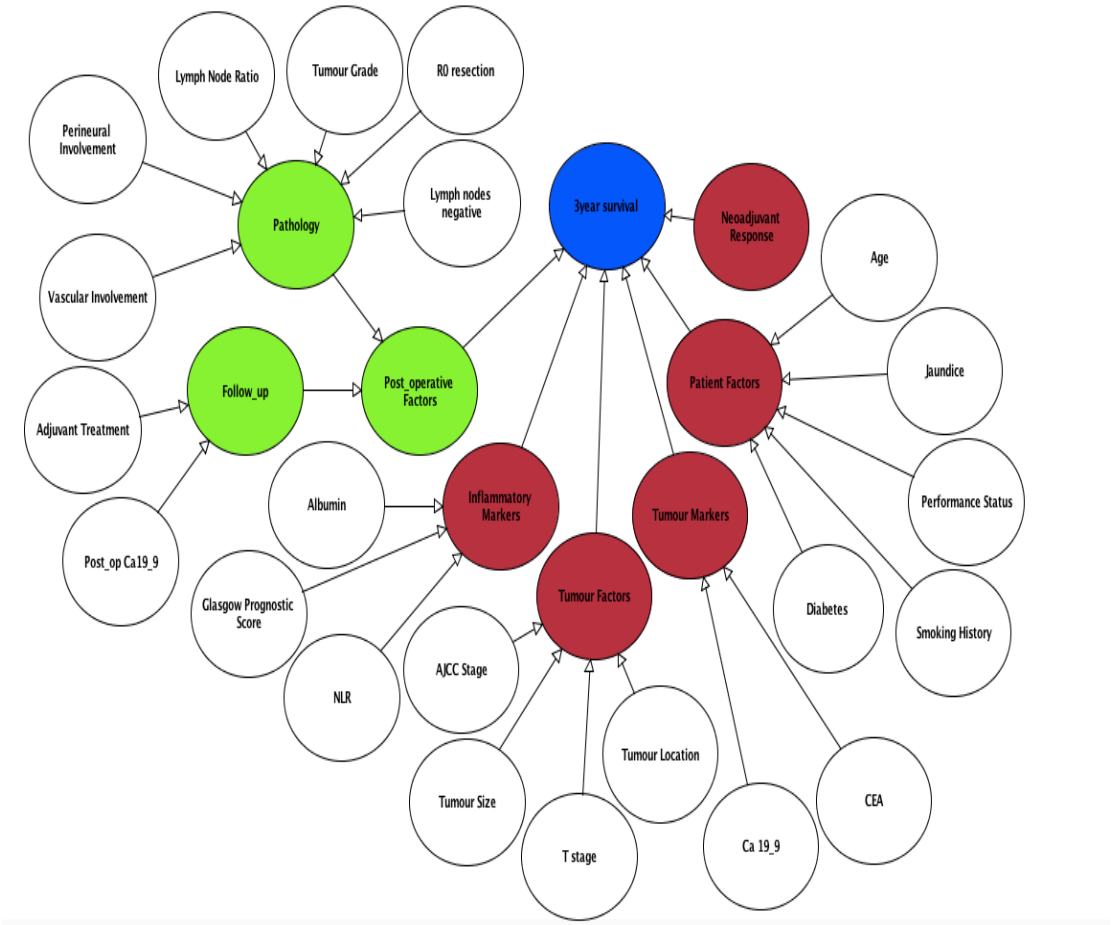
Figure 62: Structure of post-operative BBNs.

a) Predicting poor prognosis



Parent nodes in white. Pre-operative child nodes in green, and post-operative child nodes in blue. Output node in purple. NLR= neutrophil lymphocyte ratio; AJCC stage= American Joint Committee on Cancer stage

b) Predicting good prognosis



Parent nodes in white. Pre-operative child nodes in red, and post-operative child nodes in green. Output node in blue. NLR= neutrophil lymphocyte ratio; AJCC stage= American Joint Committee on Cancer stage

The node probability table for each child node was calculated using the truncated Normal (TNormal) statistical distribution as this provides finite end points between 0 and 1 and has been proven to generate accurate node probability tables for BBNs involving ranked nodes with ranked parent nodes (Fenton & Neil, 2019). The normalised weighting of each parent node was used as the weighted

mean of the TNomal distribution from which to calculate the node probability tables. The definitions and categorisation of input data for each node within the BBNs are outlined in Table 42 and Table 43 and were determined by how values were reported in the PubMed studies. They and the overall model structure were also approved by a panel of experts from a pancreatic tertiary referral centre.

Table 42: Definition and categorisation of input data for each node in the BBN predicting good prognostic outcome

Node	Node States	Definition
Albumin	Normal Low	=/> 35 g/l < 35 g/l
mGPS (modified Glasgow Prognostic Score)	0 1 2	0= CRP</= 10mg/L and albumin >/= 35 g/L 1= CRP > 10mg/L 2= CRP> 10mg/L and albumin <35 g/L
NLR (Neutrophil Lymphocyte Ratio)	<2 >2	
Location	HOP Body/Tail	Head of Pancreas Location other than HOP
Size	< 2cm >2cm	
T stage	Tis/T1/T2 T3/T4	
AJCC Stage	0 1 2 3 4	As per American Joint Committee on Cancer definition [19]
Ca19-9	<50 50-999 >1000	<50 U/mL 50-999 U/mL >1000 U/ mL
CEA	<5 >5	<5 ng/mL >5 ng/mL
Age	< 70 >70	Under 70 years Equal to or over 70 years
Performance Status	Good Moderate Poor	ASA 1 ASA2 ASA 3 or worse
Jaundice	No Yes	Bilirubin < 40μmol/l Bilirubin >40μmol/l
Diabetes	No Yes	
Smoking History	Non-smoker Smoker	
Neoadjuvant Response	Response/stable resectable Progression/ Unresectable	Radiological response or stable disease that is still resectable Radiological evidence of stable disease that is still borderline resectable or locally advanced/progression/ unresectable disease
Lymph nodes negative	Yes No	
R0 resection	Yes No	
Tumour Grade	G1/G2	As per American Joint Committee on Cancer definition [19]: Well/moderate

	G3/G4	differentiation, low/intermediate grade Poorly differentiated, high grade
Lymph Node Ratio (LNR)	<0.2 >0.2	
Perineural Involvement (PNI)	No Yes	
Vascular Involvement	No Yes	
Adjuvant Therapy	Yes No	
Post-operative Ca19-9	Normal Raised	<37 U/mL >37 U/mL

Table 43: Definition and categorisation of input data for each node in the BBN predicting poor prognostic outcome

Variable/ Node	Node Status	Definition
Lymph Node Positive	Yes No	
Lymph node ratio	<0.3 >0.3	Ratio of the number of positive lymph nodes to the total number of lymph nodes removed
Tumour Grade	G1/G2 G3/G4	As per American Joint Committee on Cancer definition [19]: Well/moderate differentiation, low/intermediate grade Poorly differentiated, high grade
Tumour Size	< 2cm >2cm	
R0 Resection	No Yes	No microscopic evidence of any residual tumour
Adjuvant Therapy	No Yes	
T stage	T1 T2 T3 T4	
Pre treatment Ca 19-9	<50 50-999 >1000	<50 U/mL 50-999 U/mL >1000 U/ mL
AJCC (American Joint Committee on Cancer) Stage	0 1 2 3 4	As per AJCC definition
Vascular Involvement	Yes No	
Perineural Involvement (PNI)	Yes No	
Age	< 70 >70	Under 70 years Equal to or over 70 years
mGPS (modified Glasgow Prognostic Score)	0 1 2	0= CRP</= 10mg/L and albumin >/= 35 g/L 1= CRP > 10mg/L 2= CRP> 10mg/L and albumin <35 g/L
CEA>5	<5 >5	<5 ng/mL >5 ng/mL
Performance Status	 Good Moderate Poor	As defined by American Society of Anaestheologits (ASA) classification ASA 1-2 ASA 2-3 ASA >3
Tumour Location	HOP	Head of Pancreas (HOP)

	Body/Tail	Location other than HOP
Post treatment Ca19-9	<120 >120	<120 U/mL >120 U/mL
Prei operative Blood Transfusion	Yes No	
Albumin	Normal Low	=/> 35 g/l < 35 g/l
Neutrophil Lymphocyte Ratio	<5 >5	
Jaundice	No Yes	Bilirubin < 40μmol/l Bilirubin >40μmol/l
Diabetes	No Yes	
Smoking	Non-smoker Smoker	
Response to Neoadjuvant Treatment	Stable Progression/ Unresectable	Radiological response or stable disease that is still resectable Radiological evidence of progression/ unresectable disease
BMI	Normal Low	Body Mass Index (BMI) above 18 BMI equal or under 18

Implementation of BBNs

The pre-operative BBN output nodes were defined using the states: inflammatory markers, tumour factors, tumour markers, patient factors and response to neoadjuvant treatment, each of which was defined as high, medium and low risk and were defined by different parent nodes depending on whether the BBN was predicting poor or good prognosis. When performing prognostic updating the output node was additionally defined by post-operative factors which was based on surgical pathology from the resected tumour and post-operative factors which included events during recovery from surgery. Again post-operative factors were defined as high, medium and low risk for the given prognostic outcome. Using the algorithm provided by commercially available software AgenaRisk version 7.0 conditional probabilities are generated for the BBNs.

Scenario analysis

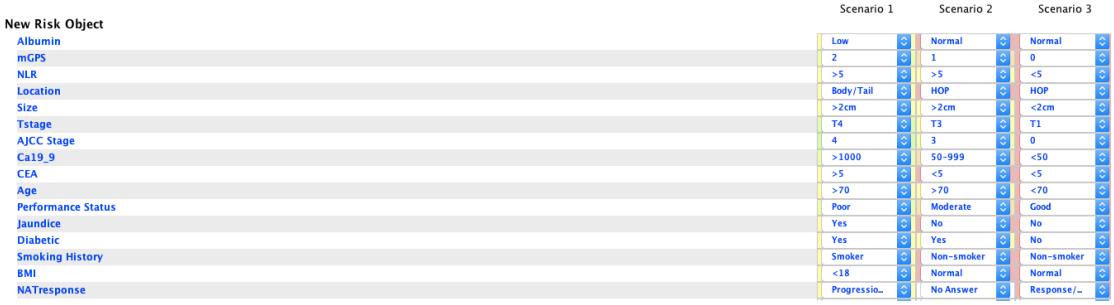
Proposed BBN models presented here have been checked with three hypothetical scenarios for each of the pre and post-operative applications of the BBNs. The states of the criteria for the scenarios are summarised in Figure 63 and Figure 64. Scenario 1 reflects a patient where all the criteria are in the worst possible state. Scenario 2 reflects a patient where all criteria are in medium or moderate states. Scenario 3 reflects a patient where all criteria are in the most favourable state for a good prognostic outcome.

For the BBN that pre-operatively predicts poor prognosis scenario 1 corresponded with a 99.989% probability that the patient will have a poor prognosis post resection. Scenario 2 corresponds with a 56.091% probability of a poor post resection prognosis. Scenario 3 corresponded with a 99.989% probability that the patient will not have a poor prognosis post resection. The BBN that incorporates post-operative data to perform prognostic updating showed scenario 1 corresponds with a 99.984% probability and scenario 2 a 55.089% probability of the patient having a poor post resection prognosis. Scenario 3 corresponded with a 99.984% probability of not having a poor post resection prognosis.

For the BBN predicting good prognosis based on pre-operative data scenario 1 corresponded with a 99.806% probability of not achieving a good prognosis post resection. Scenario 2 and 3 gave a 56.695% and 99.806% probability of achieving a good post resection prognosis respectively. The BBN that incorporated post-operative

data to predict good prognosis found that for scenario 1 the probability of not achieving good prognosis was 99.869%. For scenario 2 and 3 the probability of achieving a good post resection prognosis is 48.959% and 99.869% respectively.

Figure 63: Scenario testing for BBN predicting poor prognosis

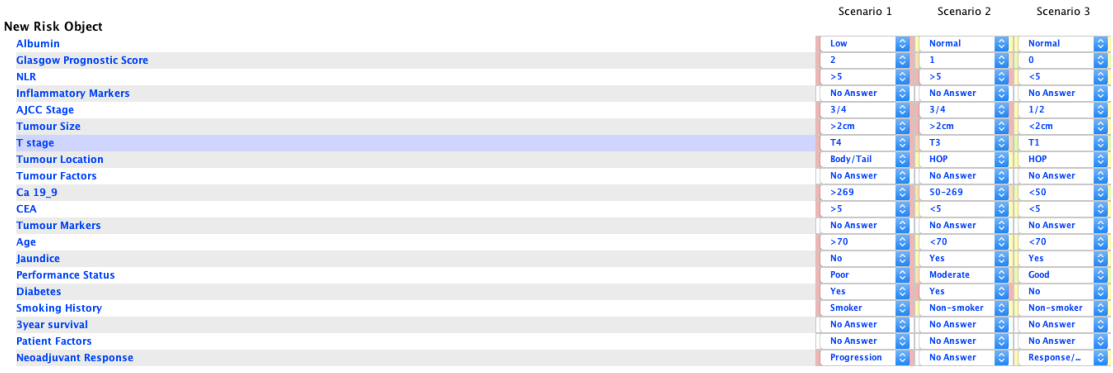


New Risk Object	Scenario 1	Scenario 2	Scenario 3
Albumin	Low	Normal	Normal
mGPS	2	1	0
NLR	>5	>5	<5
Location	Body/Tail	HOP	HOP
Size	>2cm	>2cm	<2cm
Tstage	T4	T3	T1
AJCC Stage	4	3	0
Ca19_9	>1000	50-999	<50
CEA	>5	<5	<5
Age	>70	>70	<70
Performance Status	Poor	Moderate	Good
Jaundice	Yes	No	No
Diabetic	Yes	Yes	No
Smoking History	Smoker	Non-smoker	Non-smoker
BMI	<18	Normal	Normal
NATresponse	Progressio...	No Answer	Response/...

Prognostic Updating

Lymph node positive	Yes	Yes	No
Lymph Node Ratio	>0.3	<0.3	<0.3
Grade	G3/G4	G3/G4	G1/G2
R0 resection	No	Yes	Yes
Vascular Involvement	Yes	Yes	No
Perineural Involvement	Yes	Yes	No
Adjuvant Therapy	No	Yes	Yes
Post treatment Ca19_9	>120	>120	<120
Peri operative Blood Transfusion	Yes	No	No

Figure 64: Scenario testing for BBN predicting good prognosis



New Risk Object	Scenario 1	Scenario 2	Scenario 3
Albumin	Low	Normal	Normal
Glasgow Prognostic Score	2	1	0
NLR	>5	>5	<5
Inflammatory Markers	No Answer	No Answer	No Answer
AJCC Stage	3/4	3/4	1/2
Tumour Size	>2cm	>2cm	<2cm
T stage	T4	T3	T1
Tumour Location	Body/Tail	HOP	HOP
Tumour Factors	No Answer	No Answer	No Answer
Ca 19_9	>269	50-269	<50
CEA	>5	<5	<5
Tumour Markers	No Answer	No Answer	No Answer
Age	>70	<70	<70
Jaundice	No	Yes	Yes
Performance Status	Poor	Moderate	Good
Diabetes	Yes	Yes	No
Smoking History	Smoker	Non-smoker	Non-smoker
3year survival	No Answer	No Answer	No Answer
Patient Factors	No Answer	No Answer	No Answer
Neoadjuvant Response	Progression	No Answer	Response/...

Prognostic Updating

Lymph nodes negative	No	No	Yes
R0 resection	No	No	Yes
Tumour Grade	G3/G4	G3/G4	G1/G2
Lymph Node Ratio	>0.2	<0.2	<0.2
Perineural Involvement	Yes	Yes	No
Vascular Involvement	Yes	Yes	No
Adjuvant Treatment	No	Yes	Yes
Post_op Ca19_9	High	Normal	Normal
Pathology	No Answer	No Answer	No Answer
Follow_up	No Answer	No Answer	No Answer
Post_operative Factors	No Answer	No Answer	No Answer

Sensitivity analysis

Sensitivity analysis assumes all input parameters in the model are uncertain and therefore determines how sensitive results are in relation to changes in observable variables (Yang *et al.*, 2009). Given that BBN output relies on a priori assigned probabilities, sensitivity analysis identifies critical input parameters that significantly impact on BBN results (Ismail *et al*., 2011). Therefore sensitivity analysis can serve as an adjunct to decision-analysis or value of information analysis to identify uncertainties and prioritise data collection (Fenton & Neil, 2019).

Various methods have been proposed to carry out sensitivity analysis. This BBNs used discreet variables, as variables in PubMed published survival analysis studies reported discretised variables, therefore Pearl's inwards analysis and broadcasting analysis were used to perform sensitivity analysis (Pearl, 1988; Fenton & Neil, 2019). Hence sensitivity was defined as $S(\bar{X}=\bar{x}, T=t)$ and determined by setting values for all source variables, $\bar{X}$, and assessing the impact on the target node, T, then changing only the target node, T, and assessing the changes on the source set, $\bar{X}$, respectively with joint sensitivity of T to perturbations in source nodes defined as:

$$S(X=x,T=t) = \frac{p(T=t|e,X=x)}{p(T=t|e)}$$

where $p(T=t|e)$ is the current probability value for T, given evidence e and $p(T=t|e, X=x)$ is the new value of T for the set of source variable, X (Pearl, 1988; Fenton & Neil, 2019). Hence inwards analysis and broadcasting results were equivocal as (Fenton & Neil, 2019):

$$\frac{p(T=t|e,X=x)}{p(T=t|e)} = \frac{p(X=x|T=t,e)}{p(X=x|e)}$$

The results of BBN sensitivity analysis showed that for the pre-operative BBNs tumour factors had the greatest impact on outcomes, followed by patient factors (Appendix Q). When post-operative data was incorporated into the BBN post-operative factors and surgical pathology had greatest impact on output followed by tumour factors and patient factors (Appendix Q). This corroborates numerous previous studies that have established that detecting disease early improves survival, hence the importance of tumour factors (Winter et al., 2012). Furthermore it has been established that the best chance of good survival outcomes depend upon achieving R0 resection whereby all tumour is completely removed (Versteijne *et al.*, 2018), which supports the importance of surgical pathology on impacting on BBN outcome. Receipt of multimodal treatment has also been established as key to achieving best possible survival outcomes in numerous studies (Bradley *et al*., 2018; Bradley *et al.,* 2019a; Versteijne *et al*., 2018; Neoptolemos *et al*., 2001) which supports the findings from sensitivity analysis that post-operative factors have significant impact on BBN output as this includes post-operative complications, which in turn affects the recovery time and time to, or indeed whether, patients receive adjuvant therapy which is also included in post-operative factors (Winter *et al*., 2012). Interestingly patient factors were found to have more of an impact on predicting poor prognosis than good prognosis. This could be explained by the fact that in practice clinicians are less likely to operate on frail and unfit patients therefore the data for such patients is less likely to form part of post resection survival analysis. This finding makes a

case for better patient selection at the pre-operative stage of the patient journey. Importantly through the novel application of methods in this paper we were able to quantify the cumulative impact of patient factors, which had previously been the domain of subjective judgment.

4.6.3 BBN Performance Validation

Strict methodological rigor was adhered to when creating this BBN by strictly following TRIPOD guidelines as outlined in appendix R. The West of Scotland Pancreatic Unit has been selected to demonstrate the application of the BBN models. This tertiary referral centre has a 20year prospectively maintained database that contains all patients referred to the unit with potentially resectable PDAC. Between January 2008 and July 2012 this unit performed surgery first pathway on all cases of potentially resectable PDAC if patients were deemed fit for surgery. From August 2012 onwards neoadjuvant approach became the standard-of-care for all cases of potentially resectable PDAC. This database therefore allowed the opportunity to demonstrate the application of BBNs across both neoadjuvant and surgery first treatment pathways.

Model Validation Dataset

The performance of BBN was assessed using the area under the curve (AUC) of the received operated curve (ROC) using SPSS Statistics version 24 software. Individual patient data was entered

into the BBNs and the personalised pre and post-operative predictions of poor prognosis and good prognosis were recorded. A prediction of 50% or greater probability of the prognostic output in question was assessed against that individual's actual post resection survival time and the BBN prediction deemed 'True' or 'False'. All patients who had undergone resection of PDAC, had survival data recorded, had died, or if still alive had a survival time below or exceeding that being predicted were included. Patients who were found to have non-resectable disease at operation, or who were treated in a neoadjuvant pathway and were found to have non-resectable disease at re-staging, were included to reduce the risk of bias in assessing the pre-operative performance of BBN as in the clinical setting the intention would be delivery of multimodal treatment. This gave a pool of 387 patients against which the predictive performance of the pre-operative BBN was validated for predicting poor prognosis and a pool of 365 against which the performance of the pre-operative BBN predicting good prognosis was validated. The predictive performance of the post-operative BBNs in performing prognostic updating was assessed against all patients for whom postoperative data was available and who had survival data recorded, had died, or if still alive had a survival time exceeding that being predicted. This gave a pool of 251 and 230 patients against which the BBNs for poor and good prognosis were validated respectively. The amount of missing data did not determine either inclusion or exclusion in order to test how the model coped with missing data. This database did not contain data on tumour markers Ca19-9 and CEA.

Results: Pre-operative Performance

BBN predicting poor prognosis pre-operatively achieved an AUC of 0.70 (*P* value 0.001; 95% CI 0.589-0.801) where data on all other nodes, apart from tumour markers, were available. With an additional one and two data points missing a statistically significant AUC of 0.70 was maintained. When an additional three data points were missing the AUC remained above 0.60 but lost statistical significance (Table 44).

Table 44: Results of BBN pre-operatively predicting poor prognosis. PPV is positive predictive value; NPV is negative predictive value

	Sensitivity	Specificity	PPV	NPV	AUC
2 data points missing (n=123)	0.84	0.64	0.45	0.92	0. 70 (*P* value 0.001; 95% CI 0.589-0.801) Std. Error: 0.54
3 data point missing (n=139)	0.82	0.65	0.43	0.92	0. 70 (*P* value 0.001; 95% CI 0.578 -0.786) Std. Error: 0.53
4 data points missing (n=144)	0.83	0.65	0.44	0.92	0.70 (*P* value 0.001; 95% CI 0.591 -0.791) Std. Error: 0.51
5 data points missing (n=176)	0.64	0.66	0.45	0.81	0. 65 (*P* value 0.009; 95% CI 0.537 -0.711) Std. Error: 0.44
6 data points missing (n=189)	0.66	0.63	0.43	0.82	0.64 (*P* value 0.024; 95% CI 0.518 -0.690) Std. Error: 0.44
6+ data points missing (n=387)	0.64	0.54	0.46	0.72	0.60 (*P* value 0.559; 95% CI 0.502-0.617) Std. Error: 0.29

BBN pre-operative predicting good prognosis achieved an AUC that ranged from 0.94 (P-value 0.002; 95% CI 0.859-1.000) for 0 missing data points in addition to the missing tumour marker data (n=33) to AUC 0.74 (P-value 0.000; 95% CI 0.660-0.809) accepting more than 4 additional missing data points (n=365) (Table 45).

Table 45: Results of BBN pre-operatively predicting good prognosis. PPV is positive predictive value; NPV is negative predictive value

	Sensitivity	Specificity	PPV	NPV	AUC
2 missing data points (n=32)	1	0.93	0.71	1	0.94 (*P* value 0.002; 95% CI 0.856-1.0) Std. Error 0.043
2-3 data points missing (n=119)	0.85	0.90	0.71	0.95	0.868 (*P* value 0.000; 95% CI 0.780-0.956) Std. Error 0.045
2-4data points missing (n=132)	0.83	0.92	0.75	0.95	0.873 (*P* value 0.000; 95% CI 0.793-0.953) Std. Error 0.041
2-5 data points missing (n=135)	0.83	0.92	0.73	0.95	0.871 (*P* value 0.000; 95% CI 0.790-0.951) Std. Error 0.041
2-6 data points missing (n=175)	0.83	0.89	0.61	0.96	0.831 (*P* value 0.000; 95% CI 0.746-0.917) Std. Error 0.044
6+ data points missing (n=365)	0.62	0.86	0.42	0.93	0.735 (*P* value 0.000; 95% CI 0.660-0.809) Std. Error 0.038

Prognostic Updating Performance

BBN performance in prognostic updating where the outcome predicted was poor prognosis achieved an AUC of 0.80 (*P value*: 0.000; 95% CI: 0.678-0.862) when all other data, apart from tumour markers, was available. An AUC of 0.80 was maintained until more than 6 pre-operative data points, and up to and including 2 post-operative data points, were missing. At this point the BBN achieved an AUC of 0.70 (*P value*: 0.000; 95% CI:0.667-0.818) which was maintained with over 6 missing pre-operative data points and up to and including 4 missing post-operative data points (*P value*: 0.000; 95% CI: 0.660-0.788) (Table 46).

Table 46: Results of BBN performing prognostic updating for poor prognosis.

	1 Missing Post-operative Data point	1-2 Missing Post-operative Data Point	1-3 Missing Post-operative Data Points	1-4 Missing Post-operative Data Points
2 Missing Pre-operative Data Points	Sensitivity: 0.97; Specificity: 0.62; PPV: 0.44; NPV: 0.98; AUC 0.80; Standard Error:0.47; *P value*: 0.000; 95% CI: 0.678-0.862 (n=117)	Sensitivity: 0.93; Specificity: 0.63; PPV: 0.44; NPV:0.97; AUC: 0.80; Standard Error:0.51; *P value*: 0.000; 95% CI: 0.651-0.850 (n=120)		
2-3 Missing Pre-operative Data Point	Sensitivity: 0.94; Specificity: 0.63; PPV: 0.45; NPV: 0.97; AUC: 0.80; Standard Error:0.045; *P value*: 0.000; 95% CI: 0.685-0.862 (n=138)	Sensitivity: 0.94; Specificity: 0.63; PPV: 0.45; NPV: 0.97; AUC: 0.80; Standard Error:0.045; *P value*: 0.000; 95% CI: 0.685-0.862 (n=139)		
2-4 Missing Pre-operative Data Points	Sensitivity: 0.94; Specificity: 0.62; PPV: 0.45; NPV: 0.98; AUC: 0.80; Standard Error: 0.042; *P value*: 0.000; 95% CI: 0.708-0.872 (n=135)	Sensitivity: 0.94; Specificity: 0.62; PPV: 0.44; NPV: 0.97; AUC: 0.80; Standard Error: 0.045; *P value*: 0.000; 95% CI: 0.681-0.858 (n=140)		
2-5 Missing Pre-operative Data Points	Sensitivity: 0.97; Specificity: 0.62; PPV: 0.45; NPV: 0.98; AUC: 0.80; Standard Error: 0.041; *P value*:	Sensitivity: 0.95; Specificity: 0.61; PPV: 0.45; NPV: 0.97; AUC: 0.80; Standard Error: 0.043; *P value*: 0.000;		

	0.000; 95% CI: 0.708-0.869 (n=137)	95% CI: 0.681-0.849 (n=146)		
2-6Missing Pre-operative Data Points	Sensitivity: 0.97; Specificity: 0.61; PPV: 0.45; NPV: 0.98; AUC: 0.80; Standard Error:0.041; *P value*: 0.000; 95% CI: 0.707-0.869 (n=138)	Sensitivity: 0.95; Specificity: 0.59; PPV: 0.43; NPV: 0.97; AUC: 0.80; Standard Error:0.043; *P value*: 0.000; 95% CI: 0.665-0.832 (n=155)	Sensitivity: 0.95; Specificity: 0.59; PPV: 0.44; NPV: 0.97; AUC: 0.80; Standard Error:0.042; *P value*: 0.000; 95% CI: 0.672-0.835 (n=157)	
>6 Missing Pre-operative Data Points	Sensitivity: 0.97; Specificity: 0.61; PPV: 0.45; NPV: 0.98; AUC: 0.80; Standard Error:0.041; *P value*: 0.000; 95% CI: 0.710-0.870 (n=139)	Sensitivity: 0.94; Specificity: 0.55; PPV: 0.41; NPV: 0.96; AUC: 0.70; Standard Error:0.039; *P value*: 0.000; 95% CI: 0.667-0.818 (n=195)	Sensitivity: 0.94; Specificity: 0.54; PPV: 0.41; NPV: 0.96; AUC: 0.70; Standard Error: 0.037; *P value*: 0.000; 95% CI: 0.667-0.814 (n=205)	Sensitivity: 0.96; Specificity: 0.49; PPV: 0.44; NPV: 0.97; AUC: 0.70; Standard Error:0.033; *P value*: 0.000; 95% CI: 0.660-0.788 (n=251)

The prognostic updating performance of BBN predicting good prognosis achieved AUC 0.97 (P-value 0.000; 95% CI 0.908-1.000) for 3 missing data points in pre and post-operative validation dataset (n=33) to AUC 0.75 (P-value 0.000; 95% CI 0.655-0.838) accepting more than 6 missing data points in the pre and up to and including 3 missing data points in the post-operative validation dataset (n=230). The latter was the only point at which BBN performance had an AUC under 0.80. Validated against every other combination of missing pre and post-operative data points BBN maintained an AUC greater than 0.8 (range 0.97-0.80) with P-value consistently below 0.001 (Table 47).

Table 47: Results of BBN performing prognostic updating for good prognosis.

	1 Missing Post-operative Data points	1-2 Missing Post-operative Data Point	1-3 Missing Post-operative Data Points	1-4 Missing Post-operative Data Points	1-5 Missing Post-operative Data Points
2 Missing Pre-operative Data Points	Sensitivity: 1 Specificity: 0.96 PPV: 0.86 NPV: 1 AUC 0.97; P-value 0.000 (95% CI 0.908-1.000) (n=33)				
2-3 Missing Pre-operative Data Point	Sensitivity: 0.62 Specificity: 0.96 PPV: 0.64 NPV: 0.89 AUC 0.81; P-value 0.000 (95% CI 0.705-0.910) (n=113)			Sensitivity: 0.62 Specificity: 0.96 PPV: 0.64 NPV: 0.89 AUC 0.81; P-value 0.000 (95% CI 0.706-0.911) (n=114)	
2-4 Missing Pre-operative Data Points	Sensitivity: 0.62 Specificity: 0.90 PPV: 0.67 NPV: 0.88 AUC 0.80; P-value 0.000 (95% CI 0.699-0.900) (n=121)	Sensitivity: 0.84 Specificity: 0.63 PPV: 0.90 NPV: 0.68 AUC 0.0.80; P-value 0.000 (95% CI 0.702-0.901) (n=122)	Sensitivity: 0.62 Specificity: 0.91 PPV: 0.67 NPV: 0.89 AUC 0.80; P-value 0.000 (95% CI 0.701-0.901) (n=126)		Sensitivity: 0.62 Specificity: 0.91 PPV: 0.67 NPV: 0.89 AUC 0.80; P-value 0.000 (95% CI 0.704-0.902) (n=127)
2-5 Missing Pre-operative Data Points	Sensitivity: 0.62 Specificity: 0.91 PPV: 0.67 NPV: 0.89 AUC 0.81; P-value 0.000 (95% CI 0.708-0.904) (n=130)			Sensitivity: 0.62 Specificity: 0.90 PPV: 0.64 NPV: 0.83 AUC 0.81; P-value 0.000 (95% CI 0.705-0.902) (n=131)	
2-6 Missing	Sensitivity: 0.62	Sensitivity: 0.62 Specificity: 0.91			Sensitivity: 0.62

Pre-operative Data Points	Specificity: 0.91 PPV: 0.67 NPV: 0.89 AUC 0.81; P-value 0.000 (95% CI 0.711-0.905) (n=131)	PPV: 0.64 NPV: 0.90 AUC 0.81; P-value 0.000 (95% CI 0.709-0.905) (n=138)			Specificity: 0.91 PPV: 0.64 NPV: 0.90 AUC 0.80; P-value 0.000 (95% CI 0.705-0.903) (n=140)
>6 Missing Pre-operative Data Points	Sensitivity: 0.63 Specificity: 0.92 PPV: 0.70 NPV: 0.90 AUC 0.82; P-value 0.000 (95% CI 0.721-0.909) (n=136)	Sensitivity: 0.66 Specificity: 0.92 PPV: 0.68 NPV: 0.91 AUC 0.82; P-value 0.000 (95% CI 0.736-0.910) (n=172)	Sensitivity: 0.54 Specificity: 0.93 PPV: 0.68 NPV: 0.87 AUC 0.75; P-value 0.000 (95% CI 0.655-0.838) (n=230)		

4.6.4. Discussion

In this section BBN models to predict poor and good post-resection prognosis for PDAC have been proposed. The BBNs demonstrated applicability and high performance level, which compares favorably with previous prognostic models (Appendix P), even when data is missing, enhancing its clinical applicability. The BBNs developed here not only view the process of delivering pancreatic cancer management as a complex system, but also engage with the FUPS characteristics of the available data to make personalised prognostic predictions.

Engaging With FUPS Characteristics

The framework proposed by Wolpert & Rutter (2018) for dealing with FUPS data has been utilised with the development of these

BBNs. Firstly the predictions made by the models are presented as a partial remnants with the intention that as patient databases mature globally, the anticipated next step will be to incorporate patient level data into the BBN so that, through Bayes theorem the prior distribution and observed data are combined to update the posterior distribution and further improve the accuracy of predictions (Pearl, 1988; Fenton & Neil, 2019). This proposed method for creating a BBN is flexible to include more contributing factors and consequence factors as new information begins to emerge such as a better understanding of the role of neoadjuvant therapies. The current focus on precision medicine carries the potential for gene-targeted therapies (Tonelli & Shirts, 2017; MacConaill *et al*., 2015). Future clinical decision making will therefore require the amalgamation of clinical and genomic data, which could easily be incorporated into the BBN, making our model a vehicle for delivering precision gene-targeted medicine (Tonelli & Shirts, 2017; MacConaill *et al*., 2015; Dzau & Ginsburg, 2016).

These potential benefits and future impact of the BBNs also link to the two other key factors in the framework for engaging with FUPS data: transparency and triangulation. The two-stage weighting process synthesised published survival analysis studies to identify major predictive and consequence factors that represent nodes within a weighted ranked BBN. The transparent methods of weighting the nodes within the network also triangulates the importance of variables with conflicting findings and their significance within the entirety of the existing body of research available on the PubMed database. This approach provides a method

for utilising population based survival analysis data to make personalised predictions of outcomes across competing treatment strategies, importantly at the pre-operative stage, therefore supporting shared patient-clinician decision making at individual patient level and planning resource allocation. The intuitive nature of BBNs and the transparency of the methods used builds trust with clinicians. The methods presented in this section also overcome the limitations of previous models applied in this field including the over reliance on post-operative data from small, biased databases limiting generalisability.

Chapter 5

Discussion

Introduction

The aim of this research was to facilitate the fruition of personalised realistic medicine for cases of potentially resectable pancreatic cancer through statistical modeling that can facilitate better shared decision making to optimise individual patient outcomes. From early on in the literature review what began to emerge was a rich and complicated tapestry of the current body of research pertaining to the management of potentially resectable pancreatic cancer. This body of research is permeated with uncertainty, ambiguity, and often conflicting and heavily contested findings and opinions across the entirety of the treatment pathway including: 1) the staging of pancreatic cancer and its implications on treatment goals and predicted outcomes, 2) the application of surgery, 3) adjuvant therapies and 4) neoadjuvant therapies.

These issues exist within the wider political context of a drive towards the delivery of personalised realistic medicine through more personalised treatment selection strategies that will ensure more cost-effective resource utilisation. This has resulted in the wider contemporary research focus within pancreatic cancer research being driven in two key directions: 1) the drive for more large multi-centre RCTs comparing neoadjuvant and upfront surgery and 2)

precision medicine with the focus on biomarker driven early diagnosis and treatment sequencing and gene targeted therapies.

The widely held assumption is that breakthroughs in these areas will result in a move away from ambiguity towards certainty. This thesis, whilst acknowledging the importance of such breakthroughs, contests this view. Firstly the challenge of optimising outcomes for potentially resectable pancreatic cancer goes far beyond simply choosing between neoadjuvant *versus* upfront surgery approach. Preliminary findings from Prep-02/JSAP-05 trial, the first RCT comparing upfront surgery and neoadjuvant therapy in the form of gemcitabine and S1 for resectable pancreatic cancer, has reported improved overall survival outcomes with neoadjuvant therapy (Unno *et al*., 2019). However, another RCT comparing mFOLFIRINOX with gemcitabine in the adjuvant setting within the upfront surgery pathway has reported improved survival outcomes with mFOLFIRINOX that rivals the survival outcomes reported in the neoadjuvant arms of the Prep-02/JSAP-05 trial (Conroy *et al*., 2018). Therefore the superior treatment pathway for resectable pancreatic caner has not been conclusively established and remains controversial. Furthermore superior treatment regime combinations within competing pathways have also not been conclusively established. Secondly with regard to gene targeted therapy and the focus on precision medicine, as our knowledge of disease at biomolecular and genomic level evolves, the clinical decision making process will pullulate with varied and complex datasets from multiple sources. Put simply humans are, and always will be, more than their genomes. This current direction in pancreatic cancer

research, rather than resulting in the diminution of complexity, could result in its augmentation considering the challenge of amalgamating such large complex databases and the meaningful application of this information to the individual patient to optimise individual treatment outcomes.

The argument being put forward in this thesis is that complexity will not be 'solved' but, if the optimisation of outcomes for pancreatic cancer is to come to fruition, research must focus on developing ways to engage with the complexity, handle uncertainty and the emergent when examining the complex system of delivering pancreatic cancer care including areas of debate, ambiguity and disagreement (Law & Mol, 2002; Fraser & Greenhalgh, 2001; Star, 2002; Greenhalgh & Papoutsi, 2018). It follows that because the system of delivering pancreatic cancer care and its outcomes are dynamic, the traditional scientific quest for certainty, predictability and linear causality through a focus on RCTs and precision medicine will only answer a fraction of the unanswered questions as the effect of context is controlled for within the artificial setting of such trials (Cohn *et al*., 2013; Braithwaite *et al*., 2017; Marchal *et al*., 2013; Greenhalgh & Papoutsi, 2018). RCTs with their strict inclusion criteria and control of context do not reflect the complexities of a real-life patient case mix and therefore cannot alone provide solutions to the challenge of optimising outcomes on an individual patient level. Therefore a key aim of this research was to augment such studies by exploring the application of statistical modeling methods that deal with uncertainty, unpredictability and general causality through methods that foreground dynamic interactions and emergence to understand

how complex systems come together as a whole (Cohn *et al*., 2013; Greenhalgh & Papoutsi, 2018; Flyvbjerg, 2006).

Existing modeling techniques regard prognosis as an isolated event at a pre-determined time, applying attribute selection prior to inducing the model and setting fixed roles of input and output variables to attributes (Verduijn *et al.,* 2007). They neglect the uncertain and dynamic nature of care processes where outcomes today predict those of tomorrow hence expected patient outcomes evolve as more information becomes available (Verduijn *et al.,* 2007). Put simply, traditional decision support models integrate data and knowledge but do not provide reasoning (Muthurmama & Sankaran, 2014). To achieve personalised predictive medicine statistical models therefore must improve both knowledge representation and reasoning facility, with ontologies employed acting as stepping-stones to achieving this, and ultimately delivering personalised realistic medicine (Muthurmama & Sankaran, 2014).

It is not without coincidence that the field of complex systems developed at a time when statistical theory began to coalesce with methods encompassed within machine learning to reliably infer models with large numbers of variables that interact in complex, non-linear ways. At their core these methods make predictions within complex systems against a background of competing risks and events (Abbod *et al.*, 2014). However, across the existing body of research pertaining to the management of potentially resectable pancreatic cancer the definition of the research problem, proposals for improvement and outcome are not recognised as being

dependent on the whole system (Ulrich, 2002). Existing studies utilising machine learning methods are not exempt from this criticism (Bradley *et al.,* 2019b; Bradley *et al*., 2019c) which means that the potential of their application is mainly untapped as doing so would place greater emphasis on how systems boundaries are justified and the implications this has for what modeling a system defined in such a way will, and importantly will not, reveal. Hence the limitations of the existing body of research is not merely to be seen as a series of methodological issues to be corrected, but rather as the system that is the delivery of healthcare being defined in simplistic and reductionist terms which defines how system boundaries are set which in turn determines and limits how outcomes are measured and assessed. It follows that both a limited knowledge of systems as a result of boundaries and a failure to engage with complexity exist and therefore require a critical and ethical imperative in the study and understanding of such systems in order to move research forward (Kruger *et al*., 2019).

To achieve this, this research engaged with what Tsoukas (2017) called conjunctive theorising by avoiding simplification and abstraction (or disjunctive theorising) and instead draws on different kinds of data from multiple sources to move research towards a theory that can build a rich picture of pancreatic cancer management pathways as a complex system. Combining operational and healthcare research and drawing on influences from complementary paradigms of critical realism and systems theory then enhancing their impact by using Cilliers' complexity theory 'lean ontology', an open-world ontology was held (Cilliers, 1998; Kruger *et al*., 2019).

Specifically the framework offered by Wolpert & Rutter (2018) for using FUPS data was expanded within the context of simulation modeling in the study of complexity in healthcare to attempt to expand its capabilities for handling the emergent and uncertainty (Long *et al*., 2018).

Several themes began to emerge in the findings of this research and frame the structure of the rest of this chapter. Firstly the optimisation of outcomes for potentially resectable pancreatic cancer can be seen as the benefit of marginal gains at individual patient level rather than proving conclusive superiority of any single treatment pathway. This challenges the current narrative surrounding the management of pancreatic cancer to move away from the theoretically possible (surgery is the only potentially curative treatment) to engage with the reality for many patients that surgery may be of limited benefit. Secondly the importance of the emerging narrative being uncovered within the 'unseen data' that has previously been dismissed due its FUPS characteristics reveals key findings that add a new dimension to the ongoing debate regarding the treatment of potentially resectable pancreatic cancer and moves research closer to realising personalised realistic medicine. Thirdly the application of Cillier's complexity theory 'lean ontology' reveals new insights into the management of potentially resectable pancreatic cancer by engaging with treatment pathways as complex systems.

5.1 Optimising Outcomes in Pancreatic Cancer: the aggregate of marginal gains

Outcomes for pancreatic cancer are poor with overall 10year survival standing at less than 1% and 5year survival for resected cases between only 7% and 25% (Cancer Research UK, 2019). Despite advances in adjuvant therapies and advances in surgical techniques overall survival outcomes have improved very little over several decades. Furthermore although the survival benefit of adjuvant therapy has been established through randomised RCTs between 71% and 76% of patients will have disease recurrence within 2-years of surgical resection (McGuigan *et al*., 2018).

In this research the Bayesian network meta-analysis that included all stages of potentially resectable disease found the aggregate rate (AR) of 1,2,3,4 and 5-year survival to be marginally higher with neoadjuvant therapy (1-year survival: 0.8109 *versus* 0.6403, O.R: 2.12, 95% CI: 1.59-2.93; 2-year survival: 0.5135 *versus* 0.3002, O.R: 1.65 95%, CI: 1.16-2.34; 3-year survival: 0.3151 *versus* 0.2147, O.R: 1.50, 95% CI: 1.10-2.04; 4-year survival: 0.2114 *versus* 0.1647 O.R: 1.57, 95% CI: 0.80-2.99; 5-year survival: 0.2118 *versus* 0.1736, O.R: 1.65, 95% CI: 0.68-3.73). This marginal survival benefit was also reflected in the AR of 1, 2, and 5year survival for cases that were resectable at presentation (1-year survival: 0.7969 *versus* 0.7481, O.R: 1.38, 95% CI: 0.69-2.96; 2-year survival: 0.5178 *versus* 0.5131, O.R: 1.26, 95% CI 0.94-1.74; 5-year 0.2069 *versus* 0.1783, O.R: 1.19 95% CI 0.65-1.73). These findings corroborate the few existing attempts at meta-analysis comparing neoadjuvant and upfront

surgery approaches. Meta-analysis by both Xu *et al*. (2014) and Andriulli *et al*. (2012) reported only marginal benefit of neoadjuvant chemotherapy in terms of overall and disease-free survival in resectable cases. However, neither of these reports focused solely on neoadjuvant therapy therefore omitted significant studies from their meta-analysis (Lee *et al.* 2016). More recently meta-analysis by Versteijine *et al*. (2018) pooled 38 trials comprising 3484 patients with resectable and borderline resectable disease in an intention-to-treat analysis. Their findings also reported only marginal benefit with neoadjuvant approach over upfront surgery approach for resectable disease (18.2months *versus* 17.4months) but a greater survival advantage with borderline resectable disease (19.2months *versus* 12.8months). Mokdad *et al*. (2017) also reported a survival advantage with neoadjuvant approach in their retrospective analysis of National Cancer Database using propensity score matched analysis of neoadjuvant therapy used to treat 2005 patients with stage I and II PDAC compared to 6,015 patients who underwent upfront surgical resection of PDAC (26months *versus* 21months). However, this analysis is heavily biased as only those who tolerated neoadjuvant therapy and underwent resection were included in the neoadjuvant group.

The tendency towards a marginal survival benefit with neoadjuvant approach continued to be demonstrated in the results of the Markov decision-analysis studies. Populating the Markov decision analysis model with data on resectable disease at the time of presentation, from both synthesisation of results from published trials and institutional data, demonstrated a survival advantage of 2.69months

(26.41 months/ 22.54 QALMs *versus* 23.72 months/18.51 QALMs) and 8.22months (32.90 months/28.51 QALMs *versus* 24.68 months/19.23 QALMs) respectively for neoadjuvant therapy pathway. These findings were reiterated in the results of the DES analysis which showed that for resectable only cases the neoadjuvant pathway gave a mean survival time of 20.01months (18.45 QLAMs) compared to 16.55months (14.19 QALMs) in the upfront surgery pathway. However, when minimum significant difference threshold was set to 3.5months the selection frequency was 40.6% for neoadjuvant pathway and 59.4% for indifference between pathways. DES analysis has never before been applied to the question of neoadjuvant pathway *versus* upfront surgery pathway for pancreatic cancer but two previous Markov decision-analysis studies that have focused on resectable only cases and utilised synthesised data have similarly reported marginal survival benefit of neoadjuvant pathway (deGeus *et al*., 2016; Sharma *et al*., 2015). One study, based on phase I/II trials, reported a 2month survival advantage (22 months *versus* 20months) (Sharma *et al*., 2015) whilst the second study reported a 5.5month advantage (32.2 *versus* 26.7 months) but their analysis mostly included retrospective studies from a single search engine (deGeus *et al*., 2016). These, and the findings of this research are corroborated by the preliminary findings from Prep-02/JSAP-05, the first RCT comparing upfront surgery and neoadjuvant therapy for resectable pancreatic cancer. However this trial reports an overall survival advantage of 10.01months (36.72months *versus* 26.65months).

Based on synthesised data from published trials that included all stages of potentially resectable pancreatic cancer upfront surgery was found to give 23.72 months (18.51 QALMs) *versus* 20.22 months (16.26 QALMs). One existing Markov decision analysis focuses on potentially resectable pancreatic cancer, therefore including borderline and locally advanced cases in the neoadjuvant pathway to capture the effect of conversion to resectability on overall pathway analysis (VanHouten *et al*., 2012). As with this study's findings it did not demonstrate an overall conclusively superior pathway on an intention-to-treat basis (neoadjuvant pathway 18.6 months *versus* 17.1 months) (VanHouten *et al*., 2012). Yet preliminary results from the PREOPANC-1 trial, a multicenter phase III RCT comparing neoadjuvant therapy and upfront surgery for borderline resectable cases, have reported improved survival with neoadjuvant therapy on an intention-to-treat basis (17.1 months *versus* 13.5 months) (Van Tienhoven *et al.,* 2018).

At this level of analysis the results of this research appear to corroborate the overall narrative emerging from the existing body of research: that neither pathway is conclusively superior but there may be a marginal advantage with neoadjuvant approach. However, this conclusion is contested considering the issues of quality surrounding existing studies. Therefore although the potential benefits appear to be marginal what is not made clear by simply looking at survival outcomes alone is where marginal gains can be made in either pathway and what the aggregate effect on overall outcome might be.

This point is illustrated best when considering the importance of achieving R0 resection, which has been long thought of as the only potentially curable treatment for pancreatic cancer. The Bayesian Network meta-analysis conducted in this research showed that the aggregate rate of R0 resection was marginally higher in the neoadjuvant pathway for cases that were resectable at presentation (0.8008 *versus* 0.7515, O.R. 1.27, 95% CI 0.60-1.96) with this margin reducing when all stages of potentially resectable pancreatic caner were included in the neoadjuvant arm (0.7389 *versus* 0.7306, O.R 1.12, 95% CI 0.60-2.08). The corresponding survival benefits were also marginal as previously discussed. This corroborates the findings of a meta-analysis of pooled proportions conducted by Versteijine *et al*. (2018). Here R0 resection rates were also higher in the neoadjuvant compared to upfront surgery group for both resectable (85% *versus* 71.4%) and borderline resectable (88.6% *versus* 63.9%) cases. However, as with the Bayesian network meta-analysis performed in this research the corresponding survival benefits were relatively small at 0.8months and 6.4months for resectable and borderline resectable cases respectively. Furthermore preliminary results from the PREOPANC-1 trial (Van Tienhoven *et al*., 2018) reported R0 resection rates of 31% in the surgery first pathway and 63% in neoadjuvant pathway but overall and disease-free survival difference was only 3.6months and 3.3months respectively in favour of neoadjuvant pathway. However the Prep-02/JSAP-05 trial reported no statistically significant difference in R0 resection rates but a greater survival advantage of 10.1months.

Whilst the impact of variations in definitions of R0 resection across studies must be acknowledged (Versteijine *et al*., 2018), the possibility of surgical resection, even achieving R0 resection, rather than being curative is only one aspect within both treatment pathways where marginal gains may be made. It follows that even where decision analysis studies suggested an overall marginal benefit with neoadjuvant therapy the possibility remains that for a subset of patients with early resectable disease with the highest chance of R0 resection and receiving multimodal therapy, upfront surgery may be the superior treatment pathway for them. Furthermore this also raises the possibility that for some patients surgical resection may be of limited benefit. This in turn questions how we define 'success' in the treatment of potentially resectable pancreatic cancer and begs the question how we define the degree of success. Is this based on quantity or quality of survival time? Is achieving R0 resection in itself a success or is it the delivery of multimodal treatment that determines success? Therefore is not proceeding to surgery in the neoadjuvant pathway a failure or does this represent the successful filtering of patients away from futile surgery? To truly make inroads to achieving personalised realistic medicine these are just some of the questions that this research attempted to answer.

To date research has focused on neoadjuvant *versus* upfront surgery pathway to attempt to establish a definitively superior treatment pathway in terms of survival outcomes but to limited avail. Instead this thesis contests that to date we have been asking the wrong question in the wrong way whilst looking at the wrong data (or at

least ignoring other useful data). Using marginal gains at individual patient level to maximise outcomes and move pancreatic cancer research towards personalised medicine allows parallels to be drawn from the work of the mathematician Abraham Wald who changed the course of history in World War II by simply asking different questions in different ways using probability theory to look at different, and often ignored data (Wald, 1980; Mangel & Samaniego, 1984).

5.2 Wald's Lessons: the importance of the unseen data and making better use of FUPS data

During World War II bomber aircrafts flying over Europe sustained enemy fire from both land and air, resulting in high rates of pilot mortality. In response the air force collected vast amounts of data and employed mathematicians, including Wald, to work on establishing a pattern of where bullet holes lay on returning aircrafts. Consequently they began to reinforce aircrafts in these areas to increase the probability of bomber aircrafts returning safely from missions.

Parallels can be drawn with the results in this research. Markov cohort analysis established that, based on synthesised data form trials including all potentially resectable stages of disease, where all treatment modalities were received, neoadjuvant therapy gave 35.05 months (29.87 QALMs) *versus* 30.96 months (24.86QALMs) for R0 resection and 34.08 months (29.87 QALMs) *versus* 25.85months (20.72 QALMs) for R1 resection. For only cases of pancreatic cancer

that were resectable at presentation this pattern continued with Markov cohort analysis based on synthesised data showing that for patients that received all treatment modalities neoadjuvant pathway yielded 39.34 months (34.63 QALMs) compared to 30.96 months (24.86 QALMs) for R0 resection and 34.94 months (31.07 QALMs) compared to 25.85 months (20.72 QALMs) for R1 resection. These findings were also corroborated in the DES analysis which showed that patients with resectable pancreatic cancer who underwent resection had a survival time of 25.30months (22.52QALMs) in the neoadjuvant pathway *versus* 18.07months (13.94QALMs) in the upfront surgery pathway, and that those patients in the neoadjuvant pathway who presented with borderline resectable disease but underwent resection had comparable survival to those presenting with resectable disease treated within the neoadjuvant pathway (25.31months/22.62QALMs). The Markov cohort analysis performed using institutional data of resectable only cases, despite reporting an overall marginal advantage with neoadjuvant pathway, showed that in patients who received all treatment modalities neoadjuvant pathway yielded 45.36 months (40.86 QALMs) compared to 52.59 months (42.38 QALMs) for R0 resection and 42.29 months (30.38 QALMs) compared to 33.37 months (26.81 QALMs) for R1 resection. DES analysis of this same study population showed that in patients who received all treatment modalities neoadjuvant pathway yielded 37.40 months compared to 52.05 months for R0 resection and 27.86 months compared to 20.75months for R1 resection.

Across these analysis based on synthesised data the reported survival times for resection and adjuvant therapy within the upfront surgery pathway ranged from 18.07months to 30.96months. This range is very close to that reported in existing RCTs where, until the JASPAC-4 trial (Uesaka *et al*., 2016), reported survival time had not exceeded 30months (Table 5; Table 6). More recently the PRODIGE 24/CCTG trial (Conroy et al., 2018) reported survival time of 54.5months with adjuvant mFOLFIRINOX. The synthesised data used within the afore mentioned Markov and DES analysis would have largely pre-dated this trial. However the Markov and DES analysis based on institutional data reported similar survival times for R0 resection and adjuvant therapy (52.59months and 52.05months respectively) within the upfront surgery arm as the PRODIGE 24/CCTG trial (54.5months) (Conroy *et al*., 2018).

The range of reported survival times across Markov cohort and DES analysis in this research was 25.30months to 45.36months. Again this is similar to the range reported across existing prospective phase II trials (Appendix E table Ei). More recently the Prep-02/JSAP-05 trial (Unno *et al*., 2019) reported an overall survival time within the neoadjuvant arm of 36.72months which is similar to that reported with the Markov analysis in this research. Furthermore the PREOPANC-1 trial reported a survival time specifically for resected cases treated within the neoadjuvant pathway of 42.2months (Van Tienhoven *et al*., 2018) which is remarkably similar to that reported in these analysis, particularly that using institutional data in a Markov cohort analysis where survival time for multimodal

treatment within the neoadjuvant pathway was reported at 45.36months and 42.29months for R0 and R1 resection respectively.

As with the analysis of the returning bomber planes a pattern has been further established with these research findings: multimodal treatment in either treatment pathway (neoadjuvant therapy and resection or resection and adjuvant therapy) optimises survival outcomes with R0 resection having a benefit over R1 resection. Just as the air force reacted to this emerging pattern by reinforcing the parts of the plane where the bullet holes were found, so too has much of research in pancreatic cancer focused in improving the efficacy of neoadjuvant and adjuvant therapies. In both cases the results have been limited. Although impressive findings from the PRODIGE 24/CCTG trial (Conroy *et al*., 2018) have altered guidelines and will lead to improved survival for those who undergo resection and receive adjuvant therapy, this will not resolve the fact that up to 50% of patients fail to progress to adjuvant therapy (Winter *et al*., 2012). Furthermore, the results of the Markov and DES analysis using patient level institutional data shows that whilst overall neoadjuvant pathway may have a slight survival advantage, for a small subsection of patients who have the highest probability of receiving early R0 resection and receiving adjuvant therapy, upfront surgery may optimise their individual outcomes. This raises the question as to how we can therefore identify this subgroup. Furthermore this begs the question as to whether those patients who do not undergo resection in the neoadjuvant pathway would have significantly improved survival times had they been treated within the upfront surgery pathway.

The returning bomber planes are therefore like the RCTs and other study populations involving only patients who have received multimodal treatment: a self-selecting group. Just as the planes returned despite sustaining attacks resulting in multiple bullet holes so too does study populations who received multimodal therapy in either pathway represent a self selecting group who can withstand the insult of surgery and adjuvant therapy, whether delivered pre or post-operatively. Studying only this data will not provide the answers about how to deliver personalised realistic medicine and optimise outcomes for all patients presenting with potentially resectable pancreatic cancer through personalised treatment pathway selection. Wald instead asked what the narrative was contained within the unseen data. In other words he recognised that they needed to study the pattern of bullet holes in the planes that did not return and reinforce the planes in these places (Wald, 1980; Mangel & Samaniego, 1984). In order to take research further and develop the narrative surrounding the treatment of potentially resectable pancreatic this research therefore set about engaging with the "unseen" data that is often ignored due to being flawed, uncertain, proximate and sparse.

The Narrative Within Flawed Uncertain Proximate and Sparse (FUPS) Data:

A key aim of this research was to address the gap that exists between the ideal of comprehensive, clear data used in complicated contexts, and the reality of decision making within pancreatic cancer

management pathways that relies on FUPS data used within the context of complexity (Wolpert & Rutter, 2018). Although data from large multicenter RCTs are held in high esteem and often readily accepted within medicine with limited criticism, the reality is that the majority of daily clinical decisions are made based on FUPS data out with controlled trial conditions and concerning patients who are not selected based on a strict inclusion criteria.

The challenge therefore was to develop decision support models that not only pay sophisticated attention to the merits but also the potential detriments of using FUPS data as well as considering the implications of the complexity of the healthcare system in both research and practice (Plsek & Greennalgh, 2001; Rutter *et al*., 2017; Wolpert & Rutter, 2018). To move beyond the biomedical model as the only model of evidence and also acknowledge the dangers of both over-interpretation of FUPS data as well as non-use, the aim of the modeling methods using FUPS data within this research was to open up conversations on findings rather than treating them as definitive facts (Wolpert & Rutter, 2018). Therefore the key principles and framework developed by Wolpert & Rutter (2018) for analysing FUPS data within statistical models were applied. This included treating data as a partial remnant with findings presented to convey associated limitations to interpretation, avoidance of 'black box' statistical modeling in favour of transparency and clarity, and triangulation to contextualise findings from models based on FUPS data to explore how other information and modeling techniques refute or support these findings (Wolpert & Rutter, 2018).

Towards Personalised Realistic Medicine: the narrative on marginal gains emerging from FUPS data

As previously discussed the analysis within this research looking at overall survival times in both neoadjuvant and upfront surgery pathways showed a picture whereby neither pathway was conclusively superior. Triangulation with Bayesian One-way ANOVA and log-linear regression analysis of the West of Scotland Pancreatic Unit institutional database of both surgery first *versus* neoadjuvant therapy treatment pathways for resectable PDAC, did not demonstrate statistically significant superiority of one pathway (one-way ANOVA *P value:* 0.808 and 0.163 respectively; log-linear regression *P value*: 0.87 and 0.871 respectively) (Appendix O). Surgery first pathway did demonstrate superiority in achieving R0 resection (one-way ANOVA *P value:* 0.025; log-linear regression *P value*: 0.025; surgery first posterior mean: 0.795; 95% CI 0.698-0.891 *versus* neoadjuvant posterior mean: 0.550; 95% CI 0.360-0.740). However receipt of multimodal treatment within either pathway was found to be statistically significant in determining whether survival outcomes fell within the good (36months or more) or poor (12 months or less) post resection survival time categories (one-way ANOVA *P value:* 0.000; linear regression and log-linear regression *P value*: 0.00) although there was no statistically significant difference between pathways in achieving multimodal treatment (one-way ANOVA and linear regression *P value:* 0.150) (Appendix O).

Markov cohort analysis was able to go further than previous studies in exploring the impact on survival outcomes of R0 and R1 resection

in either pathway. In itself the benefits of multimodal treatment and R0 resection in either pathway are not new findings. However, by engaging with FUPS data Markov cohort analysis was able to go further to provide further insights into the impact of other aspects of, and events within, the system on overall survival outcomes.

The results from the Markov cohort analysis demonstrate that grade 3-4 post-operative complications marginally affect quality adjusted survival time. Experiencing toxicity from adjuvant therapy however has a more noticeable impact on quality adjusted survival time. However where both post-operative complications and toxicity from adjuvant therapy are experienced the quality-adjusted survival outcomes were closer to that of cases with the same resection margin that did not progress to receiving adjuvant therapy in models populated with synthesised data for all potentially resectable cases and resectable only cases at presentation. An exception to this was observed in the Markov cohort analysis where the model was populated with institutional data on resectable only cases. Here R0 resection and adjuvant therapy, even where both post-operative complications and toxicity from adjuvant therapy were experienced, had superior survival outcomes.

This analysis was therefore beginning to build a richer picture of the management of pancreatic cancer beyond the established benefits of multimodal treatment and R0 resection. Opportunities for marginal gains in avoiding postoperative complications and toxicity from adjuvant therapy were beginning to emerge. However, accepting all such findings as a partial remnant and employing the principle of

transparency of analysis to convey the limitations of interpretation stemming from the FUPS characteristics of data (Wolpert & Rutter, 2018) both deterministic and probabilistic analysis were undertaken which revealed further key aspects of the system that could guide more personalised treatment selection.

For the Markov model based on synthesised data of potentially resectable cases one-way deterministic sensitivity analysis showed that neoadjuvant pathway was superior only if the probability of resection was greater than 51.04% or below 75.68% in surgery first pathway. For the Markov model based on synthesised data of resectable only cases the probability of resection in the neoadjuvant pathway had to be greater than 47.48% for neoadjuvant therapy to maintain superiority. When the model was populated with institutional data for resectable only cases the probability of undergoing resection in the neoadjuvant pathway had to be above 34% to maintain superiority. Furthermore two-way sensitivity analysis showed that pathway superiority depended the individual patient's probability of obtaining multimodal treatment in either pathway.

The theme of the importance of obtaining multimodal treatment was further echoed in the results of the cost-effectiveness analysis. Markov Model results using synthesised data reported neoadjuvant therapy pathway gave 21.27 QALMs at a cost of £109879.65. Surgery first gave 17.59 QALMs at a cost of £101251.75. Neoadjuvant therapy therefore had an incremental cost of £8627.90 more than surgery first for an incremental effectiveness of 3.68 QALMs and an ICER of

£2344.16. Using West of Scotland Pancreatic Unit data neoadjuvant therapy gave 26.71 QALMs at a cost of £117426.89. Surgery first gave 21.27 QALMs at a cost of £109879.65. Neoadjuvant therapy therefore had an incremental cost of £29126.08 more than surgery first for an incremental effectiveness of 8.48 QALMs and an ICER of £3433.07. DES model results using synthesised data reported neoadjuvant pathway gave 16.45 QALMs at a cost of £81934.19. Surgery first gave 13.84 QALMs at a cost of £69630.42. Neoadjuvant therapy therefore had an incremental cost of £12303.77 more than surgery first for an incremental effectiveness of 2.61 QALMs and an ICER of £4708.51. Based on West of Scotland Pancreatic Unit data neoadjuvant therapy gave 21.60 QALMs at a cost of £72083.26. Surgery first gave 13.87 QALMs at a cost of £45813.65. Neoadjuvant therapy therefore had an incremental cost of £26219.61 more than surgery first for an incremental effectiveness of 7.73 QALMs and an ICER of £3390.51. In both Markov and DES models using synthesised and institutional data the main driver of the ICER was receipt of multimodal treatment in the neoadjuvant pathway.

The emergence of the impact on outcomes of individual probability thresholds across competing treatment pathways continued in the DES analysis, which corroborated the findings from the Markov decision-analysis. For resectable cases of pancreatic cancer the probability of resection in the neoadjuvant pathway had to be greater than 38% for neoadjuvant pathway to be superior. Furthermore the probability of R0 resection in the neoadjuvant pathway had to be greater than 15.4%. This analysis however used FUPS data to go further than existing studies to analyse the outcomes for patients

who failed to undergo resection. For patients in the upfront surgery pathway who were found to have unresectable disease the mean survival time was 8.50months (6.41QALMs). Patients in the neoadjuvant pathway who did not undergo surgery had a mean survival time of 11months (8.04QALMs). These patients are at the centre of the controversy surrounding the use of neoadjuvant therapy for cases of pancreatic cancer that are resectable at presentation as critics of this approach highlight the dangers of losing the window of resectability. For this group of patients their probability of resection in the upfront surgery pathway had to be greater than 29% for the upfront surgery pathway to be the superior choice. As these patients are presenting with resectable disease it seems likely that such a threshold would be reached, making upfront surgery the superior pathway choice for them. However, proponents of the neoadjuvant approach for resectable pancreatic cancer have argued that this particular group of patients represents more aggressive tumour types that are successfully filtered away from futile surgery through neoadjuvant approach. The expected incremental value in terms of months survival time for patients who did not undergo surgery in the neoadjuvant pathway, if treated in the upfront surgery pathway, were 1.5months (-0.86QALMs), 3.5months (1.14QALMs) and 5.5months (3.14QALMs) corresponding to a probability of resection in the surgery first pathway of 47%, 70.5% and 94% respectively. This adds a new dimension to the current research narrative and begins to quantify in more personalised terms the realistic potential survival benefit from surgery.

The use of FUPS data to create a simulation model that more accurately reflects the complex systems of care delivery whereby, unlike RCTs complexity is not controlled for, allowed this breakthrough in quantifying the potential survival benefit from surgery in personlised terms to be taken forward. When the results from recent RCTs were incorporated into the DES models interesting thresholds emerged.

Preliminary results from the PREOPANC-1 trial (Van Tienhoven *et al*., 2018) were incorporated to compare outcomes for borderline resectable cases treated in both the surgery first and neoadjuvant pathways. Although this trial reported superiority of neoadjuvant therapy this was found to be dependent on the patient's probability of resection in the neoadjuvant pathway being greater than 18.9% within the DES model. The improved outcomes reported by Conroy *et al*. (2018) with mFOLFIRINOX in the adjuvant setting of the surgery first pathway were significant and their incorporation within the model increased the overall outcome from the surgery first pathway from 16.56months to 38.43months. However, the apparent benefit of adjuvant therapy only applied to those who underwent resection and receive adjuvant therapy therefore threshold analysis within the DES models incorporating these findings showed that the probability of resection in the surgery first pathway had to be greater than 54% and that the probability of receiving adjuvant therapy also had to be greater than 8% for the surgery first pathway to be the superior option for cases of resectable pancreatic cancer.

Overall both the Markov and DES models were going beyond previous studies to provide predicted thresholds that must be achieved in real life, where complexity is not controlled for, in order that the reported positive findings of emerging trials be applicable at individual patient level. The next challenge was to explore how complexity theory could drive research even further to pre-operatively make individual predictions of outcomes across competing treatment strategies.

Engaging With Complexity: what has been learned through the application of Cillier's Lean Ontology?

Bayesian One-way ANOVA and log-linear regression analysis of the West of Scotland Pancreatic Unit institutional database identified AJCC stage (*P value:* 0.000), tumour size above or below 3 centimeters (*P value:* 0.005), ASA grade (*P value:* 0.002), albumin(*P value:* 0.047) and modified Glasgow Prognostic Score (*P value:* 0.031) as being statistically significant in predicting survival of 36 months or more post resection of PDAC (Appendix O). Bayesian one-way ANOVA and log-linear regression analysis identified modified Glasgow prognostic score and tumour size greater than 3 centimeters as being significant in predicting survival of 12 months or less (*P* value 0.505 and 0.037) (Appendix O).

This Bayesian linear regression analysis of all pre-operative factors produced a model of good fit to the regression line (R: 0.955; R squared 0.912; *P value* 0.006) and further built the case for focusing on individual pre-operative factors to make more clinically relevant

prediction to help decide best treatment pathway. However it shares a key limitation of existing studies, mainly based on log linear regression techniques, that fundamentally regard prognosis as an isolated event at a pre-determined time hence neglecting the dynamic nature of the care processes reflected in the unfolding relationships between variables with expected patient outcomes evolving as more information becomes available (Verduijn *et al*., 2007). To attempt to address this BBN models were created to make pre and post operative predictions of prognosis post resection of PDAC across competing treatment strategies.

BBNs Models Performance.

The creation of a BBN allowed the novel utilisation of knowledge from existing PubMed studies in a clinically more meaningful way for individual patients and their clinicians. This also means that the model, based on the wider collective body of existing evidence, overcomes the limitations of many existing models that lack generalisability as they are largely based on single institutional database analysis with the potential for inherent bias that this creates. This also allows the BBN to make predictions even when data is missing through probabilistic inference with predictions made based on global averages of the patient population. Secondly this model goes beyond the few existing nomograms and prognostic models by providing personalised predictions based on pre-operative information therefore being of more value in patient counseling and decision making throughout the patient journey. Thirdly this model is unique in its ability to make personalised

predictions of outcome across the competing treatment strategies of upfront surgery and neoadjuvant therapy.

BBN predicting poor prognosis pre-operatively achieved an AUC of 0.70 (*P* value 0.001; 95% CI 0.589-0.801) where data on all other nodes, apart from tumour markers, were available. With an additional one and two data points missing a statistically significant AUC of 0.70 was maintained. BBN performance in prognostic updating where the outcome predicted was poor prognosis achieved an AUC of 0.80 (*P value*: 0.000; 95% CI: 0.678-0.862) when all other data, apart from tumour markers, was available. An AUC of 0.80 was maintained until more than 6 pre-operative data points, and up to and including 2 post-operative data points, were missing. At this point the BBN achieved an AUC of 0.70 (*P value*: 0.000; 95% CI:0.667-0.818) which was maintained with over 6 missing pre-operative data points and up to and including 4 missing post-operative data points (*P value*: 0.000; 95% CI: 0.660-0.788). This performance compares favorably to existing predictive model development studies aiming to predict poor post pancreatic cancer resection prognosis. Existing models based on multivariate cox proportional hazard regression techniques report an AUC of between 0.7 and 0.887 (Kanda *et al*., 2014b; Hsu *et al*., 2012; Shen *et al.*, 2018; Balzano *et al*., 2017; Walczak & Velanovich 2017). However many are based on single institution databases (Kanda *et al*., 2014b; Hsu *et al*., 2012; Balzano *et al*., 2017) and failed to undergo external validation (Kanda *et al*., 2014b; Hsu *et al*., 2012) One study, based on single institution data, used artificial neural network technique to predict 7month mortality post-resection and reported an AUC of 0.6576 but did not perform

external validation (Walczak & Velanovich 2017). One study used Bayesian modeling techniques and National Registry data to predict 6month, 1,3 and 5year survival and achieved a c-statistic of 0.65 (Smith & Mezhir, 2014).

BBN pre-operative predicting good prognosis achieved an AUC that ranged from 0.94 (*P value* 0.002; 95% CI 0.859-1.000) for 0 missing data points in addition to the missing tumour marker data to AUC 0.74 (*P value* 0.000; 95% CI 0.660-0.809) accepting more than 4 additional missing data points. The prognostic updating performance of BBN predicting good prognosis achieved AUC 0.97 (*P value* 0.000; 95% CI 0.908-1.000) for 3 missing data points in pre and post-operative validation dataset to AUC 0.75 (*P value* 0.000; 95% CI 0.655-0.838) accepting more than 6 missing data points in the pre and up to and including 3 missing data points in the post-operative validation dataset. The latter was the only point at which BBN performance had an AUC under 0.80. The BBN model's performance again compared favorably with findings from previous studies. Existing model development studies that aimed to predict post resection survival time of 3-years or more for resected PDAC report an AUC between 0.63 and 0.884 with all relying on post-operative information and only one of these studies having undergone external validation (Dasari *et al*., 2016; Pu *et al*., 2018; Brennan *et al*., 2004; Miura *et al*., 2014; Xu *et al*., 2017; Botsis *et al*., 2009; Liu *et al*., 2018; Smith & Mezhir, 2014; Pu *et al*., 2017; Katz *et al*., 2012b).

5.3 Embracing Cilliers' Lean Ontology of Complexity: an assessment of the strengths and limitations of this research

Clinical decision making in the management of pancreatic cancer is challenging and complex and therefore requires a more sophisticated process to represent the dynamic and evolving relationship between variables in determining outcomes as more information becomes available (Kabir *et al.*, 2015; Verduijn *et al*., 2007; Lucas *et al.*, 2004; Velikova *et al*., 2014; Lewis & Vollmer, 2012). In keeping with Cillier's lean ontology of complexity the BBNs developed here contained a large number of elements with a rich level of interaction. By encompassing a large number of concepts, elements and ideas applied to an even larger number of combinations implies a finite number of elements that will impose an artificial boundary in the operational research system (Kruger *et al*., 2019).

Some have argued that this is necessary from the observer's perspective to study a complex system (Cilliers, 2008; Mowat & Davis, 2018). In a practical sense however this means that the models developed within this research therefore have an artificial boundary imposed by the limitations of the currently available data and the limitations this imposes on this research must be acknowledged. In particular information regarding the quality-of-life throughout the patient journey is limited for pancreatic cancer. Although the best available quality-of-life indicies in keeping with the few existing decision analysis model studies was used to maximise accuracy and facilitate comparability of findings further research into this vital area must be progressed and findings incorporated into the

statistical models. Another area where data was limited was indirect costs accumulated by both the patients and their carers, which meant that this was lacking in the cost effectiveness analysis models. Moving forward the integration of such qualitative data into statistical models will be as pertinent as the integration of emerging breakthroughs in biomarker and gene targeted treatment sequencing. Accepting Merali's (2006) conceptualisation of the world as a networked world it stands that a collection of concepts and techniques does not constitute an operational research system and equally an operational research application does not exist in isolation (Merali, 2006; Kruger *et al*., 2019). This perspective contained within the broader lens of complexity theory means that the future application of this research can be viewed in a much richer sense. Whilst elements, such as society and the economy for example, may not interact with the application in a deterministic way they will interact and merge with the application (Kruger *et al*., 2019). This has been witness in the cost-effectiveness analysis of upfront surgery *versus* neoadjuvant therapy where the WtP threshold acceptability curves demonstrated how defining the treatment intention as realistically curative or palliative, and therefore determining the currently acceptable WtP threshold, had significant impact on deciding which treatment pathway was cost-effective. It follows that as future developments in biomarker and gene targeted therapy for earlier disease detection and targeted treatment sequencing emerge and become incorporated within the future artificial boundary of the operational research system, elements outside this boundary will continue to interact and merge with the application to determine cost-effective resource allocation.

Accepting that an operational research model is only meaningful within the real-life context within which it is applied it therefore follows that any such model will be meaningless unless it can interact with the environment within which it operates (Cilliers, 1998). This not only means that elements within the model can interact mathematically but also that they represent the dynamic interaction with the wider environment within which the model is operating. This interaction between the operational research application and reality in this case was transference of information to facilitate better shared decision making (Cilliers, 1998). The impact of the previously outlined uncertainties and limitations in the quality of the existing body of research must therefore be acknowledged. There is a tendency to publish positive findings, which could mean that the ability of models to support decisions where the outcomes will never be favourable are weaker, or that models could produce overly optimistic predicted survival outcomes. However, the aim of the modeling methods using FUPS data within this research was to move beyond the biomedical model as the only model of evidence to open up conversations on findings rather than treating them as definitive facts whilst acknowledging the dangers of both over-interpretation of FUPS data as well as non-use (Wolpert & Rutter, 2018). This included treating data as a partial remnant with findings presented to convey associated limitations to interpretation, avoidance of 'black box' statistical modeling in favour of transparency and clarity, and triangulation to contextualise findings from models based on FUPS data to explore how other information and modeling techniques refute or support these findings (Wolpert & Rutter, 2018). Overall

model outputs were in keeping with findings from the existing body of research but were able to go beyond this and add a further dimension to the current debate as will now be explored.

The BBN was developed with the appreciation that the relationships amongst the elements were dynamic, far from equilibrium, and did not always have a linear relationship. Within a complex system non-linearity between components, the environment and whole systems results in a state of non-equilibrium (Capra, 2007; Prigogine, 1987). For Cilliers the concepts of non-linearity is closely aligned with the principle of asymmetry. Personalised realistic medicine within the context of this research was seeking the most effective delivery of treatment for individual patients. This introduces an element of competition hence introduces asymmetry. For Cilliers non-linearity, asymmetry and competition are all inevitable components of complex systems (Kruger *et al*., 2019). The concept of non-linearity within an operational research model means that small causes can have large results (Cilliers, 1998). Interactions with the world are dynamic therefore systems in non-equilibrium have multiple states of states and become more robust through a process of adaptation than static systems operating in a state of equilibrium (Kruger *et al*., 2019; Capra, 2007; De Villiers-Botha & Cilliers, 2010). Furthermore within the BBN models each element was ignorant of the behaviour of the system as a whole but the system was allowed to have a history. Together this meant that the impact of even small isolated adverse events, such as postoperative complications, could be assessed for their potential impact on other aspects of the system

such as the probability of receiving adjuvant therapy and the subsequent impact on overall prognostic prediction.

Employing Cillier's ontology of complexity meant that statistical techniques had to be developed to engage not only with the complexity of the system but also to cope with uncertainty. As the existing data fulfills the FUPS criteria the principles of using FUPS data were incorporated within this ontology of complexity. For complex problems it is difficult to establish mutual relationships among nodes considering that the number of conditional probabilities increase exponentially to the number states of the parent and child nodes (Nadkarni & Shenoy, 2001; Tang & McCabe, 2007). Previous BBNs in healthcare have relied on expert judgment to quantify the required probability relationships. However, under such complex and large conditional probability tables (CPT) expert derived conditional probabilities can become inconsistent and it has been found to be more reliable to construct CPTs from data (Tang & McCabe, 2007; Hager & Andersen, 2010; Kabir *et al*., 2015). Therefore to engage with FUPS data and address challenges of variable selection, generalisability, missing data and pre-operative application that limit existing predictive models, principles of triangulation and transparency were employed (Wolpert & Rutter, 2018). This was done through a two stage weighting process, adapted from Zhao & Weng (2011) to synthesise PubMed survival analysis data by placing existing, even contradictory data, within the context of the wider body of existing data through a second stage normalisation weighting process. The conditional probability dependencies between variables were then represented through

Bayes theorem, which has been proven to be an effective way of handling uncertainty (Ismail *et al*., 2011; Sun & Shenoy, 2007).

This holds several advantages when modeling treatment pathways for potentially resectable PDAC. Through Bayes theorem the prior distribution and observed data are combined to update knowledge in the form of the posterior distribution (Pearl, 1988; Fenton & Neil, 2019). Therefore BBNs allow the modeling of relationships between variables at various stages of the healthcare process, with predictions of outcomes evolving throughout the process by utilizing all available patient data at that time (School *et al*., 2013). This means that the model could not only make predictions of outcome pre-operatively but also perform prognostic updating at the post-operative stage of the patient journey. It also means that this modeling technique overcomes limitations of existing predictive models including: lack of generalisability, bias inherent in overreliance on single institutional databases and dependence on post-operative information to make predictions. Where patient information is limited probabilistic inference can still make predictions based on global averages of the patient population making the model better able to cope with missing data (Verduijn *et al.,* 2007; Lucas *et al.*, 2004). As more information becomes available the predictions become more patient specific (Verduijn *et al.,* 2007).

This also means that the BBN allows the management of PDAC to be modeled as an open system, another key aspect of Cillier's 'lean ontology' of complexity. When an operational research model is applied to a specific problem the model becomes exposed to the real

world as an open system with a large number of elements having an influence on its formulation (Kruger *et al*., 2019). In practical terms this means that as the BBN is used in clinical practice, and further patient level data is amassed, through Bayes theorem the posterior distribution can be constantly updated with predictions becoming increasingly accurate. This prevents the model itself from becoming a closed system with results confined to a set of variables, which would be a gross oversimplification of the system being modeled (Kruger *et al*., 2019).

This links to defining characteristics of a complex system: boundary setting, lack of complete knowledge and responsibility. Complete knowledge of a complex system is not possible but rather knowledge in terms of a certain framework is (Kruger *et al*., 2019). Hence the generation of knowledge within a complex system is exploratory and temporary (Cilliers, 2005b) which complements Wolpert & Rutter's (2018) framework for using FUPS data that also contests that all findings be treated as partial remnants with limitations stemming from FUPS characteristics conveyed. Both the artificial nature of boundary setting and the provisional nature of knowledge means that a level of uncertainty will prevail in model outputs which means that responsibility must be taken for intended and unintended consequences when a system does not reflect reality (Cilliers, 2008; Woemann & Cilliers, 2012; Ackoff, 1974; Gallo, 2004; Ormerod & Ulrich, 2013) particularly as boundary definitions involve a value based judgment (Audouin *et al*., 2013; Ulrich, 1983; Midgley, 2000).

Cilliers (2005b) and Woermann (2010) argue that this represents a challenge to develop a new kind of scientific understanding (Cilliers, 2007). Cilliers postulated that one of the defining characteristics of a complex system is its emergent properties, which cannot be reduced to the system component properties (Cilliers, 2010). Therefore complexity emerges as a result of the dynamic and non-linear interactions between elements within an open system (Cilliers, 1998). The magnitude of emergence can be difficult to quantify (Paul *et al.*, 2014) particularly as emergence can take many forms including deeper understanding (Kruger *et al*., 2019).

By viewing the management of potentially resectable pancreatic cancer through the lens of complexity theory and applying the methods presented here to develop a more sophisticated process of engaging with FUPS data to represent the relationship of interacting elements in determining outcomes (Kabir et al., 2015) other findings and insights began to emerge. In keeping with Wolpert & Rutter's (2018) framework for using FUPS data Pearl's inwards analysis and broadcasting analysis were used to perform sensitivity analysis of the BBN models (Pearl, 1988; Fenton & Neil, 2019). Sensitivity analysis assumed all input parameters in the model are uncertain and therefore determines how sensitive results are in relation to changes in observable variables (Yang *et al.*, 2009; Ismail *et al*., 2011). Therefore sensitivity analysis served as an adjunct to decision-analysis (Fenton & Neil, 2019) and allowed for the exploration of emergence.

The results of BBN sensitivity analysis showed that for the pre-operative BBNs tumour factors had the greatest impact on outcomes, followed by patient factors. When post-operative data was incorporated into the BBN post-operative factors and surgical pathology had greatest impact on output followed by tumour factors and patient factors. This corroborates numerous previous studies that have established that detecting disease early improves survival, hence the importance of tumour factors (Winter et al., 2012). Furthermore it has been established that the best chance of good survival outcomes depends upon achieving R0 resection whereby all tumour is completely removed (Versteijne *et al.*, 2018), which supports the importance of surgical pathology on impacting on BBN outcome. Receipt of multimodal treatment has also been established as key to achieving best possible survival outcomes in numerous studies (Bradley *et al*., 2018; Bradley *et al.,* 2019a; Versteijne *et al*., 2018; Neoptolemos *et al*., 2001) which supports the findings from sensitivity analysis that post-operative factors have significant impact on BBN output as this includes post-operative complications, which in turn affects the recovery time and time to, or indeed whether, patients receive adjuvant therapy which is also included in post-operative factors (Winter *et al*., 2012).

Emergence as a key characteristic of modeling PDAC management as a complex system meant that for the first time a prognostic model could quantify the significant impact of cumulative patient factors on survival outcome. Previously these individual variables were found to play a limited role in survival analysis yet formed a significant component of subjective clinical judgment on a regular basis in

clinical practice. In the BBN predicting poor prognosis sensitivity analysis showed that patient factors were the second most significant determinant for pre-operatively predicting a poor prognostic outcome and the fourth most significant in the post-operative setting when performing prognostic updating. The BBN predicting good prognostic outcome found that patient factors were the third most important factor in predicting outcome in the pre-operative setting. Interestingly patient factors were found to have more of an impact on predicting poor prognosis than good prognosis. This could be explained by the fact that in practice clinicians are less likely to operate on frail and unfit patients therefore the data for such patients is less likely to form part of post resection survival analysis. This finding makes a case not only for better patient selection at the pre-operative stage of the patient journey but also emphasises the importance of developing statistical models that can handle a variety of FUPS data and deal with uncertainty to assist decision making at an individual patient level.

5.4 Conclusion: Future Direction of Research

Building on Markov and DES models, the BBN models developed through this research marks a significant step towards the delivery of personalised realistic medicine in pancreatic cancer management. The methods proposed could be applied to other forms of cancer where barriers exist to developing large detailed patient databases, where FUPS characteristics apply to the available data, and where pathway complexity makes expert elicitation for large CPTs less reliable. Given the scarcity and FUPS characteristics of data, lack of

technical and financial resources and limited experience of many clinicians in Bayesian statistics, the proposed BBN could in future help to guide cost-effective resource allocation and prioritisation, perform risk-benefit analysis and support rational decision making in a way that is both intuitive and transparent to clinicians.

The next phase of this research will be to integrate larger international patient databases into the BBN models and prospectively validate the models' performances against other institutional databases. This however will form only one aspect of the BBN models' ongoing assessment and development. Any such decision support model will have to be integrated into the much wider complex system of clinical healthcare delivery. Whilst technological innovations have moved apace and are widely postulated as contributing to health (Garber *et al*., 2014), there is a scarcity of literature exploring the sustainability of technology-supported change within healthcare (Grin *et al*., 2010). Furthermore many such innovations have been plagued by non-adoption, abandonment and difficulties in scale-up (van Limburg *et al*., 2011). This is thought to be due to dynamic interactions between multiple factors at multiple organisational levels (Greenhalgh *et al*., 2017a). Therefore a vital next step in this research will be to attempt to identify, understand, and address these interacting challenges using the Non-adoption abandonment Scale-up Spread and Sustainability (NASS) framework to design future studies that are interdisciplinary, nondeterministic and designed to examine the relationship between human action (both patient and clinician) and the wider organisational and system context (Greenhalgh *et al*., 2016).

The future application of this research will be in integrating breakthroughs in biomarker and genomic trials for early detection and targeted treatment sequencing with clinical data to make increasingly precise, personalised predictions of outcomes throughout the trajectory of the individual patient's journey. This stands to accelerate the clinical application of our ever expanding knowledge base. However, as previously discussed, such breakthroughs will only ever expand the artificial borders of statistical models and therefore be of limited impact unless the statistical models developed continue to be viewed as open complex systems. This means that future research focus must be on continuing to develop methods of engaging with complexity and developing these alongside RCTs and genomic research as well as more qualitative studies seeking to establish the impact on individual patient quality-of-life (Obermeyer & Lee, 2017, Star, 2002; Fraser & Greenhalgh, 2001; Greenhalgh & Papoutsi, 2018; Long *et al.,* 2018; Law & Mol, 2002; Tsouka, 2017).

Chapter 6

Conclusion

Pancreatic cancer is one of the most challenging malignancies, characterised by increasing incidence rates globally and poor survival outcomes despite advances in treatment modalities. Over 80% of cases present at an advanced disease stage. For cases that are amenable to surgical resection the current standard of treatment is surgical resection followed by adjuvant mFOLFIRINOX (Khorana *et al*., 2019). However, despite a growing number of RCTs reporting increased survival outcomes with adjuvant therapy, up to 50% of patients who undergo resection fail to progress to receiving adjuvant therapy due to a combination of factors that include early disease recurrence and a delay in commencing treatment due to post-operative complications and the impact of pre-existing comorbidities (Winter *et al*., 2012). Consequently the potential benefits of costly and high-risk surgery are nullified. Furthermore this challenges the prevailing narrative that surgical resection is the only curative treatment as for many patients the reality is that surgery is of limited benefit.

This has resulted in a renewed interest in neoadjuvant therapy as a means of increasing the delivery of multimodal treatment and diverting patients with more aggressive disease away from futile surgery. However, critics have highlighted the potential for losing the window of resectability with this approach. The current evidence

base underpinning the management of potentially resectable pancreatic cancer is highly heterogeneous, widely contested and permeated with uncertainty, ambiguity and controversy. In response the prevailing narrative regarding the current focus within pancreatic cancer research is to seek clarity through biomarker and gene targeted diagnosis and treatment sequencing as well more large multicentre RCTs to determine treatment pathway superiority for potentially resectable pancreatic cancer. However, recent RCTs involving both upfront surgery and neoadjuvant approaches, rather than provide such clarity, have produced contradictory findings. Given the afore mentioned facts surrounding the reality of receiving multimodal therapy this calls into question the extent to which the results of such RCTs, where real-world complexity is controlled for, are applicable in the real-life clinical setting where clinical decisions have to be made. The same argument applies to the much-anticipated results from genomic and biomarker studies as the reality is that patients are, and always will be, more than their genomes. Therefore the reality is that clinical decision making will require that such emerging data be integrated with the ever expanding amounts of pre-existing clinical data. Clinicians will therefore continue to be expected to make decisions based on FUPS data. Furthermore this ongoing challenge exists within the wider political context of a move towards personalised realistic medicine with the associated expectation of personalised shared clinical decision-making.

The aim of this research was to facilitate a shift in pancreatic research focus towards personalised realistic medicine through statistical modelling that will facilitate better shared decision making

with patients to optimise individual patient outcomes. By amalgamating operational and healthcare research disciplines this research sought to be theory driven and empirically focused from a complexity perspective. Specifically the potential for simulation modelling in the study of complexity in healthcare was explored to attempt to expand capabilities for handling uncertainty, the emergent and engage in disagreements through the use of FUPS data.

Key points and novel findings began to emerged from this research that added a further dimension to the current debate surrounding the treatment of potentially resectable pancreatic cancer. Synthesisation of existing trials could not conclude that either upfront surgery or neoadjuvant treatment pathway was conclusively superior for the management of potentially resectable pancreatic cancer. Both Markov and DES modelling for decision-analysis emphasised the importance of multimodal treatment in optimising patient outcomes. This was corroborated in cost-effectiveness analysis of both pathways where once again neither pathway was conclusively superior but the main driver of the ICER was multimodal treatment.

Markov and DES simulation modeling for decision-analysis went further to reveal that superior pathway selection is determined by individual patient and tumour factors. DES modeling also, for the first time, quantified the modest anticipated survival outcomes for patients who did not proceed to surgery in the neoadjuvant pathway has they been treated in the upfront surgery pathway. Due to the prevailing narrative that surgery was the only potential cure for

pancreatic cancer it had been assumed by critics of neoadjuvant therapy that such patients would have had significantly superior survival outcomes in the upfront surgery pathway. New insights were also gained through DES modeling by integrating the results of recent RCTs into a simulation model that reflected a system where complexity was not controlled. This revealed individualised probability thresholds within each pathway that must be obtained for the results of such trials to be applicable within a real-life clinical setting.

BBNs were then constructed to model the decision making process for individual patients with potentially resectable pancreatic cancer. This process was modeled as an open complex adaptive system with multiple variables displaying a high level of rich, dynamic, non-linear interactions whilst each variable was simultaneously ignorant of the behaviour of the whole system. The result was the first BBN that could make individual pre-operative predictions of post resection prognosis across competing treatment strategies and that had the ability to perform prognostic updating when more information became available. When validated against a prospective patient database the model performed with a level of predictive accuracy rivaling existing predictive models, even when only based on pre-operative data.

The impact of this research is that it offers new insights and adds a further dimension to the current debate and narrative surrounding pancreatic cancer management, which moves the focus towards personalised realistic medicine. This study marks a potentially

significant step towards achieving the delivery of personalised cancer care. In the clinical setting the BBNs, by engaging with complexity and handling the emergent, have the potential to have an immediate impact on improving patient counseling and facilitating better shared decision making by providing a mechanism to communicate and transmit the complex and data rich empirical narrative surrounding a diagnosis of potentially resectable pancreatic cancer on a personalised level that includes being better able to explain the impact of "what if" scenarios on anticipated prognosis.

As patient databases mature and develop in complexity globally, so too should predictive modeling become more sophisticated at integrating multiple complex databases to make individualised patient predictions and support clinical decision making even under uncertainty. The methods developed for using FUPS data to construct the BBNs offer a potential vehicle to incorporate anticipated breakthroughs from the wider area of precision medicine by integrating complex genomic and clinical databases to deliver personalised realistic medicine for other types of cancer, as well as pancreatic cancer, by facilitating better shared decision making through personalised predictive modeling. This could serve to accelerate the clinical application of our ever expanding knowledge-base. However, to achieve this goal future research must focus on further development and testing of the BBNs. This includes investigating issues surrounding the integration of such decision support models within the complex system that is the clinical setting.

Healthcare is both an art and a science. Patients will always be more than their disease regardless of how deep our understanding of that disease develops at a genomic level. How then can advances in pancreatic cancer research ever be made unless methods of engaging with the complex tapestry of the real-world clinical setting are developed alongside RCTs and genomic studies?

This thesis used complexity theory as a lens through which to focus the question about how to optimise outcomes for potentially resectable pancreatic cancer. In so doing it marked a paradigm shift away from the prevailing Newtonian quest within pancreatic cancer research to either deny or solve complexity. Instead the journey taken throughout this thesis has demonstrated that by developing ways of engaging with complexity, including uncertainty, debate and the emergent, new insights can be gained and inroads made into improving care delivery through the clinical application of our ever expanding knowledge base where its impacts matters most; supporting the individual patient through out their personal journey.

Reference List

Abbod, M.F. *et al*. (2006). 'Comparative study of intelligent models for the prediction of bladder cancer progression'. *Oncology*, 15, pp. 1019-1022.

Abbod, M.F. *et al.* (2014). 'Predictive modeling in cancer: where systems biology meets the stock market' *Expert Review of Anticancer Therapy*, 9, pp. 867-870.

Abbott, D.E. *et al.* (2013). 'The cost-effectiveness of neoadjuvant chemoradiation is superior to a surgery-first approach in the treatment of pancreatic head adenocarcinoma'. *Ann Surg Oncol*, 20, pp. 500-508.

Abbott, D.E. *et al*. (2012). 'Cost-effectiveness of Treatment strategies for pancreatic head adenocarcinoma and potential opportunities for improvement'. *Ann Surg Oncol*, 19, pp. 3659-3667

Abraham, R. (1991). *Complex dynamics*. Santa Cruz: Aerial Press.

Abraham, R. and Shaw, C. (1992). *Dynamics, The geometry of behavior.* 2nd Edn. Reading: Addison-Wesley

Ackoff, R.L. (1974). 'The social responsibility of Operational Research'. *Oper. Res., 25*, pp. 361–371.

Ackoff, R.L. (1979a). 'The future of Operational Research is past'. *J. Oper. Res. Soc.* 30, pp. 93–104.

Ackoff, R.L. (1979b). 'Resurrecting the future of Operational Research'. *J. Oper. Res. Soc.* 30, pp. 189–199.

Ades, W.E. *et al.* (2006). 'Bayesian methods for evidence synthesis in cost-effectiveness analysis'. *Pharmacoeconomics,* 24(1), pp. 1-19.

Albert, D.A. (1978). 'Decision theory in medicine: a review and critique'. *Milbank Mem Fund Q Health Soc.*, 56(3), pp. 362–401.

Alexakis, N. *et al.* (2015). 'High serum CA 19-9 but nottumour size should select patients for staging laparoscopy in radiological resectable pancreas head and peri-ampullary cancer'. *EJOS*, pp. 265-269.

Alexandrou, D., Skitsas, I., Mentzas, G. (2011).'A holistic environment for the design and execution self-adaptive clinical pathways'. *IEEE Trans on Information Technology in Biomedicine*, 15(1), pp. 108-118.

Al-Hawary, M.M. *et al*. (2014). 'Pancreatic ductal adenocarcinoma radiology reporting template:consensus statement of the Society of Abdominal Radiology and the American Pancreatic Association'. *Radiology*, 270 (1), pp. 248-260

Allen, V.B. *et al.* (2013). 'Diagnostic accuracy of laparoscopy following computed tomography(CT) scanning for assessing the resectability with curative intent in pancreatic and periampullary cancer'. *Cochrane Database of Systematic Reviews,* Issue 11. doi: 10.1002/14651858.CD009323.pub2.

Al-Sukhun, S. *et al*. (2003). 'Chemoradiotherapy in the treatment of regional pancreatic carcinoma: a phase II study'. *Am J Clin Oncol*., 26, pp. 543-549.

Altman, D.G. (2001). 'Systematic reviews of evaluations of prognostic variables'. *BMJ,* 323, pp. 224–228

Altman DG. (2007). 'Prognostic models: a methodological framework and review of models for breast cancer'. In: Lyman GH, Burstein HJ, eds. *Breast cancer*: *Translational therapeutic strategies*. New York: New York Informa Healthcare, pp. 11–26.

Altman, D.G. and Lyman, G.H. (1998). 'Methodological challenges in the evaluation of prognostic factors in breast cancer'. *Breast Cancer Res Treat*, 52, pp. 289–303.

Altman, D.G. and Riley, R.D. (2005). 'Primer: an evidence-based approach to prognostic markers'. *Nat Clin Pract Oncol*, 2, pp. 466–472.

Altman, D.G. and Royston, P. (2000). 'What do we mean by validating a prognostic model?'. *Statist Med*, 19(4), pp. 453-73

Altman, D.G. *et al.* (2001). 'The revised CONSORT statement for reporting randomized trials: explanation and elaboration'. *Ann Intern Med*, 134, pp. 663–694.

Altman, D.G. *et al.* (2009). 'Prognosis and prognostic research: validating a prognostic model'. *BMJ*, 338, pp. 1432–1435.

Amanam, I. and Chung, V. (2018). 'Targeted therapies for pancreatic cancer'. *Cancers*, 10, 36. doi: 10.3390/cancers10020036

Anderson, R.A. *et al*. (2005). 'Case study research: the view from complexity science'. *Qual Health Res*., 15(5), pp. 669–85.

Andriulli, A. *et al*. (2012). 'Neoadjuvant/preoperative gemcitabine for patients with localized pancreatic cancer: a meta-analaysis of prospective studies'. *Ann Surg Oncol,* 19, pp. 1644-62.

Ansari, D. *et al.* (2013). 'Fast-track surgery: procedure-specific aspects and future direction'. *Langenbecks Arch. Surg.,* 398(1), pp. 29–37.

Ansari, D. et al. (2016a). 'Safety of pancreatic resection in the elderly: a retrospective analysis of 556 patients'. *Ann. Gastroenterol.* 29(2), pp. 221–225.

Ansari, D. *et al*. (2016b). 'Pancreatic cancer: yesterday, today, and tomorrow'. *Future Oncol,* 12(16): 1929-1946

Are, C. *et al.* (2009). 'Preoperative nomogram to predict risk of perioperative mortality following pancreatic resections for malignancy'. *J Gastro- intest Surg,* 13(12), pp. 2152-62

Argiles, J.M. (2005). 'Cancer-associated malnutrition'. *Eur. J. Oncol. Nurs.,* 9(2), pp. 39–50.

Artinyan, A. *et al*. (2008). 'The anatomic location of pancreatic cancer is a prognostic factor for survival'. *HPB*, 10, pp. 371–376.

Artinyan, A. *et al*. (2011). 'Neoadjuvant therapy is associated with improved survival in resectable pancreatic adenocarcinoma'. *Cancer*, 117, pp. 2044-2049.

Asare, E.A. et al. (2016). 'Neoadjuvant treatment sequencing adds value to the care of patients with operable pancreatic cancer'. *Journal of Surgical Oncology,* 114(3), pp. 291-295.

Asbun, H.J. and Stauffer, J.A. (2012). 'Laparoscopic vs open pancreaticoduodenectomy: overall outcomes and severity of complications using the Accordion Severity Grading System'*. J Am Coll Surg,* 215, pp. 810–9.

Ashrafian, H., Darzi, A., and Athanasiou, T. (2011). 'Evidence synthesis: evolving methodologies to optomise patient care and enhance policy decisions. In: *Evidence synthesis in healthcare'*. London: Springer: pp. 1-46.

Ashrafian, H., Sevdalis, N. and Athanasiou, T. (2010). 'Evidence-based surgery'. In: *Key topics in surgical research and methodology*. London: Springer pp. 9-26.

Assifi, M.M. *et al.* (2012). 'Surgical Apgar score predicts perioperative morbidity in patients undergoing pancreaticoduodenectomy at a high-volume center'. *J Gastrointest Surg*, 16(2), pp. 275-81

Attard, C.L. *et al.* (2015). 'Cost-effectiveness analysis of neoadjuvant pertuzumab and trastuzumab therapy for locally advanced, inflammatory, or early HER2-positive breast cancer in Canada'. *J of Med Economics,* 18(3), pp. 173-188.

Audouin, M. *et al*. (2013). 'Exploring the implications of critical complexity for the study of social-ecological systems'. *Ecol. Soc,* 18, pp.12.

Bachellier, P. *et al*. (2001). 'Is pancreaticoduodenectomy with mesentericoportal venous resection safe and worthwhile?'**.** *Am J Surg*, 182, pp. 120–129

Bafeta, A. *et al*. (2013). 'Analysis of the systematic reviews process in reports of network meta-analyses: methodological systematic review'. *BMJ,* 347, pp 3675.

Bailey, P. *et al*. (2016). 'Genomic analyses identify molecular subtypes of pancreatic cancer'. *Nature,* 531, 47–52.

Ballehaninna, U.K. and Chamberlain, R.S. (2012). 'The clinical utility of serum CA 19-9 in the diagnosis, prognosis and management of

pancreatic adenocarcinoma: an evidence based appraisal'. *J Gastrointest Oncol.* 3, pp. 105–19.

Balzano, G. *et al.* (2017). 'A preoperative score to predict early death after pancreatic cancer resection'. *Dig Liver Dis.*17(1), pp. 1590-8658

Barnes, C.D. *et al.* (1997). Proceedings of the 29th Conference on Winter Simulation 1997. Atlanta, GA: IEEE Computer Society, pp. 1280–1285.

Barro, S. and Marín, R. (2002). *Fuzzy Logic in Medicine*. Berlin Heidelberg: Springer.

Barton, P. *et al.* (2004). 'The use of modelling to evaluate new drugs for patients with a chronic condition: the case of antibodies against tumour necrosis factor in rheumatoid arthritis'. *Health Technol Assess*, 8, pp. 1-91

Bartosch-Härlid, A. *et al*. (2008). 'Artificial neural networks in pancreatic disease'. *Br J Surg*, 95, pp. 817-826.

Bassi, C. *et al*. (2005). 'Postoperative pancreatic fistula: an international study group (ISGPF) definition'. *Surgery*, 138(1), pp. 8-13.

Balzano, G. *et al.* (2017). 'A preoperative score to predict early death after pancreatic cancer resection'. *Digestive and Liver Disease*, 49, pp. 1050-1056.

Beck, J.R. and Pauker, S.G.(1983). 'The Markov process in medical prognosis'. *Med Decis Making*, 3(4), pp. 419–458

Bernardo, J.M. and Smith, A.F.M. (1994). *Bayesian theory.* New York: Wiley

Bentley, C. *et al*. (2014). 'Addressing design and suitability barriers to telecare use: has anything changed?' *Technol Disabil.*, 26(4), pp. 221–35.

Bhaskar, R. (1975). *A realist theory of science.* Leeds: Leeds Books.

Bhaskar, R. (1979). *The possibility of naturalism: a philosophical critique of the contemporary Human Sciences.* Brighton: Harvester Press.

Bhatia, A. *et al.* (2014). 'Use of soft computing techniques in medical decision making: a survey'. In *Advances in Computing*, *Communications and Informatics* ICACCI, 2014 International Conference, 2014. IEEE pp.1131-1137

Bilimoria, K.Y. *et al*. (2007a). 'Multimodality therapy for pancreatic cancer in the U.S.: utilization, outcomes, and the effect of hospital volume'. *Cancer*, 110, pp. 1227–34.

Bilimoria, K.Y. *et al*. (2007b).'National failure to operate on early stage pancreatic cancer'. *Ann Surg*, 246, pp. 173–180.

Birkmeyer, J.D. et al. (2002). 'Hospital volume and surgical mortality in the United States'. *N Engl J Med*, 346(15), pp. 1128-37.

Bisharat, N. *et al.* (2001). 'Risk of infection and death among post-splenectomy patients'. *J Infect*, 43, pp. 182–6.

Bisman, J. (2010). 'Positivism and Accounting Research: A (Personal) Primer on Critical Realism'. *Australasian Accounting, Business and Finance Journal*: 4(4), pp. 1-25.

Black, M. (1937). 'Vagueness: an exercise in logical analysis'. *Philos Sci,* 4, pp. 427-55

Black, M. (1963). 'Reasoning with loose concepts'. *Dialogue,* 2, pp. 1-12

Blazer, M. *et al.* (2015). 'Neoadjuvant modified (m) FOLFIRINOX for locally advanced unresectable (LAPC) and borderline resectable (BRPC) adenocarcinoma of the pancreas.' *Ann Surg Onco*l, 22 (4), pp. 1153-1159

Bliss, L.A*. et al*. (2014). 'Outcomes in operative management of pancreatic cancer'*. J Surg Oncol,* 110, pp. 592–8.

Bochner, B.H., Kattan, M.W. and Vora, K.C. (2006). 'International Bladder Cancer Nomogram Consortium. Postoperative nomogram predicting risk of recurrence after radical cystectomy for bladder cancer'. *J Clin Oncol*, 24(24), pp. 3967-72.

Bockhorn, M. *et al.* (2014). 'Borderline resectable pancreatic cancer: a consensus statement by the International Study Group of Pancreatic Surgery (ISGPS)'. *Surgery,* 155(6), pp. 977–988.

Bosman, F.T., Carneiro, F. and Hruban, R.H.T.N. (2018). 'WHO Classification of Tumours of the Digestive System WHO Classification of Tumours of the Digestive System.' *WHO Classification of Tumours of the Digestive System*, 2, pp. 7–8.

Botsis, T. *et al.* (2009). 'Modeling prognostic factors in resectable pancreatic adenocarcinomas'. *Cancer Informatics*, 7, pp. 281-291.

Bouchon-Meunier, B. and Zadeh, L.A. (1995). 'Fuzzy logic and its applications'. Paris, France: Addison-Wesley.

Bouwmeester, W. *et al*. (2012). 'Reporting and Methods in Clinical Prediction Research: A Systematic Review'. *PLoS Med*, 9(5), pp. e1001221. doi:10.1371/journal.pmed.1001221

Boyd, W.P. and Nord, H.J. (2000). 'Diagnostic laparoscopy.' *Endoscopy*, 32 pp. 153-158.

Box, G. E. P. (1979). 'Robustness in the strategy of scientific model building', in Launer, R. L. and Wilkinson, G. N., *Robustness in Statistics*, New York: Academic Press. pp. 201–236

Box, G.E.P., Hunter, W.G. and Hunter, J.S. (1978). 'Statistics for Experimenters: an Introduction to Design, Data Analysis, and Model Building.' New York: John Wiley & Sons.

Bradley, A. et al. (2018). 'Markov decision analysis of neoadjuvant treatment pathway versus surgery first pathway for resectable pancreatic cancer'. *Journal of Clinical Oncology*. 36(4), pp. 456-456.

Bradley, A. and Van Der Meer, R. (2019a). 'Neoadjuvant therapy versus upfront surgery for potentially resectable pancreatic cancer: A Markov decision analysis'. *PLoS One,* 14(2), pp: e0212805.

Bradley, A., Van der Meer, R., and McKay, C.J. (2019b). 'Personalized pancreatic cancer management: a systematic review of how machine learning is supporting decision-making'. *Pancreas*. 48(5), pp. 598-604.

Bradley, A., Van Der Meer, R., McKay, C.J. (2019). 'A systematic review of methodological quality of model development studies predicting prognostic outcome for resectable pancreatic cancer'. *BMJ Open* 9**:**e027192. doi: 10.1136/bmjopen-2018-027192

Braga, M. *et al*. (2011). 'A prognostic score to predict major complications after pancreaticoduodenectomy'.*Ann Surg*, 254(5), pp. 702-7.

Brailsford, S.C. (2005). 'Overcoming the barriers to implementation of operations research simulation models in healthcare'. *Clin Investig Med.* 28(6), pp. 312–5.

Brailsford, S.C. *et al.* (2009). *Stakeholder engagement in health care simulation. Proceedings of the 2009 Winter Simulation Conference.* Available at: https://doi.org/10.1109/WSC.2009.5429190. (Downladed: 28 August 2018).

Brailsford, S.C. *et al.* (2009). 'An analysis of the academic literature on simulation and modelling in health care'. *J Simulation*, 3(3), pp. 130–40.

Brailsford, S.C. *et al*. (2013). 'Overcoming the barriers: a qualitative study of simulation adoption in the NHS'. *J Oper Res Soc*. 64(2), pp. 157–68.

Braithwaite, J. *et al.* (2017). '*Complexity Science in Healthcare-Aspirations, Approaches, Applications and Accomplishments: A White Paper'*. Sydney: Macquarie University.

Braitman, L.E. and Davidoff, F. (1996). 'Predicting clinical states in individual patients'. *Ann Intern Med*, 125(5), pp. 406-12.

Brand, R.M. *et al.* (2006). 'Risk of colon cancer in hereditary non-polyposis colorectal cancer patients as predicted by fuzzy modeling: influence of smoking'. *World Journal of Gastroenterology,* 12(28), pp. 4485-4491

Brandabur, J.J. *et al*. (1988). 'Nonoperative versus operative treatment of obstructive jaundice in pancreatic cancer: cost and survival analysis'. *Am J Gastroenterol,* 83, pp. 1132–1139.

Brennan, M.F. *et al*. (2004). 'Prognostic nomogram for patients undergoing resection for adenocarcinoma of the pancreas'. *Ann Surg*, 240(2), pp. 293-8.

Bridges, J.H. (eds) (2009) *A General View of Positivism; Trubner and Co., 1865.* London: Cambridge University Press

Brignardello-Petersen, R. *et al*. 'Advances in the GRADE approach to rate the certainty in estimates from a network meta-analysis'. *Journal of Clinical epidemiology,* 93**,** pp. 36-44.

Brooks, S., Gelman, A. (1998). 'General methods for monitoring convergence of iterative simulations'. *J Comput Graph Stat.* 7**,** pp. 434–455.

Brouwer, W.B. *et al.* (2008). 'Welfarism vs extra-welfarism'. *J Health Econ*. 27(2), pp. 325-338.

Brown, S. *et al.* (2014). 'A Microsoft-Excel-based tool for running and critically appraising network meta-analyses—an overview and application of NetMetaXL'. *Systematic Reviews.* 3, 110.

Brown S *et al*. (2018). *NetMetaXL*. Available at: http://www.netmetaxl.com/index.html [Downloanded: 6th Spetember 2018]

Brunner, M. *et al*. (2019). 'R0-resection following chemotherapy or chemoradiation improves survival of primary inoperable pancreatic cancer patients. Interim results of the CONKO-007 prospective randomized multicenter trial. AB17. ALPS HPB Meeting'. Avaialble at:

http://www.alpshpbmeeting.org/wp-content/uploads/2019/01/ALPS-2019-Abstract-File-FINAL.pdf (Downloaded 2nd August 2019).

Buzby, G.P. *et al*. (1980). 'Prognostic nutritional index in gastrointestinal surgery'. *Am J Surg,* 139(1), pp. 160-7.

Byrne, D.S. (1998). *Complexity theory and the social sciences: an introduction*. New York, NY: Routledge.

Byrne, D., Callaghan, G. and Winter, T. *Complexity theory and the social sciences: the state of the art*. Abingdon: Taylor and Francis.

Cai, B. *et al*. (2013). 'Application of Bayesian networks in quantitative risk assessment of subsea blowout preventer operations'. *Risk Analysis*, 33 (7), pp. 1293-1311

Callery, M.P. *et al*. (2009a). 'Pretreatment assessment of resectable and borderline resectable pancreatic cancer: expert consensus statement'. *Ann Surg Oncol*, 16(7), pp. 1727-33

Callery, M.P., Pratt, W.B. and Vollmer, C.M. Jr. (2009b).'Prevention and management of pancreatic fistula'. *J Gastrointest Sug,* 13(1), pp. 163-73.

Callery, M.P. *et al*. (2013). 'A prospectively validated clinical risk score accurately predicts pancreatic fistula after pancreatoduodenectomy.' *J Am Coll Surg.* 216(1), pp. 1-14.

Cancer Research UK. *Pancreatic Cancer UK.* Available at: https://www.cancerresearchuk.org/about-cancer/pancreatic-cancer/survival. (Accessed: 7th January, 2019).

Canto, M.I. *et al.* (2013). 'International Cancer of Pancreas Screening (CAPS) Consortium. International Cancer of the Pancreas Screening (CAPS) Consortium summit on the management of patients with increased risk for familial pancreatic cancer'. *Gut*. 62, pp. 339–347.

Capra, F. (2007). 'Complexity and Life'. In Capra, F., Juarerro, A., Sotolongo, P. and Van Uden, J. (eds.). *Exploring Complexity*; *Volume 1*: *Reframing Complexity. Perspectives from the North and South*; Mansfield, MA, USA,: ISCE Publishing

Caro, J.J., Moller, J. and Getsios, D. (2010). 'Discrete event simulation: the preferred technique for health economics evaluation?' *Value in Health*, 13(8), pp. 1056-1060

Caruana, R. *et al*. (2015). 'Intelligible models for healthcare: predicting pneumonia risk and hospital 30-day readmission'. In: *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York, NY: Association for Computing Machinery. pp. 1721-1730.

Cardenes, H.R. *et al*. (2011). 'A phase II study of gemcitabine in combination with radiation therapy in patients with localized,

unresectable, pancreatic cancer: a Hoosier Oncology Group study'. *Am J Clin Oncol.* 34(5), pp. 460–465.

Casadei, R. *et al.* (2015). 'Neoadjuvant chemoradiotherapy and surgery versus surgery alone in resectable pancreatic cancer: a single-center prospective, randomized, controlled trial which failed to achieve accrual targets'. *Journal of Gastrointestinal Surgery*, 19(10), pp. 1802–1812

Catto, J.W.F. *et al.* (2003). 'Artificial intelligence in predicting bladder cancer outcome: a comparison of neuro-fuzzy modelling and artificial neural netwroks'. *Clinical Cancer Research,* 9, pp. 4172-4177.

Catto, J.W.F. *et al.* (2006). 'Neuro-Fuzzy Modelling: an accurate and interpretable method for predicting bladder cancer progression'. *The Journal of Urology,* 175, pp. 474-479.

Catto, J.W.F. *et al.* (2009). 'Neurofuzzy modelling to determine recurrence risk following radical cystectomy for nonmetastatic urothelial carcinoma of the bladder'. *Clinical Cancer Research,* 15(9), pp. 3150-3155

Centre for Evidence Based Medicine, Howick, J. *et al.* (2011). 'Explanation of the 2011 Oxford Centre for vidence-Based Medicine (OCEBM) Levels of Evidence (Background Document.' Oxford Centre for Evidence Based Medicine (OCEBM). https://www.cebm.net/index.aspx?o=5653 (Downloaded 18th June 2018).

Chaimani, A. *et al.* (2013). 'Graphical tools for network meta-analysis in STATA'. *PLoS One.* 8, e76654.

Cetin, V. *et al*. (2013). 'A phase II trial of cetuximab, gemcitabine, 5-fluorouracil, and radiation therapy in locally advanced nonmetastatic pancreatic adenocarcinoma'. *Gastrointestinal Cancer Research,* 6:4(1), pp. S2-S9.

Chakraborty, S. *et al.* (2014). 'Accelerated fraction radiotherapy with Capecitabine as neoadjuvant therapy for borderline resectable pancreatic cancer'. *Gastrointestinal Cancer Research*, 7, pp. 15–22

Chakravarty, K.D. *et al*. (2010) 'Prognosis and feasibility of en-bloc vascular resection in stage II pancreatic adenocarcinoma'. *World J Gastroenterol,* 16, pp. 997–1002.

Checkland, P. *Systems Thinking, Systems Practice*; Wiley: Chichester, UK, 1981.

Checkland, P. (1987). 'The application of systems thinking in real-world problem-situations: The emergence of soft systems methodology'. In Jackson, M. and Keys, P. (eds). *New Directions in Management Science*. Brookfield, VT, USA: Gower Publishing. pp. 87–96.

Checkland, P. (1999). *Systems Thinking, Systems Practice*. New York, NY, USA: John Wiley and Sons Inc.

Cheeseman, P. (1985). 'In defense of probability' In: Proc. IJCAI-85, Los Angeles, CA, pp. 1002–1009

Cheeseman, P. (1986). 'Probabilistic versus fuzzy reasoning' In: Kanal, L. and Lemmer, J. (eds.). *Uncertainty in Artificial Intelligence*. Amsterdam: North-Holland: pp. 85–102.

Chen, Y.F. *et al.* (2006). **'**A systematic review of the effectiveness of adalimumab, etanercept and infliximab for the treatment of rheumatoid arthritis in adults and an economic evaluation of their cost-effectiveness'. *Health Technol Assess,* 10, pp. 1-229

Chen, X. *et al.* (2017). 'Neoadjuvant radiation followed by resection versus upfront resection for locally advanced pancreatic cancer patients: a propensity score matched analysis'. *Oncotarget*, 8(29), pp. 47831-47840.

Cheng, J., Bell, D. and Liu, W. (1007). 'Learning Bayesian networks from data: an efficient approach based on information theory'. *Sixth International Conference on Information and Knowledge Management,* pp. 325-331.

Choi, J.G. *et al*. (2018). 'Neoadjuvant FOLFIRINOX for patients with borderline resectable or locally advanced pancreatic cancer: results of a decision analysis.' *The Oncologist,* 23, pp. 1-10.

Christou, N.V. (1994). 'Evaluation of operative risk'. In: Wilmore, D.W. (ed). *American College of Surgeons: Care of the Surgical Patient.* New York, NY: Scientific American, Inc.

Chua, T.C. and Saxena, A. (2010). 'Extended pancreaticoduodenectomy with vascular resection for pancreatic cancer: a systematic review'. *J Gastrointest Surg*, 14, pp. 1442–52.

Chua, T.C. and Saxena, A. (2011). 'Preoperative chemoradiation followed by surgical resection for resectable pancreatic cancer: a review of current results'. *Surg Oncol.* 20,, pp. 161–168.

Cilliers , P. (1998). *Complexity and postmodernism: understanding complex systems*. London, UK,: Routledge

Cilliers, P. (2005a). 'Knowledge, limits and boundaries'. *Futures*, 37, pp. 605–613.

Cilliers, P. (2005b). 'Complexity, deconstruction and relativism'. *Theory Cult. Soc,* 22, pp. 255–267

Cilliers, P. (2007). 'The philosophical importance of thinking complexity'. In Cilliers, P. (ed). *Thinking Complexity: Complexity and Philosophy*. Mansfield, TX, USA: ICSE Publishers:, pp. 3–6.

Cilliers, P. (2008). 'Complexity theory as a general framework for sustainability science'. In Burns, M. and Weaver, A. (eds). *Exploring Sustainable Science: A South African Perspecfive*. Stellenbosch, South Africa: African Sun Media. pp. 39–57.

Cilliers, P. (2010). 'The value of complexity. A response to Elizabeth Mowat and Brent Davis'. *Complicity Int. J. Complex. Educ*, 7, pp. 39-42

Clark, J. and McGee-Lennon, M. (2011). 'A stakeholder-centred exploration of the current barriers to the uptake of home care technology in the UK'. *J Assist Technol*, 5(1), pp. 12–25.

Claxton, K., Sculpher, M. and Drummond, M. (2002). 'A rational framework for decision making by the National Institute for Clinical Excellence'. *Lancet*, 360, pp. 711-5.

Cockburn, G., and Tesfamariam, S. (2012). 'Earthquake disaster risk index for Canadian cities using Bayesian belief networks'. *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards,* 6 (2), pp. 128-140

Cohn, S. *et al.* (2013). 'Entangled complexity: why complex interventions are just not complicated enough'. *J Health Serv Res Policy.* 18(1), pp. 40–3.

Coletti, G., Gilio, A. and Scozzafava, R. (1993). 'Comparative probability for conditional events: A new look through coherence'. *Theory and Decision,* 35, pp. 237–258.

Collins, G.S. *et al.* (2015). 'Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): The TRIPOD Statement.' *Ann Intern Med,* 162, pp. 55–63.

Collisson, E. A. *et al*. (2011). 'Subtypes of pancreatic ductal adenocarcinoma and their differing responses to therapy'. *Nat. Med.* 17, pp. 500–503.

Collisson, E.A. *et al*. (2019). 'Molecular subtypes of pancreatic cancer'. *Nature Reviews Gastroenterology & Hepatology 1*6, pp. 207–220

Conroy, T. F. *et al.* (2011). 'FOLFIRINOX versus gemcitabine for metastatic pancreatic cancer'. *NEJM,* 364(19), pp. 1817-1825

Conroy, T. *et al.* (2018). 'FOLFIRINOX or Gemcitabine as adjuvant therapy for pancreatic cancer'. *NEJM,* 379(25), pp. 2395-2406.

Coolsen, M.M. *et al*. (2013). 'Systematic review and meta-analysis of enhanced recovery after pancreatic surgery with particular emphasis on pancreaticoduodenectomies'. *World J Surg*, 37, pp. 1909–18.

Cooper, N.J., Sutton, A.J. and Abrams, K. (2002).'Decision analytical economc modeling within a Bayesian framework: application to prophylactic antibiotics use for cesarean section'. *Stat Methods Med Res,* 11, pp. 491-512.

Cooper, N.J. *et al*. (2004). 'Comprehensive decision analytic modeling in economic evaluation: a Bayesian approach'. *Health Econ,* 13, pp. 203-26.

Cox, D.R. (1972). 'Regression models and life-tables'. *J R Stat Soc B Stat Methodol* 34(2), pp.187-220.

Cox, D.R. and Oakes, D. (1984). *Analysis of Survival Data*. New York, NY: Chapman & Hall.

Crane, C.H. *et al.* (2011). 'Phase II trial of cetuximab, gemcitabine, and oxaliplatin followed by chemoradiation with cetuximab for locally advanced (T4) pancreatic adenocarcinoma: correlation of Smad4(Dpc4)Immunostaining with pattern of disease progression'. *Journal of Clinical Oncology*, 29(22), pp. 3037-3043.

Currie, G. *et al*. (2012). 'Institutional work to maintain professional power: Recreating the model of medical professionalism'. *Organ Stud*. 33(7), pp. 937–62.

D'Angelo, F, L. *et al.* (2017). 'Neoadjuvant treatment in pancreatic cancer: evidence-based medicine? A systematic review and meta-analysis'. *Med Oncol,* 34(5), pp. 85

Danese, M.D. (2008). 'Budget impact model of adding erlotinib to a regimen of gemcitabine for the treatment of locally advanced, nonresectable or metastatic pancreatic cancer'. *Clin Ther*, 30, pp. 775–784.

Dasari, B.V.M. *et al.* (2016). 'A model to predict survival following pancreaticoduodenectomy for malignancy based on tumour site, stage and lymph node ratio'. *HPB.* 18(4), pp. 332-338.

de Castro, S.M. *et al.* (2009). 'Evaluation of POSSUM for patients undergoing pancreatoduodenectomy'. *World J Surg*, 33(7), pp. 1481-7.

de Felice, F. *et al.* (2014). 'Neoadjuvant strategy as initial treatment in resectable pancreatic cancer: concrete evidence of benefit'. *Anticancer Res.* 34**,** pp. 4673–6.

deGeus, S.W. et al. (2016). 'Neoadjuvant therapy versus upfront surgical strategies in resectable pancreatic cancer: a Markov decision analysis'. *Eur J Surg*. 42, pp. 1552-1560.

de Gus, S.W.L. *et al*. (2017a). 'Neoadjuvant therapy affects margins and margins affects all: perioperative and survival outcomes in resected pancreatic adenocarcinoma'. *HPB*. 20, pp. 573-581.

de Gus, S.W.L. *et al*. (2017b). 'Neoadjuvant therapy versus upfront surgery for resected pancreatic adenocarcinoma: A nationwide propensity score matched analysis'. *Surgery.* 161(3), pp. 592-601.

de Santiago, I. *et al.* (2017). 'Immuno-phenotypes of pancreatic ductal adenocarcinoma: metaanalysis of transcriptional subtypes'. Preprint Available at: https://www.biorxiv.org/content/ early/2017/10/05/198903 (Downlaoded May 2018).

De Villiers-Botha, T. and Cilliers, P. 'The complex "I": The formation of identity in complex systems'. In Cilliers, P. and Preiser, R. (eds). *Complexity, Difference and Identity*. Dordrecht, The Netherlands: Springer. pp. 19–38.

Delgado, G. *et al.* (1989). 'A prospective surgical pathological staging of stage I squamous carcinoma of the cervix: a Gynecologic Oncology Group Study'. *Gynecol Oncol*, 35, pp. 314

Deo, R.C. (2015). 'Machine learning in medicine'. *Circulation*. 132, pp. 1920-1930.

Dempster, A. (1967). 'Upper and lower probabilities induced by a multivalued mapping'. *Ann. Math. Statist*. 38, pp. 325–339.

Depaoli, S. (2013). 'Mixture class recovery in GMM under varying degrees of class separation: Frequentist versus Bayesian estimation'. *Psychological Methods* 18(2), pp.186-219.

Department of Health (DoH). (2015). *2010 to 2015 Government policy: cancer research and treatmen.* London: DoH.

Dewdney, S.B. *et al.* (2010). 'The role of neoadjuvant chemotherapy in the management of patients with advanced stage ovarian cancer: survey results from members of the Society of Gynecologic Oncologists'. *Gynecol Oncol,* 119, pp. 18-21.

Diaconis, P. and Zabell, S. (1983). 'Some alternatives to Bayes's rule'. In *Information Pooling and Group Decision Making: Proceedings of the Second University of California, Irvine, Conference on Political Economy,* pp. 25–38.

Dias, S. *et al*. (2013a). 'Evidence synthesis for decision-making 2: a generalized linear modeling framework for pairwise and network meta-analysis of randomized controlled trials'. *Med Decis Making.* 33**,** 607–617.

Dias, S. *et al*. (2013b). 'Evidence synthesis for decision making 4: inconsistency in networks of evidence based on randomized controlled trials'. *Med Decis Making.* **33,** pp. 641–656.

Dindo, D., Demartines, N. and Clavien, P.A. (2004). 'Classification of surgical complications: a new proposal with evaluation in a cohort of 6336 patients and results of a survey'. *Ann Surg*, 240(2), pp. 205-13

Donders, A.R. *et al*. (2006). 'Review: a gentle introduction to imputation of missing values'. *J Clin Epidemiol*, 59, pp. 1087–1091.

Drummond, M.F. *et al.* (1997). 'Users' guide to the medical literature, XIII: How to use an article on economic analysis of clinical practice.

Are the results of the study valid? Evidence-based Medicine Working Group'. *JAMA,* 277, pp. 1552-1557.

Drummond, M.F. *et al.* (2015). *Methods for the economic evaluation of health care programmes*. 4th Ed. London: Oxford University Press.

Du ,W. *et al.* (2000). Cost analysis of pancreatic carcinoma treatment. *Cancer*. 89, pp. 1917–1924

Dudley, W.N., Wickham, R. and Coombs, N. (2016). An introduction to survival statistics: Kaplan-Meier analysis. *J Adv Oncol.* 7(1), pp. 91-100.

Dweiri, F.T. and Kablan, M.M. (2006). 'Using fuzzy decision making for the evaluation of the project management internal efficiency'. *Decis Support Syst*. 42, pp. 712-726.

Dzau, V.J. and Ginsburg, G.S. (2016). 'Realizing the Full Potential of Precision Medicine in Health and Health Care'*. JAMA*. 316(16), pp. 1659–1660.

Earman, J. (1992). *Bayes or bust? Critical account of Bayesian confirmation theory.* Cambridge: MIT Press.

Eddy, G.L. *et al*. (1995). 'Neoadjuvant chemotherapy with vincristine and cisplatin followed by radical hysterectomy and pelvic lymphadenectomy for FIGO stage IB bulky cervical cancer: a Gynecologic Oncology Group pilot study'. *Gynecol Oncol* 57, pp. 412-6.

Eden, C. (1987). 'Problem-solving or problem-finishing'. In Jackson, M. and Keys, M.P. (eds) *New Directions in Management Science*. Brookfield, VT, USA: Gower Publishing. pp. 97–107.

Eldabi, T. (2009). 'Implementation issues of modeling healthcare problems: misconceptions and lessons'. *Proceedings of the 2009 winter simulation conference (WSC). IEEE*. Availbale at: https://doi.org/10.1109/WSC.2009.5429192. (Downloaded 15th April, 2018)

Elixhauser, A. and Halpern, M.T. (1999). 'Economic evaluations of gastric and pancreatic cancer'. *Hepatogastroenterology.* 46, pp. 1206–1213.

Enestvedt, C.K. *et al*. (2008). 'Diagnostic laparoscopy for patients with potentially resectable pancreatic adenocarcinoma: is it cost-effective in the current era?' *J Gastrointest Surg*. 12, pp. 1177–1184.

Epelbaum, R. et al. (2002). 'Phase II study of gemcitabine combined with radiation therapy in patients with localized, unresectable pancreatic cancer'. *J Surg Oncol*. 81, pp. 138-143.

Ercolani, G. *et al.* (2011). 'Effectiveness and cost-effectiveness of peri-operative versus post-operative chemotherapy for resectable colorectal liver metastases'. *Eur J of Cancer,* 47(15), pp. 2291-2298.

Eshuis, W.J. *et al.* (2015). 'Gastric emptying and quality of life after Pancreatoduodenectomy with retrocolic or ante-colic gastroenteric anastomosis'. *Br J Surg*. 102, pp. 1123-32.

Esnaola, N.F. *et al*. (2014). 'Phase 2 trial of induction gemcitabine, oxaliplatin, and cetuximab followed by selective capecitabine-based chemoradiation in patients with borderline resectable or unresectable locally advanced pancreatic cancer'. *Int J Radiat Oncol Biol Phys*. 88, pp. 837–844.

Estrella, J.S. *et al*. (2012). 'Post-therapy pathologic stage and survival in patients with pancreatic ductal adenocarcinoma treated with neoadjuvant chemoradiation'. *Cancer*, 118, pp. 268-77.

Esposito, I. *et al.* (2014). 'Pathology of pancreatic ductal adenocarcinoma: facts, challenges and future developments.' *World J Gastroenterol*. 20, pp. 13833–13841

Evans, D.B. *et al*. (1992). 'Preoperative chemoradiation and pancreaticoduodenectomy for adenocarcinoma of the pancreas'. *Arch Surg*, 127, pp. 1335-9.

Evans, D.B. *et al.* (2008). 'Preoperative gemcitabine-based chemoradiation for patients with resectable adenocarcinoma of the pancreatic head'. *J Clin Oncol*, 26, pp. 3496-3502

Evans, D.B., Erickson, B.A. and Ritch, P. (2010). 'Borderline resectable pancreatic cancer: definitions and the importance of multimodality therapy.' *Ann Surg Oncol*. 17, pp. 2803-2805

Evans, D.B. (2018). 'What Makes a Pancreatic Cancer Resectable?' *American Society of Clinical Oncology Educational Book.* 38, pp. 300-305

Faith, A. *et al*. (2015). 'Survival of patients with resectable pancreatic cancer who received neoadjuvant therapy'. *Surgery*. 6, pp: 418–29.

Faltinsen, E.G. *et al.* (2018). 'Network meta-analysis: the highest level of medical evidence?' *BMJ Evidence-Based Medicine.* Published Online First: 14 March2018. doi: 10.1136/bmjebm-2017-110887

Falzon, L. (2006). 'Using Bayesian network analysis to support centre of gravity analysis in military planning*.' European Journal of operational Research.* 170, pp. 629-643

Feldmann, G. *et al*. (2007). 'Molecular genetics of pancreatic intraepithelial neoplasia'. *J Hepatobiliary Pancreat Surg*. 14, pp. 224–232.

Felli, J.C. and Hazen, G. (1999). 'A Bayesian approach to sensitivity analysis'. *Health Econ,* 8, pp. 263-8.

Fenton, N., and Neil, M. (2019). *Risk assessment and decision analysis with Bayesian networks.* 2nd ed. London: CRC Press.

Ferlay, J. et al. (2013). 'Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012'. *Eur J Cancer* 49, pp. 1374-403

Ferrone, C.R. *et al.* (2005). 'Validation of a postresection pancreatic adenocarcinoma nomogram for disease-specific survival'. *J Clin Oncol* 23(30), pp. 7529-35.

Fiore, M. *et al*. (2017). 'Phase II study of induction chemotherapy followed by chemoradiotherapy in patients with borderline resectable and unresectable locally advanced pancreatic cancer'. *Scientific Reports*. 7, pp. 45845.

Flyvbjerg, B. (2006). 'Five misunderstandings about case-study research'. *Qual Inq*. 12(2), pp. 219-45.

Fone, D. *et al*. (2003). 'Systematic review of the use and value of computer simulation modelling in population health and health care delivery'. *J Public Health Med*. 25(4), pp. 325–35.

Fong, Y. *et al*. (1999). 'Clinical score for predicting recurrence after hepatic resection for metastatic colorectal cancer: analysis of 1001 consecutive cases'. *Ann Surg*, 230(3), pp. 309-18

Fraser, S.W. and Greenhalgh, T. (2001). 'Complexity science- coping with complexity: educating for capability'*. BMJ*. 323(7316), pp. 799-803

Friend, J. (2001). 'The strategic choice approach'. In Rosenhead, J. and Mingers, J. (eds). *Rational Analysis for a Problematic World Revisited*. Chichester, UK: Wiley.

Freeman, M.F. and Tukey, J.W. (1950). 'Transformations related to the angular and the square root'. *The Annals of Mathematical Statistics*. 21,pp. 607-11.

Fuhrman, G.M. *et al*. (1996). 'Rationale for en bloc vein resection in the treatment of pancreatic adenocarcinoma adherent to the superior mesenteric-portal vein confluence. Pancreatic Tumor Study Group'**.** *Ann Surg*. 223, pp. 154–162

Fujii,T. *et al.* (2015). 'Inverse probability of treatment weighting analysis of upfront surgery versus neoadjuvant chemoradiotherapy followed by surgery for pancreatic adenocarcinoma with arterial abutment'. *Medicine.* 94(39), e1647.

Fujii, T. *et al.* (2017). 'Clinical benefits of neoadjuvant chemoradiotherapy for adenocarcinoma of the pancreatic head: an observational study using inverse probability of treatment weighting'. *J Gastroenterol.* 52, pp. 81–93.

Gall, T.M.H. *et al.* (2014). 'Surgical techniques for improving outcomes in pancreatic ductal adenocarcinoma'. *Expert Rev Gastroenterol Hepatol* 8, pp. 241–6.

Gall, T.M.H. *et al*. (2015). 'Pancreatic cancer: current management and treatment strategies'. *Postgrad Med J*, 91, pp. 601-607

Gallo, G. (2004). 'Operations research and ethics: Responsibility, sharing and cooperation'. *Eur. J. Oper. Res. 153* pp. 468–476.

Garas, G. *et al.* (2012). 'Evidence-based surgery: barriers, solutions, and the role of evidence synthesis'. *World J Surg*, 36, pp. 1723-1731

Garber, S. (2014). 'Redirecting innovation in U.S. health care: options to decrease spending and increase value'. Santa Monica, CA: RAND Corporation.

Garcea, G. *et al.* (2012). 'Improving the diagnostic yield from staging laparoscopy for periampullary malignancies: the value of preoperative inflammatory markers and radiological tumor size'. *P*ancreas, 41(2), pp. 233-237.

Gawande, A.A. *et al*. (2007). 'An Apgar score for surgery'. *J Am Coll Surg* 204(2), pp. 201-8.

Gelman, A. (2005). 'Analysis of variance – Why it is more important than ever'. *Annals of Statistics* 33, pp. 1–53.

Gelman, A. and Shakizi, C. (2013). 'Philosophy and the practice of Bayesian statistics'. *British Journal of Mathematical and Statistical Psychology*: 66, pp. 8-38.

Ghaneh, P. *et al.* (2018). 'PET-PANC: multicentre prospective diagnostic accuracy and health economic analysis study of the impact of combined modality 18fluorine-2-fluoro-2-deoxy-d-glucose positron emission tomography with computed tomography scanning in the diagnosis and management of pancreatic cancer'. *Health Technol Assess* 22(7), pp 1362-71

Ghavidel, A.A. *et al.* (2014). 'Two new mathematical models for prediction of early mortality risk in coronary artery bypass graft surgery'. *The Journal of Thoracic Cardiovascular Surgery*. 148, pp. 1291-1298.

Ghosn, M. *et al.* (2016). 'Where does chemotherapy stands in the treatment of ampullary carcinoma? A review of literature.' *World J Gastrointest Oncol.* 8, pp. 745–750.

Gillen, S. *et al*. (2010). 'Preoperative/neoadjuvant therapy in pancreatic cancer: a systematic review and meta-analysis of response and resection percentages.' *PLoS Med.* **7**, e1000267.

Gillen, S.T. et al. (2010). 'Preoperative/neoadjuvant therapy in pancreatic cancer: a systematic review and meta-analysis of response and resection percentages'. *PLoS Med,* 7(4), e1000267

Glimelius, B. *et al.* (1995). 'Cost-effectiveness of palliative chemotherapy in advanced gastrointestinal cancer'. *Ann Oncol.* 6, pp. 267–274.

Glannl, L. *et al.* 'Efficacy and safety of neoadjuvant pertuzumab and trastuzumab in women with locally advanced inflammatory, or early HER-2 positive breast cancer (NeoSphere): a randomized multicenter, open-label, phase 2 trial'. *Lancet Oncol,* 13, pp.25-32.

Goffi, L. *et al*. (1999). 'Preoperative APACHE II and ASA scores in patients having major general surgical operations: prognostic value and potential clinical applications'. *Eur J Surg*, 165(8), pp. 730-5.

Golcher, H. *et al*. (2008). 'Preoperative chemoradiation in adenocarcinoma of the pancreas. A single centre experience advocating a new treatment strategy'. *European journal of surgical oncology: the journal of the European Society of Surgical Oncology and the British Association of Surgical Oncology*. 34(7), pp. 756-764.

Golcher, H. *et al.* (2015). 'Neoadjuvant chemoradiation therapy with gemcitabine/cisplatin and surgery versus immediate surgery in resectable pancreatic cancer: Results of the first prospective randomized phase II trial'. *Strahlentherapie Und Onkologie*. 191, pp. 7-16.

Gold, M.R. *et al.* (eds) (1996). *Cost-Effectiveness in Health and Medicine*. New York, NY: Oxford University Press.

Goldman, L. et al. (1977). 'Multifactorial index of cardiac risk in noncardiac surgical procedures'. *NEJM*, 297(16), pp. 845-50.

Gordon, L.G*. et al.* (2012). 'Modeling the cost-effectiveness of strategies for treating esophageal adenocarcinoma and high-grade dysplasia'. *Journal of Gastrointestinal Surgery:* 16, pp. 1451-1461.

Greenberg, D. *et al.* (2010). 'When is cancer care cost-effective? A systematic overview of cost-utility analyses in oncology'. *Journal National Cancer Institute:* 102, pp. 82-88.

Greenhalgh, T. and Papoutsi, C. (2018). 'Studying complexity in health services research: desperately seeking an overdue paradigm shift.' *BMC Medicine*, 16 pp. 95

Greenberg, D. *et al*. (2010). 'When is cancer care cost-effective? A systematic overview of cost-utility analyses in oncology'. *Journal National Cancer Institute,* 102, pp. 82-88.

Greenhalgh, T. *et al*. (2016).' SCALS: a fourth-generation study of assisted living technologies in their organisational, social, political and policy context'. *BMJ Open.* 15;6(2), pp. e010208. doi: 10.1136/bmjopen-2015-010208.

Greenhalgh, T. *et al*. (2017a). 'Beyond Adoption: A New Framework for Theorizing and Evaluating Nonadoption, Abandonment, and Challenges to the Scale-Up, Spread, and Sustainability of Health and Care Technologies'. *J Med Internet Res*. 19(11), pp. e367. doi:10.2196/jmir.8775

Greenhalgh, T., A'Court, C. and Shaw, S. (2017b). 'Understanding heart failure; explaining telehealth - a hermeneutic systematic review'. *BMC Cardiovasc Disord*. 14;17(1):, pp. 56.

Greenblatt, D.Y. *et al*. (2011). 'Preoperative factors predict perioperative morbidity and mortality after pancreaticoduodenectomy'. *Ann Surg Oncol* 18(8), pp. 2126-35

Grin, J., Rotmans, J. and Schot, J. (2010). *Transitions to Sustainable Development: New Directions in the Study of Long Term Transformative Change.* New York, NY: Routledge.

Grizzi, F. and Chiriva-Internati, M. (2006). 'Cancer: looking for simplicity and finding complexity'. *Cancer Cell International*, 6(4), 1-7

Grizzi, F. *et al*. (2004). 'Complexity and cancer'. *Gastroenterology,* 126, pp. 630–631. doi: 10.1053/j.gastro.2003.12.033.

Grobmyer, S.R. *et al.* (2007). 'Defining morbidity after pancreaticoduodenectomy: use of a prospective complication grading system'. *J Am Coll Surg*, 204:356–64.

Grose, D. *et al*. (2017). 'The role of induction chemotherapy + chemoradiotherapy in localised pancreatic cancer: initial experience in Scotland'. *J Gastrointest Oncol*. 8(4), pp. 683-695.

Grosof, B. (1986). 'An inequality paradigm for probabilistic knowledge' in: Kanal, L and Lemmer, J. (eds.) *Uncertainty in Artificial Intelligence*. Amsterdam: North-Holland, pp. 259–275.

Grossi, E. (2006). 'How artificial intelligence tools can be used to assess individual patient risk in cardiovascular disease: problems with the current methods'. *BMC Cardiovascular Disorders* 6:20

Grossi, E. (2015). 'Medical concepts related to individual risk are better explained with "plausibility" rather than "probability". '*BMC Cardiovascular Disorders*, 5:31

Gubanski, M. *et al.* (2010). 'Randomized phase II study of sequential docetaxel and irinotecan with 5-fluorouracil/ folinic acid (leucovorin) in patients with advanced gastric cancer: The GATAC trial.' *Gastric Cancer,* 13, pp. 155-161.

Gudjonsson, B. (1995). 'Carcinoma of the pancreas: critical analysis of costs, results of resections, and the need for standardized reporting.' *J Am Coll Surg,* 181, pp. 483–503.

Gursel, G. (2016). 'Healthcare, uncertainty, and fuzzy logic'. *Dig Med*. 2, pp. 101-112

Guyatt, G.H. *et al*. 'GRADE: an emerging consensus on rating quality of evidence and strength of recommendations'. *BMJ.* 336**,** pp. 924.

Haack, S. (1979). 'Do we need ``fuzzy logic''? ' *Int J Man-Machine Stud*, 11, pp. 437-45.

Haack, S. (1980). 'Is truth at or bumpy?' In: Mellor, D.H. (ed). *Prospects for pragmatism: essays in memory of FP Ramsey*. Cambridge: Cambridge University Press, pp. 1-20.

Hackert, T. and Buchler, M.W. (2014). 'Invited commentary on "Evaluation of a predictive model for pancreatic fistula based on amylase value in drains after pancreatic resection" by Partelli et al.' *Am J Surg.* 208(4), pp. 640-1

Hackert, T. *et al.* 'Locally advanced pancreatic cancer: neoadjuvant therapy with Folfirinox results in resectability in 60% of the patients.' *Ann. Surg.* 264, pp. 457–463

Hager, D. and Andersen, L.B. (2010). 'A knowledge based approach to loss severity assessment in financial institutions using Bayesian networks and loss determinants'. *European Journal of Operational Research.* 207(3), pp. 1635-1644

Hampson, L.V. *et al.* (2014). 'Bayesian methods for the design and interpretation of clinical trials in very rare diseases'. *Statistics in Medicine,* 33, pp. 4186-4201.

Hansen, K. and Singer, D.B. (2001). 'Asplenic-hyposplenic overwhelming sepsis: postsplenectomy sepsis revisited'. *Pediatr Dev Pathol,* 4, pp. 105–21.

Harper, P. *et al*. (2003). 'A systems modelling approach for the prevention and treatment of diabetic retinopathy'. *European Journal of Operational Research.* 150(1), pp. 81-91.

Harrell, F.E., Jr., Lee, K.L. and Mark, D.B. (1996). 'Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors'. *Stat Med* 15, pp. 361–387.

Harrell, F.E. (2001) 'Regression modeling strategies with applications to linear models, logistic regression and survival analysis'. New York: Springer Verlag.

Hartley, M.N. and Sagar, P.M. (1994). 'The surgeon's "gut feeling" as a predictor of post-operative outcome'. *Ann R Coll Surg Engl* 76(6), pp. 277-8.

Hartwig, W. e al. (2013). 'Improvement of surgical results for pancreatic cancer'. *Lancet Oncol.* 14, pp476–485

Harvey, K.B. *et al*. (1981). 'Biological measures for the formulation of a hospital prognostic index'. *Am J Clin Nutr*, 34(10), pp. 2013-22.

Hashimoto, D.A. et al. (2018). 'Artificial Intelligence in Surgery: promises and perils'. *Annals of Surgery*. 286, pp. 70-76.

Hayward, J. *et al.* (2010). 'Machine learning of clinical performance in a pancreatic cancer database'. *Artificial Intelligence*. 49, pp. 187-195.

Heckerman, D. (1999). 'Bayesian learning,' in: Wilson, R and Kell, F. (eds.). *The MIT Encyclopedia of the Cognitive Sciences*. Cambridge, MA: MIT Press, pp. 70–72.

Heinrich, S. *et al*. (2005). 'Positron emission tomography/ computed tomography influences on the management of resectable pancreatic cancer and its cost-effectiveness'. *Annals of Surgery*, 242, pp. 235-243.

Heinrich, S. *et al.* (2008). 'Prospective phase II trial of neoadjuvant chemotherapy with gemcitabine and cisplatin for resectable

adenocarcinoma of the pancreatic head'. *J Clin Oncol*. 26, pp. 2526-2531.

Hendry, D. Leamer, E. and Poirier, D. (1990). 'The ET dialogue: a conversation on economic methodology'. *Econometric Theory,* 6, pp. 171-261.

Henschke, U.K. and Flehinger, B.J. (1967). 'Decision theory in cancer therapy'. *Cancer*, 20(11), pp. 1819–1826.

Herman, J. *et al.* (2015). 'Phase 2 multi-institutional trial evaluating gemcitabine and stereotactic body radiotherapy for patients with locally advanced unresectable pancreatic adenocarcinoma'. *Cancer*. 121, pp. 1128–1137.

Higgins, J.P.T. *et al*. (2011). 'The Cochrane Collaboration's tool for assessing risk of bias in randomised trials.' *BMJ.* 343**,** pp. 5928

Hill, R.G. Jr, Sears, L.M., Melanson, S.W. (2013). '4000 clicks: a productivity analysis of electronic medical records in a community hospital ED'. *Am J Emerg Med*. 31(11), pp. 1591-1594.

Hill, J.S. *et al.* (2010). 'A simple risk score to predict in-hospital mortality after pancreatic resection for cancer'. *Ann Surg Oncol*, 17(7), pp. 1802-7.

Hirano, S. *et al*. (2007). 'Distal pancreatectomy with en bloc celiac axis resection for locally advanced pancreatic body cancer: long-term results'. Ann Surg, 246, pp. 46–51

Hirono, S*. et al.* (2016). 'Treatment Strategy for Borderline Resectable Pancreatic Cancer With Radiographic Artery Involvement.' *Pancreas.* 45(10), pp. 1438–1446.

Hishinuma, S. *et al*. (2006).' Patterns of recurrence after curative resection of pancreatic cancer, based on autopsy findings'. *J Gastrointest Surg*, 10(4), pp. 511-8.

Hoffman, J.P. *et al*. (1998). 'Phase II trial of preoperative radiation therapy and chemotherapy for patients with localized, resectable adenocarcinoma of the pancreas: an Eastern Cooperative Oncology Group Study'. *J Clin Oncol*. 16, pp. 317-23.

Holbrook, R.F., Hargrave, K. and Traverso, L.W. (1996). 'A prospective cost analysis of pancreatoduodenectomy'. *Am J Surg*. 171, pp. 508–511.

Holland, J.H. (2014). *Complexity: a very short introduction*. Oxford: Oxford University Press.

Hong, S.M. *et al.* (2011). 'Molecular signatures of pancreatic cancer'. *Arch Pathol Lab Med*. 135, pp. 716–727.

Hong, T.S. *et al*. (2014). 'A Phase 1/2 and biomarker study of preoperative short course chemoradiation with proton beam therapy and capecitabine followed by early surgery for resectable pancreatic ductal adenocarcinoma'. *Int. J. Radiat. Oncol. Biol. Phys.* 89(4), pp. 830–838.

Hosein, P.J. *et al.* (2012). 'A retrospective study of neoadjuvant FOLFIRINOX in unresectable or borderline-resectable locally advanced pancreatic adenocarcinoma'.
BMC Cancer, 12, pp. 199.

Hosmer, D.W. and Lemeshow, S. (1989). *Applied Logistic Regression*. New York: Wiley.

Houghton, T. (2011). 'Does positivism really 'work' in social science?' *E-International Relations* [online] Available at: http://www.e9ir.info/2011/09/26/does-positivism-really--'work'-in-the-social-sciences/#_ftn4 [Accessed 20th December 2016]

Howard, T.J*. et al*. (2003). 'Efficacy of venous reconstruction in patients with adenocarcinoma of the pancreatic head.' *J Gastrointest Surg*. 7, pp. 1089–1095

Howard, T.J. *et al*. (2006). 'A margin-negative R0 resection accomplished with minimal postoperative complications is the surgeon's contribution to long-term survival in pancreatic cancer.' *J Gastrointest Surg.*10, pp. 1338–1345.

Hruban, R.H., Maitra, A. and Goggins, M. (2008). 'Update on pancreatic intraepithelial neoplasia'*. Int J Clin Exp Pathol*. 1, pp. 306–316.

Hsu, C.C. *et al.* (2012). 'Early mortality risk score: identification of poor otcomes following upfront surgery for resectable pancreatic cancer'. *J Gastrointest Surg,* 16(4), pp. 753-761

Huang, J. *et al*. (2018). 'Cost-effectiveness analysis of adjuvant treatment for resected pancreatic cancer in China based on the ESPAC-4 trial'. *Cancer Management and Research*, 10, pp. 4065-4072

Hultman, B. *et al.* (2012). 'Cost and clinical outcome of neoadjuvant systemic chemotherapy followed by cytoreductive surgery and hyperthermic intraperitoneal chemotherapy in peritoneal carcinomatosis from gastric cancer'. *Acata Oncologica:* 51(1), pp. 112-121.

Husereau, D. *et al.* (2013). Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement. *BMJ,* 346, pp. 1049

Hutton, B. *et al.* (2014). 'The quality of reporting methods and results in network meta-analyses: an overview of reviews and suggestions for improvement.' *PLoS One.* 9, pp. 92508.

Iacobuzio-Donahue, C.A. *et al.* (2009). 'DPC4 gene status of the primary carcinoma correlates with patterns of failure in patients with pancreatic cancer.' *J Clin Oncol.* 10: 27(11), pp. 1806-13.

Iacono, C. *et al*. (2013). 'Systematic review of central pancreatectomy and meta-analysis of central versus distal pancreatectomy'. *Br J Surg*, 100, pp. 873–85

Ikeda, M. T. *et al.* (2013). 'A multicenter phase II trial of S-1 with concurrent radiation therapy for locally advanced pancreatic cancer' *Int J Radiat Oncol Biol Phys*, 85 (1), pp. 163-169

Ielpo, B*, et al.* (2016). 'Preoperative treatment with gemcitabine plus nab-paclitaxel is a safe and effective chemotherapy for pancreatic adenocarcinoma'. *Eur J Surg Oncol*. 42, pp. 1394–1400.

Ilic, M. and Ilic, I. (2016). 'Epidemiology of pancreatic cancer'. *World J Gastroenterology* 22, pp. 9694-9705

Ingelfinger, F.J. (1975). 'Decision in medicine'. *NEJM* 293(5), pp. 254–255.

International Agency for Research on Cancer, World Health Organization. (2018). *Global Cancer Observatory 2018*. Available at: http://gco.iarc.fr/ (Downloaded 25th July 2019).

ISCRTN registry. (2014). 'ESPAC-5F: European Study group for Pancreatic Cancer – Trial 5F.' *BMC* Available at: www.isrctn.com/ISRCTN89500674. (Dowlnloaded 20th May 2018).

Ishii, H. *et al*. (2005). 'Treatment cost of pancreatic cancer in Japan: analysis of the difference after the introduction of gemcitabine'. *Jpn J Clin Oncol.* 35, pp. 526–530.

Ismail, M.A. *et al*. (2011). Developing a road performance index using a Bayesian belief network model. *Journal of the Franklin Institute*. 348 (9), pp. 2539-2555

Jackson, M.C and Keys, P. (1984) 'Towards a system of systems methodologies'. *J. Oper. Res. Soc. 35*, pp. 473–486.

Jackson, M. (1988). 'Some methodologies for community operational research'. *J. Oper. Res. Soc. 39*, pp. 715–724

Jager, K.J. *et al.* (2008). 'The analysis of survival data: the Kaplan-Meier method'. *Kidney International.* 74, pp. 560-565.

Jarnagin, W.R. (2011). 'Clinical scoring systems for stratifying risk after resection of hepatic colorectal metastases: still relevant?' *Ann Surg Oncol*, 18(10), pp. 2711-3.

Jang, J.Y. *et al.* (2018). 'Oncological benefits of neoadjuvant chemoradiation with gemcitabine versus upfront surgery in patients with borderline resectable pancreatic cancer: a prospective, randomized, open-label, multicenter phase 2/3 Trial'. *Annals of Surgery.* 268(2), pp. 215-222.

Janssen, Q. *et al.* (2018). 'Neoadjuvant FOLFIRINOX in patients with (borderline) resectable pancreatic cancer: a systematic and patent-level meta-analysis'. *Journal of Clinical Oncology.* 36(15), pp.16207-16207

Jayakrisshnan, T.T. *et al.* (2015). 'Diagnostic laparoscopy should be performed before definitive resection for pancreatic cancer: a financial argument'. *HPB* 17, pp. 131-139.

Jeffrey, R. (1065). *The Logic of Decision* New York: McGraw-Hill.

Jemal, A. *et al.* (2010).Cancer statistics, 2010. *Cancer Journal for Clinicians,* 60(5), pp. 277-300.

Jensen, E.H. *et al*. (2014). 'Neoadjuvant interferon-based chemoradiation for borderline resectable and locally advanced pancreas cancer: a phase II pilot study'. *HPB*.16(2), pp. 131-139.

Jeurnink, S.M. *et al.* (2010). 'Cost comparison of gastrojejunostomy versus duodenal stent placement for malignant gastric outlet obstruction'. *J Gastroenterol.* 45, pp. 537–543

Jing, W. *et al.* (2013). 'Distal pancreatectomy with en bloc celiac axis resection for the treatment of locally advanced pancreatic body and tail cancer'. *Hepatogastroenterology,* 60, pp. 187–90.

Jobanputra, P. *et al.* (2002). 'The effectiveness of infliximab and etanercept for the treatment of rheumatoid arthritis: a systematic review and economic evaluation'**.** *Health Technol Assess*. 6, pp. 1-110

Joensuu, T.K. *et al*. (2004). 'Phase I-II trial of twice-weekly gemcitabine and concomitant irradiation in patients undergoing pancreaticoduodenectomy with extended lymphadenectomy for locally advanced pancreatic cancer'. *Int J Radiat Oncol Biol Phys*. 60**,** pp. 444-452.

Joglekar, S. *et al.* (2015). 'Sarcopenia is an independent predictor of complications following pancreatectomy for adenocarcinoma'*. J. Surg. Oncol.* 111(6), pp. 771–775.

Johnson, S.R. *et al.* (2010). 'Methods to elicit beliefs for Bayesian prior: a systematic review'. *Journal of Clinical Epidemiology*. 63, pp. 355-369

Joyner, M.J. and Paneth, N. (2015). 'Seven Questions for Personalized Medicine'. *JAMA.* 314(10), pp. 999–1000.

Jun, J.B., Jacobson, S.H., Swisher, J.R. (1999). 'Application of discrete-event simulation in health care: A survey'. *J Opl Res Soc.* 50(2), pp. 109–123.

Junejo, M.A. *et al.* (2014). 'Cardiopulmonary exercise testing for preoperative risk assessment before pancreaticoduodenectomy for cancer'*. Ann. Surg. Oncol***.** 21(6), pp. 1929–1936.

Kabir, G. et al. (2015). 'Evaluating risk of water mains failure using a Bayesian belief network model'. *European Journal of Operatinal Research*. 240(1), pp. 220-234

Kalser, M. H. and Ellenberg, S. S. (1985).'Pancreatic cancer. adjuvant combined radiation and chemotherapy following curative resection'. *Arch. Surg.* 120, pp. 899–903.

Kamath, P.S. and Kim, W.R.. (2007). 'Advanced Liver Disease Study Group. The model for end-stage liver disease (MELD)'. *Hepatology*, 45(3), pp. 797-805.

Kanda, T., Jiang, X. and Yokosuka, O. (2014a). 'Androgen receptor signaling in hepatocellular carcinoma and pancreatic cancers'. *World J Gastroenterol.* 20, pp. 9229–9236

Kanda, M. *et al.* (2014b). 'The combination of the serum carbohydrate antigen 19-9 and carcinoembryonic antigen is a simple and accurate predictor of mortality in pancreatic cancer patients'. *Surg Today*, 44, pp. 1692-1701

Kaneoka, Y., Yamaguchi, A. and Isogai, M. (2009). 'Portal or superior mesenteric vein resection for pancreatic head adenocarcinoma: prognostic value of the length of venous resection'*. Surgery* 145, pp. 417–25.

Kapiteljn, E. *et al.* (1999). 'Total Mesorectal excision (TME) with or without preoperative radiotherapy in the treatment of primary rectal cancer: Prospective randomized trial with standard operative and histopathological techniques- Dutch Colorectal Cancer Group'. *Eur J Surg,* 165, pp. 410-420.

Kattan, M.W. *et al.* (2001). 'A postoperative prognostic nomogram for renal cell carcinoma'. *J Urol*, 166(1), pp. 63-7.

Katz, M.H. G. *et al.* (1012a). 'Response of borderline resectable pancreatic cancer to neoadjuvant therapy is not reflected by radiographic indicators'**.** *Cancer,* 118 (23), pp. 5749-5756

Katz, M.H.G. *et al.* (2012b). 'A clinical calculator of conditional survival estimates for resected and unresected pancreatic cancer survivors'. *Arch Surg*, 147(6), pp. 513-519

Katz, M.H.G. *et al.* (2016). 'Preoperative modified FOLFIRINOX treatment followed by capecitabine-based chemoradiation for borderline resectable pancreatic cancer: alliance for clinical trials in oncology trial A021101'. *JAMA Surg*, 151, 8

Katz, S.J., Jagsi, R. and Morrow, M. (2018). 'Reducing Overtreatment of Cancer With Precision Medicine Just What the Doctor Ordered'. *JAMA*, 319(11), pp. 1091–1092.

Kernick, D. (2006). 'Wanted – new methodologies for health service research. Is complexity theory the answer?' *Fam Pract*. 23(3), pp.385–90.

Kharofa, J. *et al*. (2014). 'Neoadjvuant chemoradiation with IMRT in resectable and borderline resectable pancreatic cancer.' *Radiother Oncol*. 113, pp. 41–6

Khoraa, A.A. *et al.* (2016). 'Potentially curable pancreatic cancer: American Society of Clinical Oncology Clinical Practice Guideline'. *J Clin Oncol.* 34(21), pp. 2541-2556.

Khorana A.A., et al. (2019). Potentally curable pancreatic adenocarcinoma: ASCO clinical practice guideline update. *J Clin Oncol*. doi: 10.1200/JCO.19.00946

Khuri, S.F. (2005). 'The NSQIP: A new frontier in surgery'. *Surgery*, 138(5), pp. 837-43

Kim, D.S. *et al*. (1988). 'Preoperative adjuvant chemotherapy in the treatment of cervical cancer stage Ib, IIa and IIb with bulky tumor'. *Gynecol Oncol*, 29, pp. 321-32.

Kim, E.J. *et al.* (2013a). 'A multi-institutional phase II study of neoadjuvant gemcitabine and oxaliplatin with radiation therapy in patients with pancreatic cancer'. *Cancer*, 119(15), pp. 2692-2700.

Kim, J.Y. *et al.* (2013b). 'A model for predicting pancreatic leakage after pancreaticoduodenectomy based on the international study group of pancreatic surgery classification'. *Korean Journal of Hepato-Biliary-Pancreatic Surgery*, 17(4), pp. 166-170.

Kim, S. Y. *et al.* (2013c). 'Single and multiple ability estimation in the SEM framework: a non-informative Bayesian estimation approach'. *Multivariate and Behavioral Research*, 48(4), pp. 563-591.

Kim, K.S. *et al.* (2017). 'Impact of Resection Margin Distance on Survival of Pancreatic Cancer: A Systematic Review and Meta-Analysis'. *Cancer Res Treat.* 49, pp. 824–833.

Kirchhof, P. and Meseth, N. (2012). 'A survey on the use of simulation in German healthcare'. *Proceedings of the 2012 Winter Simulation Conference.* Available at: https://doi.org/10.1109/WSC.2012.6465026. (Downloaded 5th May 2018).

Kleeff, J. *et al*. (2016). 'Pancreatic cancer'. *Nat. Rev. Dis. Primers* 2, 16022.

Klein, J.H. and Young, T. (2015). 'Health care: a case of hypercomplexity?' *Health Systems,* 4(2), pp. 104–10.

Klinkenbijl, J.H. *et al.* (1999). 'Adjuvant radiotherapy and 5-fluorouracil aftercurative resection of cancer of the pancreas and periampullary region: phase III trial of the EORTC gastrointestinal tract cancer cooperative group.' *Ann Surg*, 230, pp. 776-82

Kloppel, G. and Luttges, J. (2001). 'WHO-classification 2000: exocrine pancreatic tumors.' *Verh. Dtsch. Ges. Pathol.* 85, pp. 219–228.

Kobelt, G., Lindgren, P. and Geborek, P. (2009). 'Costs and outcomes for patients with rheumatoid arthritis treated with biological drugs in Sweden: a model based on registry data'**.** *Scand J Rheumatol*, 38, pp. 409-418

Kolmogorov, A.N. (1950). *Foundations of the Theory of Probability*. New York: Chelsea Publishing.

Kosuge, T. *et al.* (2006). 'A multicenter randomized controlled trial to evaluate the effect of adjuvant cisplatin and 5-fluorouracil therapy after curative resection in cases of pancreatic cancer'*. J Clin Oncol.* 36, pp. 159–165.

Kruger, H., Verhoef, A. and Preiser, R. (2019). 'The epistemological implications of critical complexity thinking for operational research'. *Systems* 7, 5. doi:10.3390/systems7010005

Krzyzanowska, M.K. *et al.* (2007). 'Using economic analysis to evaluate the potential of multimodality therapy for elderly patients with locally advanced pancreatic cancer'. *Int J Radiat Oncol Biol Phys.* 67, pp. 211–218

Kuhn, T.S. (1962). *The structure of scientific revolutions*. Chicago: The University of Chicago Press.

Kuhn, T.S. (1977). 'Second thoughts on paradigms'. In: Suppe, F. (ed) The *structure of scientific theories.* 2nd ed. Urbaa: University of Illinois Press, pp 459-82.

Kuntz, K. *et al.* (2013). 'Decision and Simulation Modelling in Systematic Reviews'. Rockville (MD): Agency for Healthcare Research and Quality (US).

Kwon, W. *et al*. (2017). 'Multicneter prospective randomized phase II/III study of neoadjuvant chemoradiation with gemcitabine in patients with borderline resectable pancreatic cancer', *51st Annual Meeting of The Pancreas Club*. Chicago, USA 5th-6th May 2017. Available at: https://www.pancreasclub.com/wp-content/uploads/2018/01/PancreasClub_2017finalprogram_web.pdf (Accessed 10th July 2018).

Kyburg, H. (1987). 'Bayesian and non-Bayesian evidential updating'. *Artificial Intelligence,* 31, pp. 271–293

Kyburg, H. (1988). 'Higher order probabilities and intervals.' *Internat. J. Approx. Reason*. 2, pp. 195–209.

La Torre, M. *et al.* (2013). 'Malnutrition and pancreatic surgery: prevalence and outcomes'. *J. Surg. Oncol.* 107(7), pp. 702–708.

Labori, K.J. *et al.* (2017). 'Neoadjuvant chemotherapy versus surgery first for resectable pancreatic cancer (Norwegian Pancreatic Cancer

Trial - 1 (NorPACT-1)) - study protocol for a national multicentre randomized controlled trial'. *BMC Surg*, 17, pp. 94.

Landoni, F. *et al.* (1997). 'Randomized study of radical surgery versus radiotherapy for stage Ib-IIa cervical cancer'. *Lancet,* 350, pp. 535-540.

Landry, J. *et al.* (2010). 'Randomized phase II study of gemcitabine plus radiotherapy vs. gemcitabine, 5-fluorouracil, and cisplatin followed by radiotherapy and 5-fluorouracil for patients with locally advanced, potentially resectable pancreatic adenocarcinoma'*. Journal of Surgical Oncology,* 101(7), pp. 587-592

Lane, D.C., Monefeldt, C. and Husemann, E. (2003). 'Client involvement in simulation model building: hints and insights from a case study in a London hospital'. *Health Care Manage Sci*. 6, pp. 105–16

Langton, C.G. (1989). 'Artificial life'. In *Proceedings of the Santa Fe Institute. Studies in the sciences of complexity*. *Vol. 6*. Redwood City, CA: Addison-Wesley.

Laupacis, A., Sekar, N. and Stiell, I.G. (1997). 'Clinical prediction rules. A review and suggested modifications of methodological standards.' *JAMA*, 277, pp. 488–494.

Laurent, S. *et al*. (2009). 'Feasibility of radiotherapy with concomitant gemcitabine and oxaliplatin in locally advanced pancreatic cancer and distal cholangiocarcinoma: a prospective dose finding phase I-II study.' *Ann Oncol*. 20(8), pp. 1369–1374

Law, J. and Mol, A.M. (2002). 'Complexities: An introduction'. In: Law, J. and Mol, A.M (eds*). Complexities: Social Studies of Knowledge Practoces*. Durham, UK: Duke University Press.

Lawrence, W.W. (1976) *Of acceptable risk.* Los Altos, CA: William Kaufmann Inc.

Le Scodan, R. *et al*. (2009). 'Preoperative chemoradiation in potentially resectable pancreatic adenocarcinoma: feasibility, treatment effect evaluation and prognostic factors, analysis of the SFRO-FFCD 9704 trial and literature review'*. Ann Oncol.* 20**(**8), pp. 1387-96.

Lea, M.S. and Stahlgren, L.H. (1987). 'Is resection appropriate for adenocarcinoma of the pancreas? A cost-benefit analysis.' *Am J Surg.* 154, pp. 651–654

Leach, S.D. *et al.* (1998) 'Survival following pancreaticoduodenectomy with resection of the superior mesenteric-portal vein confluence for adenocarcinoma of the pancreatic head'**.** *Br J Surg*. 85, pp. 611–617

Lee, J.L. *et al.* (2012). 'Prospective efficacy and safety study of neoadjuvant gemcitabine with capecitabine combination chemotherapy for borderline-resectable or unresectable locally advanced pancreatic adenocarcinoma'. *Surgery.* 152 (5), pp. 851 - 86

Lee, C.H. and Yoon, H.J. (2017). 'Medical big data: promise and challenges'. *Kidney Research and Clinical Practice*, 36, pp. 3–11.

Lee, J. *et al*. (2016). 'Clinical impact of neoadjuvant treatment in resectable pancreatic cancer: a systematic review and meta-analysis protocol.' *BMJ*, 6(3), pp. 1-9

Leone, F. *et al.* (2013). 'Induction gemcitabine and oxaliplatin therapy followed by a twice-weekly infusion of gemcitabine and concurrent external-beam radiation for neoadjuvant treatment of locally advanced pancreatic cancer'. *Cancer.* 119, pp. 277–284.

Lermite, E. *et al*. (2013). 'Complications after pancreatic resection: diagnosis, prevention and management.' *Clin Res Hepatol Gastroenterol,* 37, pp. 230–9.

Levy, J. *et al.* (2016). 'Diagnostic laparoscopy with ultrasound still has a role in the staging of pancreatic cancer: a systematic review of the literature'. *HPB Surgery* 20 (16), pp. 1-6.

Lewis, R.S. Jr, and Vollmer, C.M. Jr. (2012). 'Risk scores and prognostic models in surgery: pancreas resection as a paradigm'. *Curr Probl Surg.*49(12), pp. 731-95.

Lin, L.L. *et al*. (2005). 'A phase II study of alternating cycles of split course radiation therapy and gemcitabine chemotherapy for inoperable pancreatic or biliary tract carcinoma'. *Am J Clin Oncol.*28, pp. 234–241.

Lin, J.H. and Haug, P.J. (2008). 'Exploiting clinical data in Bayesian network modeling for predicting medical problems'. *Journal of Biomedical Informatics* 41, pp. 1-14.

Lind, P.A. *et al.* (2008).' Efficacy of preoperative radiochemotherapy in patients with locally advanced pancreatic carcinoma'. *Acta Onco*. l47, pp. 413–420

Lindgren, P., Geborek, P. and Kobelt, G. (2009).'Modeling the cost-effectiveness of treatment of rheumatoid arthritis with rituximab using registry data from Southern Sweden'**.** *Int J Technol Assess Health Care*, 25, pp. 181-189

Litaker, D. *et al.* (2006). 'Using complexity theory to build interventions that improve health care delivery in primary care.' *J Gen Intern Med.* 21(S2), pp. S30–4.

Liu, L. *et al.* (2018). 'A novel scoring system predicts postsurgical survival and adjuvant chemotherapeutic benefits in patients with pancreatic adenocarcinoma: implications for AJCC-TNM staging'. *Surgery*, 163, pp. 1280-1294

Ljungman, D., Lundholm, K. and Hyltander, A. (2011). 'Cost-utility estimation of surgical treatment of pancreatic carcinoma aimed at cure.' *World J Surg.* 35, pp. 662-70.

Löhr, M. *et al.* (2005). Frequency of K-ras mutations in pancreatic intraductal neoplasias associated with pancreatic ductal

adenocarcinoma and chronic pancreatitis: a meta-analysis. *Neoplasia*. 7, pp. 17–23.

Long, K.M. and Meadows, G. (2018). 'Simulation modelling in mental health: a systematic review'. *J Simulation.* 12(1), pp.76–85.

Long, K.M., Dermott, F. and Meadows, G.N. (2018). 'Being pragmatic about healthcare complexity: our experiences applying complexity theory and pragmatism to health services research.' *BMC Medicine*, 16:94

Lucas, P.J.F. *et al.* (2004). 'Bayesian networks in biomedicine and health-care.' *Artificial Intelligence in Medicine*. 2004, pp. 201-214.

Luchini, C., Capelli, P. and Scarpa, A. (2016). 'Pancreatic Ductal Adenocarcinoma and Its Variants.' *Surg Pathol Clin*. 9, pp. 547–560

Luengo-Fernandez, R. *et al.* (2013). 'Economic burden of cancer across European Union: a population-based cost analysis.' *Lancet* 14(12), pp. 1165-1174.

MacConaill, L.E., Lindeman, N.I., Rollins, B.J. (2015). 'Brave-ish New World—What's Needed to Make Precision Oncology a Practical Reality.' *JAMA Oncol*. 1(7), pp. 879–880.

Magnin, V. *et al.* (2003). 'Neoadjuvant preoperative chemoradiation in patients with pancreatic cancer.' *Int J Radiat Oncol Biol Phys.* 55(5), pp. 1300-4.

Magnino, A. *et al*. (2005). 'Phase II trial of primary radiation therapy and concurrent chemotherapy for patients with locally advanced pancreatic cancer.' *Oncology*, 68, pp. 493-499.

Mahaseth, H. *et al.* (2013). 'Modified FOLFIRINOX regimen with improved safety and maintained efficacy in pancreatic adenocarcinoma**'.** *Pancreas*, 42(8), pp. 1311-1315

Malinchoc, M. *et al.* (2000). 'A model to predict poor survival in patients undergoing transjugular intrahepatic portosystemic shunts.' *Hepatology*, 31(4), pp. 864-71.

Mangel, M. and Samaniego, FJ. (1984). 'Abraham Wald's work on aircraft survivability'. *Journal of the American Statistical Association* 70 (386), pp. 259-267

Marchal, B. *et al*. (2013). 'Realist RCTs of complex interventions – an oxymoron'. *Soc Sci Med*. 94, pp. 124–8.

Marcus, G. (2018). 'Deep learning: a critical appraisal'. *arXiv*, 1-27.

Malottki, K. *et al.* (2011). 'Adalimumab, etanercept, infliximab, rituximab and abatacept for the treatment of rheumatoid arthritis after the failure of a tumour necrosis factor inhibitor: a systematic review and economic evaluation.' *Health Technol Assess*, 15, pp. 1-278

Manson, S.M. (2001). 'Simplifying complexity: a review of complexity theory.' *Geoforum*. 32(3), pp. 405–14

Markus, P.M. *et al*. (2005). 'Predicting postoperative morbidity by clinical assessment'. *Br J Surg,* 92(1), pp. 101-6.

Marsh, D. and Stoke, G. (eds). (2002). *Theory and Methods in Political Science, 2nd ed.* Basingstoke: Palgrave Macmillan.

Marshall, A., Altman, D.G. and Holder, R.L. (2010). 'Comparison of imputation methods for handling missing covariate data when fitting a Cox proportional hazards model: a resampling study'. *BMC Med Res Methodol* 10, pp 112.

Marthey, L. *et al.* (2015). 'FOLFIRINOX for locally advanced pancreatic adenocarcinoma: results of an AGEO multicenter prospective observational cohort'**.** *Ann Surg Oncol,* 22(1), pp. 295-301

Marti, J.L. *et al.* (2008). 'Phase I/II Trial of Induction Chemotherapy Followed by Concurrent Chemoradiotherapy and Surgery for Loco-regionally Advanced Pancreatic Cancer.' *Ann Surg Oncol*. 15(12), pp. 521-3531.

Massucco, P. *et al.* (2006). 'Pancreatic resections after chemoradiotherapy for locally advanced ductal adenocarcinoma: analysis of perioperative outcome and survival.' *Ann Surg Oncol.* 13, pp. 1201–1208.

Mattiucci, G.C. et al. (2010). 'External beam radiotherapy plus 24-hour continuous infusion of gemcitabine in unresectable pancreatic carcinoma: long-term results of a phase II study.' *Int J Radiat Oncol Biol Phys*. 76, pp. 831–8.

Maximous, D.W. *et al.* (2009). 'Preoperative gemcitabine based chemo-radiotherapy in locally advanced non metastatic pancreatic adenocarcinoma.' *International Archives of Medicine*. 2:7.

Mayo, S.C. *et al.* (2009). 'Evolving preoperative evaluation of patients with pancreatic cancer: does laparoscopy have a role in the current era?' *J Am Coll Surg,* 208, pp. 87-95.

McGuigan, A. *et al*. (2018). 'Pancreatic cancer: a review of clinical diagnosis, epidemiology, treatment and outcomes'. *World Journal of Gastroenterology,* 24(43), pp. 4846-4861

McKelvey, B. (1999). 'Complexity theory in organization science: seizing the promise or becoming a fad?' *Emergence*. 1(1), pp. 5–32.

McShane, L.M. *et al*. (2005). 'Reporting recommendations for tumor marker prognostic studies (REMARK).' *J Natl Cancer Inst*, 97, pp. 1180–1184.

McNeill, F.M. and Thro, E. (1994). *Fuzzy Logic: A Practical Approach.* Cambridge, Massachutes: Academic Press Professional.

Mellon, E.A. *et al.* (2016). 'Favorable perioperative outcomes after resection of borderline resectable pancreatic cancer treated with

neoadjuvant stereotactic radiation and chemotherapy compared with upfront Pancreatectomy for resectable cancer.' *Journal of Gastrointestinal Oncology*. 7(4), pp. 547-555.

Merali, Y. (2006). 'Complexity and information systems: The emergent domain.' *J. Inform. Technol,* 21, pp. 216–228

Midgley, G. (2000). *Systemic Intervention: Philosophy, Methodology and Practice*. New York, USA: Kluwer Academinc/Plenum Publishers.

Midgley, G. and Richardson, K.A. (2007). 'Systems thinking for community involvement in policy analysis.' *Emerg. Complex. Organ. 9*, pp. 167–183.

Midgley, G., Munlo, I. and Brown, M. (1998). 'The theory and practice of boundary critique: Developing housing services for older people'. *J. Oper. Res. Soc.*, *49*, pp. 467–478.

Midha, S., Chawla, S. and Garg, P.K. (2016). 'Modifiable and non-modifiable risk factors for pancreatic cancer: A review.' *Cancer Lett*, 381, pp. 269–277

Miettinen, O.S. and Caro, J.J. (1989). 'Principles of non-experimental assessment of excess risk, with special reference to adverse drug reactions.' *J Clin Epidemiol,* 42, pp. 325-331

Miksad, R.A., Schnipper, L. and Goldstein, M. (2007). 'Does a statistically significant survival benefit of erlotinib plus gemcitabine

for advanced pancreatic cancer translate into clinical significance and value?' *J Clin Oncol.,* 25, pp. 4506–4507

Milton, C.W. *et al.* (2001). 'Modeling for Health Care and Other Policy Decisions: Uses, Roles, and Validity.' *Value in Health,* 4(5), pp. 348-361

Miller, M.E. *et al.* (1993). 'Validation of probabilistic predictions'. *Med Decis Making,* 13(1), pp. 49-58.

Mingers, J. (1992). 'Recent developments in critical management science'. *J. Operat. Res. Soc.*, *43*, pp. 1–10.

Mingers, J. (2004). 'Realizing information systems: critical realism as an underpinning philosophy for information systems'. *Information and Organization:* 14(2), pp. 87-103.

Mingers, J. (2005). 'A critique of statistical modeling in management science from a critical realist perspective: its role within multimethodology'. *Journal of the Operational Research Society,* 57(2), pp. 202-219.

Mingers, J. and Rosenhead, J. (2004). 'Problem structuring methods in action.' *Eur. J. Oper. Res, 152*, pp. 530–554.

Miura, T. *et al.* (2014). 'A new preoperative prognostic scoring system to predict prognosis in patients with locally advanced pancreatic body cancer who undergo distal Pancreatectomy with en bloc celiac axis resection: a retrospective cohort study'. *Surgery* 155, pp. 457-67.

Moffitt, R. A. *et al*. (2015). 'Virtual microdissection identifies distinct tumor- and stroma-specific subtypes of pancreatic ductal adenocarcinoma.' *Nat. Genet.* 47, pp. 1168–1178.

Mohammed, S., Van Buren, G. and Fisher, W.E. (2014).'Pancreatic cancer: advances in treatment.' *World J Gastroenterol.*, 20, pp. 9354–9360.

Moher, D. *et al.* (2009). 'Preferred reporting items for systematic reviews and meta-analyses: the PRISMA Statement'. *Open Med*, 3, pp. 123–130

Mokdad, A.A. *et al* (2017). 'Neoadjuvant therapy followed by resection versus upfront resection for resectable pancreatic cancer: a propensity score matched analysis.' *J Clin Oncol,* 35(5), pp. 515-522

Mollberg, N. *et al*. (2011). 'Arterial resection during pancreatectomy for pancreatic cancer: a systematic review and meta-analysis'. *Ann Surg,* 254, pp. 882–93.

Moldovan, S.C. et al. (2012). 'The advantages of retropancreatic vascular dissection for pancreatic head cancer with portal/superior mesenteric vein invasion: posterior approach pancreatico-duodenectomy technique and the mesopancreas theory'. *Chirurgia (Bucur)*, 107, pp. 571–8.

Montori, V. (2017). 'Why we revolt: A patient revolution for careful and kind care.' Rochester, Minnesota: The Patient Revolution

Moons, K.G.M *et al.* (2009). 'Prognosis and prognostic research: what, why, and how?' *BMJ,* 338: b375.

Moons, K.G.M. *et al.* (2014). 'Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies: The CHARMS Checklist.' *PLoS Med*, 11(10): e1001744.

Moore, G.F. *et al.*, (2015). 'Process evaluation of complex interventions: Medical Research Council guidance'. BMJ, 350:h1258

Mornex ,F. *et al.* (2006). 'Feasibility of preoperative combined radiation therapy and chemotherapy with 5-fluorouracil and cisplatin in potentially resectable pancreatic adenocarcinoma: the French SFRO-FFCD 97-04 Phase II trial.' *Int J Radiat Oncol Biol Phys.* 65, pp. 1471–1478

Morak, M.J.M. *et al*. (2008). 'Adjuvant intra-arterial chemotherapy and radiotherapy versus surgery alone in resectable pancreatic and periampullary cancer a prospective randomized controlled rrial.' *Ann Surg.* 248, pp. 1031–1041.

Morris, S. *et al.* (2015). 'Cost-effectiveness of diagnostic laparoscopy for assessing resectability in pancreatic and periampulary cancer.' *BMC Gastroenterology,* 15, pp. 44:1-8

Motoi, F. et al. (2013). 'Neoadjuvant chemotherapy with gemcitabine and S-1 for resectable and borderline pancreatic ductal

adenocarcinoma: Results from a prospective multi-institutional phase 2 trial.' *Ann Surg Oncol.* 20, pp. 3794–3801.

Moustafa, A. M. (2010). 'Risk based decision making tools for sewer infrastructure management'. PhD dissertation, University of Cincinnati, Cincinnati, USA.

Moutardier, V. *et al*. (2002). 'A phase II single institutional experience with preoperative radiochemotherapy in pancreatic adenocarcinoma'. *Eur J Surg Oncol,* 28, pp. 531-539.

Mowat, E. and Davis, B. (2010). 'Interpreting embodied mathematics using network theory: Implications for mathematics education'. *Complicity Int. J. Complex. Educ.*, *7*, pp. 1–31.

Muir-Gray J. (2001). *Evidence-Based Health Care*. Edinburgh: Churchill Livingstone.

Muller, S.A. *et al.* (2009). 'Vascular resection in pancreatic cancer surgery: survival determinants.' *J Gastrointest Surg*, 13, pp. 784–92.

Murakami, Y. *et al*. (2013). 'Benefit of portal or superior mesenteric vein resection with adjuvant chemotherapy for patients with pancreatic head carcinoma.' *J Surg Oncol*, 107, pp. 414–21.

Murakami, Y. *et al.* (2017). 'Survival impact of neoadjuvant gemcitabine plus S-1 chemotherapy for patients with borderline resectable pancreatic carcinoma with arterial contact.' *Cancer Chemother Pharmacol.* 79, 37.

Murphy, J.D. *et al*. (2012). *'*Cost-effectiveness of modern radiotherapy techniques in locally advanced pancreatic cancer'. *Cancer,* 118, pp. 1119-29.

Murphy, J.E. *et al.* (2018). 'Total neoadjuvant therapy with FOLFIRINOX followed by individualized chemoradiotherapy for borderline resectable pancreatic adenocarcinoma: a phase 2 clinical trial.' *JAMA Oncol*, 4(7), pp. 963-969

Muthuraman, T. and Sankaran, G. (2014). 'A framework for personalized decision support system for healthcare application.' *Health Science Journal,* 8(2), pp. 262-273

Nadkarni, S. and Shenoy, P.P. (2001). 'A Bayesian network approach to making inferences in causal maps.' *European Journal of Operational Research*. 128(3), pp. 479-498

Nakagohri, T. *et al.* (2003). *'*Survival benefits of portal vein resection for pancreatic cancer.**'** *Am J Surg*. 186, pp.149–153

Nakano, H. *et al.* (2002). **'**Arterial and vena caval resections combined with pancreaticoduodenectomy in highly selected patients with periampullary malignancies.' *Hepatogastroenterology*, 49, pp. 258–262

Namkoong, S.E. *et al.* (1995). 'Comparative study of the patients with locally advanced stages I and II cervical cancer treated by radical

surgery with and without preoperative adjuvant chemotherapy'. *Gynecol Oncol,* 59, pp. 136-42.

Nanda, R.H. *et al.* (2015). 'Neoadjuvant modified FOLFIRINOX and chemoradiation therapy for locally advanced pancreatic cancer improves resectability.' *J Surg Oncol*, 111(8), pp. 1028-1034

National Comprehensive Cancer Network (NCCN). (2015). *Clinical Practice Guidelines in Oncology. Pancreatic Adenocarcinoma. Version 2* Available at: www.nccn.org/professionals/physician_gls/f_guidelines.asp.Google Scholar (Downloaded 25th October, 2017).

National Comprehensive Cancer Network (NCCN). (2018). *Clinical Practice Guidelines in Oncology (NCCN Guidelines). Pancreatic Adenocarcinoma. V2.2018.* Available at: https://www.nccn.org/professionals/physician_gls/pdf/pancreatic.pdf. (Dowlnloaded 25th August, 2018).

National Institute of Clinical Excellence (NICE). (2011). *Discounting of health benefits in special circumstances*. London: NICE.

National Health Service (NHS). *National Schedule of Refference Costs-Year 2011-12- NHS trusts and NHS foundation trusts: NHS own costs.* Availanle at: https://www.gov.uk/governments/publications/nhs-reference-costs-financal-year-2011-to-2012 (Accessed 25th August 2017).

National Institute of Clinical Excellence (NICE). (2013). *Guide to the methods of technology appraisal*. London: NICE

National Institute for Clinical Excellence (NICE). (2018). *Pancreatic cancer in adults: diagnosis and management. NICE guideline [NG85}.* Available at: https://www.nice.org.uk/guidance/ng85 (Accessed June 2019).

National Research Council. (1991). *Improving Information for Social Policy Decisions: the Uses of Micro- simulation Modeling. Vol 1. Review and Recommendations*. Washington, DC: National Academy Press.

National Research Council. (2011). *Toward Precision Medicine: Building a Knowledge Network for Biomedical Research and a New Taxonomy of Disease.* Washington, DC: The National Academies Press.

Neoptolemos, J.P. et al. (2001). Adjuvant chemoradio- therapy and chemotherapy in resectable pancreatic cancer: a randomized controlled trial. *Lancet*. 358**,** pp. 1576–85.

Neoptolemos, J.P. *et al.* (2004). 'A randomized trial of chemoradiotherapy and chemotherapy after resection of pancreatic cancer.' *NEJM* 350, pp. 1200–10.

Neoptolemos, J.P. *et al.* (2010). 'Adjuvant chemotherapy with fluorouracil plus folinic acid vs gemcitabine following pancreatic cancer resection: a randomized controlled trial.' *JAMA*, 304, pp. 1073–1081.

Neoptolemos, J.P. *et al*. (2017). 'Comparison of adjuvant gemcitabine and capecitabine with gemcitabine monotherapy in patients with resected pancreatic cancer (ESPAC-4): a multicentre, open-label, randomised, phase 3 trial.' *Lancet.* 389, pp. 1011–1024

Neoptolemos, J.P. *et al*. (2018). 'Therapeutic developments n pancreatic cancer: current and future perspectives. '*Nature Reviews Gastroenterology & Hepatology,* 15, pp. 333–348

Neumann, P.J. (2009). 'Costing and perspective in published cost-effectiveness analysis'. *Med Care*. 47:7(1), pp. 28-32.

Noll, E. M. *et al*. (2016). 'CYP3A5 mediates basal and acquired therapy resistance in different subtypes of pancreatic ductal adenocarcinoma.' *Nat. Med.* 22, pp. 278–287.

Nurmi, A. *et al.* (2018). 'Neoadjuvant therapy offers longer survival than upfront surgery for poorly differentiated and higher stage pancreatic cancer.' *Acta Oncologica*. 57(6), pp. 799-806.

Obermeyer, Z. M.D and Lee, T.H. (2017). 'Lost in Thought — The Limits of the Human Mind and the Future of Medicine.' *NEJM,* 377, pp. 1209-1211

Oettle, H. *et al.* (2013). 'Adjuvant chemotherapy with gemcitabine and long-term outcomes among patients with resected pancreatic cancer: the CONKO-001 randomized trial.' *JAMA,* 310, pp. 1473–1481.

O'Neill, C.B. *et al.* (2012). 'Costs and trends in pancreatic cancer treatment.' *Cancer*, 118(20), pp. 5132–5139.

O'Reilly, E.M. *et al.* (2014). 'A single-arm, non-randomized phase II trial of neoadjuvant gemcitabine and oxaliplatin in patients with resectable pancreas adenocarcinoma.' *Annals of surgery*. 260(1), pp. 142-148.

Ormerod, R.J. and Ulrich, W. (2013). 'Operational research and ethics: A literature review'. *Eur. J. Oper. Res.*, *228*, pp. 291–307.

Ozols, R.F. *et al*. (2003). 'Phase III trial of carboplatin and paclitaxel compared with cisplatin in patients with optimally resected stage III ovarian cancer: a Gynecological Oncology Group Study.' *J Clin Oncol,* 21, pp. 3194-200.

Paaß, G. (1991). 'Second order probabilities for uncertain and conflicting evidence'. In: Bonissone, P. *et al.*n(eds.). *Uncertainty in Artificial Intelligence*. Amsterdam: North-Holland, pp. 447– 456.

Palmer, D. *et al*. (2007). 'A randomized phase 2 trial of neoadjuvant chemotherapy in resectable pancreatic cancer: gemcitabine alone versus gemcitabine combined with cisplatin.' *Ann Surg Oncol.*14, pp. 2088–2096.

Pan, F. *et al.* (2018). 'Modeling clinical outcomes in prostate cancer: application and validation of the discrete event simulation approach.' *Value in Health*, pp. 416-422

Pancreatic Cancer UK. (2017). *Pancreatic Cancer UK policy briefing: every life matters: the real cost of pancreatic cancer diagnosis via emergency admission.* Available at: https://www.pancreaticcancer.org.uk/media/86662/every-im_policybriefing-final.pdf. (Accessed 26th June 2017.)

Papalezova, K.T. *et al.* (2012). 'Does preoperative therapy optimize outcomes in patients with resectable pancreatic cancer?' *Journal of Surgical Oncology.* 106, pp. 111-118.

Papavasiliou, P. *et al.* (2014). 'Impact of preoperative therapy on patterns of recurrence in pancreatic cancer.' *HPB,* 16(1), pp. 34–39.

Parmigiani, G. (2002). *Modeling in medical decision-making: a Bayesian approach.* Chichester: Wiley

Parikh, P. *et al.* (2010). 'Pancreatectomy risk calculator: an ACS-NSQIP resource.' *HPB*, 12(7), pp. 488-97

Partelli, S. *et al.* (2014). 'Evaluation of a predictive model for pancreatic fistula based on amylase value in drains after pancreatic resection.' *Am J Surg.* 208(4), pp. 634-9.

Paul, S *et al.* (2014). 'Developing a management decision-making model based upon a complexity perspective with reference to the Bee Algorithm.' *Emerg. Complex. Organ., 16*, pp. 1–12.

Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems*. San Mateo, CA: Morgan Kaufmann.

Pearl, J. (1990). 'Jeffrey's rule, passage of experience, and Neo-Bayesianism'. In: Kyburg, H. et al. (eds.). *Knowledge Representation and Defeasible Reasoning*. Dordrecht: Kluwer Academic, pp. 245–265.

Pearl, J. (2000). *Causality*. Cambridge, UK: Cambridge University Press.

Peduzzi, P. *et al.* (1995). 'Importance of events per independent variable in proportional hazards regression analysis. II. Accuracy and precision of regression estimates.' *J Clin Epidemiol,* 48, pp. 1503–1510.

Peduzzi, P. *et al.* (1996). 'A simulation study of the number of events per variable in logistic regression analysis.' *J Clin Epidemiol*. 49, pp. 1373–1379.

Peeters, K.C. *et al*. (2005). 'Validation of a nomogram for predicting disease-specific survival after an R0 resection for gastric carcinoma.' *Cancer*, 103(4), pp. 702-7.

Pereira, R.D.M.A. *et al.* (2015). 'Fuzzy modelling to predict severely depressed left ventricular ejection fraction following admission to the intensive care unit using clinical physiology. '*The Scientific World Journal*, 1-9.

Petrelli, F. *et al*. (2015). 'FOLFIRINOX-based neoadjuvant therapy in borderline resectable or unresectable pancreatic cancer: a meta-analytical review of published studies.' *Pancreas,* 44, pp. 515-21.

Pericleous, S. *et al*. (2012). 'Systematic review and meta-analysis of case-matched studies comparing open and laparoscopic distal pancreatectomy: is it a safe procedure?' *Pancreas* , 41, pp. 993–1000.

Pidd, M. (2004). *Computer Simulation in Management Science.* Chichester: John Wiley & Sons.

Pipas, J. *et al.* (2012). 'Neoadjuvant cetuximab, twice-weekly gemcitabine, and intensity-modulated radiotherapy (IMRT) in patients with pancreatic adenocarcinoma.' *Ann Oncol*. 23, pp. 2820–2827.

Pisters, P. *et al*. (2002). 'Preoperative paclitaxel and concurrent rapid-fractionation radiation for resectable pancreatic adenocarcinoma: toxicities, histologic response rates, and event-free outcome.' *J Clin Oncol*. 20, pp. 2537–2544.

Pletcher, M.J. and McCulloch, C.E. (2017). 'The Challenges of Generating Evidence to Support Precision Medicine.' *JAMA Intern Med*. 177(4), pp. 561–562.

Plsek, P.E. (2003). 'Complexity and the adoption of innovation in health care'. In: *Accelerating quality improvement in health care: strategies to accelerate the diffusion of evidence-based innovations.* Washington, DC: National Institute for healthcare management foundation and National Committee for Quality in Health Care.

Plsek, P.E. and Greenhalgh, T. (2001). 'Complexity science: The challenge of complexity in health care.' *BMJ*, 323(7313), pp. 625–628.

Plsek, P.E. and Wilson, T. (2001). 'Complexity, leadership, and management in healthcare organisations.' *BMJ*, 323, pp. 746–9.

Poonawalla, I.B. *et al.* (2015). 'Cost-effectiveness of neoadjuvant chemotherapy versus primary surgery in elderly patients with advanced ovarian cancer.' *Value in Health,* pp. 387-395.

Popper, K.R. (1959). *The logic of scientific discovery.* London: Hutchinson.

Poropudas, J. and Virtanen, K. (2011). 'Simulation metamodeling with dynamic Bayesian networks.' *European Journal of Operational Research,* 214 (3), pp. 644-655

Poston, G. *et al.* (2001). 'Cost of neoadjuvant chemotherapy and surgery in patients with liver metastases from advanced colorectal cancer.' *Journal of Medical Economics*, 4, pp. 167-177.

Pratihar, D.K., Deb, K. and Ghosh, A. (1999). 'A genetic-fuzzy approach for mobile robot navigation among moving obstacles.' *Int J Approx Reason*, 20, pp. 145-172.

Pratt, W.B. *et al.* (2007). 'Clinical and economic validation of the International Study Group of Pancreatic Fistula (ISGPF) classification scheme.' *Ann Surg*, 245(3), pp. 443-51.

Pratt, W. *et al.* (2008). 'POSSUM accurately predicts morbidity for pancreatic resection'. *Surgery*, 143(1), pp. 8-19.

Prediction Tools, 2012. Available at: http://www.mskcc.org/cancer-care/ prediction-tools. (Accessed 16th March, 2017).

Preiser, R. and Woermann, M. (2016). 'Introduction'. In Cilliers, P *et al.* (eds)*. Critical Complexity Collected Essays*. Berlin, Germany: De Gruyter (Ontos).

Prigogine, I. (1987). 'Exploring Complexity.' *Eur. J. Oper. Res.* 30, pp. 97–103.

Prytherch, D.R. *et al*. (1998). 'POSSUM and Portsmouth POSSUM for predicting mortality. Physiological and Operative Severity Score for the enUmeration of Mortality and morbidity.' *Br J Surg*, 85(9), pp. 1217-20.

Pu, N. *et al.* (2017). 'Alkaline phosphatase-to-albumin ratio as a prognostic indicator in pancreatic ductal adenocarcinoma after curative resection.' *Journal of Cancer*, 8:16, pp. 3362-3370

Pu, N. *et al.* (2018). 'Comparison of prognostic prediction between nomogram based on lymph node ratio and AJCC 8th staging system for patients with resected pancreatic head carcinoma: a SEER analysis.' *Cancer Management Research,* 10, pp. 227-238.

Puhan, M.A. *et al.* (2014). 'A GRADE Working Group approach for rating the quality of treatment effect estimates from network meta-analysis'. *BMJ,* **349,** g5630.

Puleo, F. *et al*. (2018). 'Stratification of pancreatic ductal adenocarcinomas based on tumor and microenvironment features.' *Gastroenterology* 155, pp. 1999–2013.

Raghupathi, W. and Raghupathi, V. (2014). 'Big data analytics in healthcare: promise and potential.' *Health Inform Sci Syst,* 2:3.

Rahib, L. *et al*. (2014). 'Projecting cancer incidence and deaths to 2030: the unexpected burden of thyroid, liver, and pancreas cancers n the United States'. *Cancer Res,* 74, pp. 2913-2921.

Raikar, G.V. et al. (1996). 'Cost-effective analysis of surgical palliation versus endoscopic stenting in the management of unresectable pancreatic cancer. '*Ann Surg Oncol*. 3, pp. 470–475.

Raufi, A.G. *et al*. (2019). 'Neoadjuvant Treatment for Pancreatic Cancer'. *Seminars in Oncology* , 46, pp. 19-27

Reid, M.D. *et al*. (2014). 'Molecular genetics of pancreatic neoplasms and their morphologic correlates: an update on recent advances and potential diagnostic applications.' *Am J Clin Pathol*. 141, pp. 168–180

Reis, M.A.M., Ortega, N.R.S. and Silveira, P.S.P. (2004). 'Fuzzy expert system in the prediction of neonatal resuscitation.' *Brazilian Journal of Medical and Biological Research.* 37, pp. 755-765.
Reitermanov´a, Z. (2010). *Data Splitting*. WDS'10 Proceedings of Contributed Papers, Part I, pp. 31–36.

Regine, W.F. *et al*. (2008). 'Fluoracil vs gemcitabine chemotherapy beforeand after fluorouracil-based chemoradiation following esection of pancreatic adenocarcinoma; a randomized controlled trial.' *JAMA.* 5: 299(9), pp. 1019-26.

Regine, W.F. *et al.* (2011). 'Fluorouracil-based chemoradiation with either gemcitabine or fluorouracil chemotherapy after resection of pancreatic adenocarcinoma: 5-year analysis of the U.S. Intergroup/RTOG 9704 phase III trial.' *Ann Surg Oncol*, 18, pp. 1319–26.

Reni, M. *et al.* (2012). 'Adjuvant PEFG (Cisplatin, Epirubicin, 5-Fluorouracil, Gemcitabine) or Gemcitabine followed by chemoradiation in pancreatic cancer: a randomized phase II trial.' *Ann Surg Oncol*. 19, pp. 2256–2263

Rettenmaier, M.A. *et al*. (1989). 'Radical hysterectomy and tailored postoperative radiation therapy in the management of bulky stage 1B cervical cancer.' *Cancer,* 63, pp. 2220-3.

Riall, T.S. and Lillemoe, K.D. (2007). 'Underutilization of surgical resection in patients with localized pancreatic cancer.' *Ann Surg.* 246, pp. 181–182.

Riediger, H. *et al.* (2006). 'Postoperative morbidity and long-term survival after pancreaticoduodenectomy with superior mesenterico-portal vein resection.' *J Gastrointest Surg*, 10, pp. 1106–15.

Riley, R.D. *et al.* (2016). 'External validation of clinical prediction models using big datasets from e-health records or IPD meta-analysis: opportunities and challenges.' *BMJ,* 353:i3140

Riley, R.D. *et al*. (2019). 'Minimum sample size for developing a multivariable prediction model: PART II - binary and time-to-event outcomes.' *Statistics in Medicine*, 38, pp. 1276– 1296.

Robinson, S. and Pidd, M. (1998). 'Provider and customer expectations of successful simulation projects.' *J Operat Res Soc*. 49(3), pp. 200–9.

Rocconi, R.P. *et al.* (2005). 'Management strategies for stage IB2 cervical cancer: a cost-effectiveness analysis.' *Gynecologic Oncology,* 97, pp. 387-394.

Rohrmann, S. *et al*. (2009). 'Ethanol intake and the risk of pancreatic cancer' *European Prospective Investigation into Cancer and Nutrition (EPIC) Cancer Causes Control*. 20, pp. 785–794.

Roland, C.L. *et al.* (2015). 'Neoadjuvant therapy is associated with a reduced lymph node ratio in patients with potentially resectable pancreatic cancer.' *Annals of Surgical Oncology*. 22(4), pp. 1168-1175.

Romanus, D. *et al.* (2012). Does health-related quality of life improve for advanced pancreatic cancer patients who respond to gemcitabine? Analysis of a randomized phase III trial of the cancer and leukemia group B (CALGB 80303). *J Pain Symptom Management.* 43: pp. 205-17.

Rosenhead, J. (1996). 'What's the problem? An introduction to problem structuring methods.' *Interfaces*, *26*, pp. 117–131.

Rosenhead, J. (2006). 'Past, present and future of problem structuring methods.' *J. Oper. Res. Soc.*, *57*, pp. 759–765.

Rosenhead, J. and Mingers, J. (2001). *Rational Analysis for a Problematic World Revisited: Problem Structuring Methods for Complexity, Uncertainty and Conflict*, 2nd ed.: Chichester, UK: Wiley.

Rothwell, P.M. (2008). 'Prognostic models'. *Pract Neurol*, 8, pp. 242–253.

Rouder, J.N. and Speckman, P.L. (2004). 'An evaluation of the Vincentizing method of forming group-level response time distributions.' *Psychon Bull Rev*. 11, pp. 419–427.

Rowland, M.R. *et al*. (2015). 'Cost-utility comparison of neoadjuvant chemotherapy versus primary debulking surgery for treatment of advanced-stage ovarian cancer in patients 65 years old or older.' *American Journal of Obstetrics & Gynecology,* 212(763), pp. 1-8.

Royall, N.A. and Bose, D. (2015). 'Impact of neoadjuvant therapy on the surgical management of pancreatic cancer: a literature review.' *Transl Cancer Res*, 4(6), pp. 641-652.

Roychowdhury, A. *et al*. (2004). 'Diagnosis of the diseases-using a GA-fuzzy approach.' *Inf Sci*. 162, pp. 105-120.

Royston, P., Altman, D.G. and Sauerbrei, W. (2006). 'Dichotomizing continuous predictors in multiple regression: a bad idea.' *Stat Med,* 25, pp. 127–141.

Rubin, D.B. (1984). 'Bayesianly justifiable and relevant frequency calculations for applied statistician.' *Annals of Statistics,* 12, pp. 1151-1

Russell, B. (1929). 'On the notion of cause, with applications to the free-will problem.' In Feigll, H. and Brodbeck, M. (eds). *Readings in Philosophy of Science.* New York: Appleton-Century-Crofts Inc, pp: 387-407.

Russell, B. (1923). 'Vagueness'. *Aust J Psychol Philos,* 1, pp. 84-92

Russell, M.C. (2016). 'Comparison of neoadjuvant versus a surgery first approach for gastric and esophagogastric cancer'. *Journal of Surgical Oncology:* 114, pp. 296-303.

Rutter, H. *et al.* (2017). 'The need for a complex systems model of evidence for public health.' *Lancet.* 390(10112), pp. 2602-47

Saad, A.M. *et al.* (2018). 'Trends in pancreatic adenocarcinoma incidence and mortality in the United States in the last four decades; a SEER-based study'. *BMC Cancer*, 18:688

Sadegh-Zadeh, K. (1981). 'On the relative avoidability and the absolute unavoidability of misdiagnoses.' *Wissenschaftliche Inform*. 7, pp. 33-43.

Sadegh-Zadeh, K. (1983). *Medicine as ethics and constructive Utopia 1.* Burgverlag: Tecklenburg

Sadegh-Zadeh, K. (1994). 'Fundamentals of clinical methodology: 1. Differential indication.' *Artif Intell Med*, 6, pp. 83-102.

Sadegh-Zadeh K. (1999). 'Fundamentals of clinical methodology. 3.Nosology.' *Artif Intell Med,* 17, pp. 87-108.

Sadegh-Zadeh, K. (1998). 'Fundamentals of clinical methodology. 2. Etiology.' *Artif Intell Med*, 12, pp. 227-70.

Sadegh-Zadeh, K. (2000). 'Fundamentals of clinical methodology. 4. Diagnosis.' *Artif Intell Med*, 20, pp, 227-41

Sadegh-Zadeh, K. (2001). 'The fuzzy revolution: goodbye to the Aristotelian weltanschauung.' *Artificial Intelligence in Medicine,* pp. 1-25

Sahora, K. *et al.* (2011). 'Neogemtax: gemcitabine and docetaxel as neoadjuvant treatment for locally advanced nonmetastasized pancreatic cancer.' *World J Surg.,* 35, pp. 1580–1589.

Salanti, G. *et al.* (2014). 'Evaluating the quality of evidence from a network meta-analysis'. *PLoS One*. 9**,** 7, e99682.

Sanders, G.D. *et al.* (2016). 'Recommendations for Conduct, Methodological Practices, and Reporting of Cost-effectiveness Analyses: Second Panel on Cost-Effectiveness in Health and Medicine.' *JAMA,* 316(10), pp. 1093-1103.

Sanderson, I. (2009). 'Intelligent policy making for a complex world: pragmatism, evidence and learning.' *Political Studies*. 57(4), pp. 699–719.

Satoi, S. *et al.* (2009). 'Surgical results following pre-operative chemoradiation therapy for patients with pancreatic cancer.' *Pancreas.* 38, pp282-288.

Savage, L.J. (1954). *The foundations of statistics.* New York: Wiley.

Saville, C.E., Smith, H.K. and Bijak, K. (2019). 'Operational research techniques applied throughout cancer care services: a review.' *Health Systems.* 8(1), pp. 52-73

Schneeweiss, A. *et al.* (2013). 'Pertuzumab plus trastuzumabin combination with standard neoadjuvant anthracycline-containing and anthrax-cycline-free chemotherapy regimes in patients with HER-2 positive early breast cancer: a randomized phase II cardiac safety study (TRYPHAENA).' *Ann Oncol* 24, pp. 2278-84.

Sadot, E. A. *et al.* (2015). 'FOLFIRINOX induction therapy for stage 3 pancreatic adenocarcinoma.' *Ann Surg Oncol*, 22 (11), pp. 3512-3521

Sahotra, S. (1996). *The emergence of logical empiricism: from 1900 to the Vienna Circle..* New York: Garland Publishing.

Saklad, M. (1941). 'Grading of patients for surgical procedures'. *Anesthesiology*, 2:4.

Saif, M.W. (2007). 'Controversies in the adjuvant treatment of pancreatic adenocarcinoma.' *JOP*. 8, pp. 545-52

Saif, M.W., Syrigos, K.N. and Katirtzoglou, N.A. (2009). 'S-1: a promising new oral fluoropyrimidine derivative.' *Expert Opin Investig Drugs*, 18(3), pp. 335-48.

Saif, M.W. (2013). 'Advancements in the Management of Pancreatic Cancer'. *JOP,* 14(2), pp. 112-118

Sato-Dahlman, M., Wirth, K. and Yamamoto, M. (2018). 'Role of gene therapy in pancreatic cancer: a review.' *Cancers* , 10: 103; doi: 10.3390/cancers10040103

Satoi, S. *et al.* (2009). 'Surgical results following pre-operative chemoradiation therapy for patients with pancreatic cancer.' *Pancreas.* 38, pp. 282-288.

Sauvanet, A. *et al.* (2015). 'Severe jaundice increases early severe morbidity and decreases long-term survival after pancreaticoduodenectomy for pancreatic adenocarcinoma.' *J. Am. Coll. Surg.* 221(2), pp. 380–389.

Schmidt, J. *et al*. (2012). 'Open-label, multicenter, randomized phase III trial of adjuvant chemoradiation plus interferon alfa-2b versus fluorouracil and folinic acid for patients with resected pancreatic adenocarcinoma.' *Journal of Clinical Oncology.* 30:33, pp. 4077-4083

Schon, D. (1984). *The Reflective Practitioner: How Professionals Think in Action*. New York, NY: Basic Books.

School, R. *et al.* (2013). 'A Gentle Introduction to Bayesian Analysis: applications to development research'. *Child Development*, 85 (3), pp. 842-860.

Seker, H. (2003). 'A fuzzy logic based-method for prognostic decision making in breast and prostate cancers.' *IEE Transactions on Information Technology in Biomedicne.* 2, pp. 114-122.

Schwartz, W.B. (1979). 'Sounding board. Decision Analysis: a look at the chief complaints'. *NEJM,* 300(10), pp. 556–559.

Serur, E. *et al.* (1997). 'Neoadjuvant chemotherapy in stage IB2 squamous cell carcinoma of the cervix.' *Gynecol Oncol,* 65, pp. 348-56.

Shafer, G. (1976). *A Mathematical Theory of Evidence*. Princeton, NJ :Princeton University Press.

Sharma, G. *et al*. (2015). 'Efficacy of neoadjuvant versus adjuvant versus adjuvant therapy for resectable pancreatic adenocarcinoma: a decision analysis.' *Ann Surg Oncol*. 22, pp. 1229-1237.

Shen, Y.N. *et al.* (2018). 'A preoperative nomogram predicts prognosis of up front resectable patients with pancreatic with pancreatic head cancer and suspected venous invasion.' *HPB,* pp*.* 1-10.

Sherman, W.H. *et al*. (2015). 'Neoadjuvant gemcitabine, docetaxel, and capecitabine followed by gemcitabine and capecitabine/radiation therapy and surgery in locally advanced, unresectable pancreatic adenocarcinoma.' *Cancer*. 121, pp. 673–680

Shimoda, M. *et al.* (2015). 'Randomized clinical trial of adjuvant chemotherapy with S-1 *versus* gemcitabine after pancreatic cancer resection.' *Br J Surg* 102:pp. 746-754.

Shubert, C.R. *et al.* (2016). 'Overall survival is increased among stage III pancreatic adenocarcinoma patients receiving neoadjuvant chemotherapy compared to surgery first and adjuvant chemotherapy: an intention to treat analysis of the National Cancer Database.' *Surgery.* 160(4), pp. 1080-1096.

Shunemann, H. *et al*. (2018) GRADEpro GDT. Available at: https://gradepro.org/product/#about [Accessed 6th September 2018].

Siegel, R.L., Miller, H.D. and Jemal, A. (2017). 'Cancer Statistics, 2017'. *CA Cancer J Clin*, 67, pp. 7-30
Siegel, R.L., Miller, K.D. and Jemal, A. (2016). 'Cancer statistics, 2015.' *CA Cancer J Clin*, 65:pp. 5-29.

Simunovic, M., Gafni, A. and Levine, M. (2004). 'Economics of preoperative radiotherapy with total Mesorectal excision: what can we learn from the Dutch experience?' *Journal of Clinical Oncology,* 22:2(15), pp. 217-219.

Sinn, M. *et al.* (2017). 'CONKO-005: adjuvant chemotherapy with gemcitabine plus erlotinib versus gemcitabine alone in patients after r0 resection of pancreatic cancer: a multicenter randomized phase iii trial.' *J. Clin. Oncol.* 35, pp. 3330–3337.

Sivakumar, S. *et al.* (2017). 'Master regulators of oncogenic KRAS response in pancreatic cancer: an integrative network biology analysis.' *PLoS Med.* 14, e1002223.

Skinner, B.F. (1938). *The behaviour of organisms: An experimental analysis.* Oxford:Appleton-Century.

Small, W. *et al*. (2011). 'Phase II trial of full-dose gemcitabine and bevacizumab in combination with attenuated three-dimensional conformal radiotherapy in patients with localized pancreatic cancer.' *International Journal of Radiation Oncology Biology Physics*. 80(2), pp. 476 – 482

Smeenk, H.G. *et al.* (2007). 'Long-term survival and metastatic pattern of pancreatic and periampullary cancer after adjuvant chemoradiation or observation: long-term results of EORTC trial 40891.' *Ann Surg,* 246, pp. 734–40.

Smith, M.L. (2005). 'Overcoming theory-practice inconsistencies: Critical realism and information systems research.' *Information and Organization:* 16, pp. 191-211.

Smith, R.A. *et al.* (2008). 'The platelet-lymphocyte ratio improves the predictive value of serum CA 19-9 levels in determining patient selection for staging laparoscopy in suspected periampullary cancer.' *Surgery*, 143, pp. 658-666.

Smith, B.J. and Mezhir, J.J. (2014). 'An interactive Bayesian model for prediction of lymph node ratio and survival in pancreatic cancer patients.' *J Am Med Inform Assoc.* 21(2), pp. 203-11.

Snowden, D.J. and Boone, M.E. 'A leader's framework for decision making.' *Harv. Bus. Rev.* 85, pp. 68–76

Solari, F., Liseo, B. and Sun, D. (2008). 'Some remarks on Bayesian inference for one-way ANOVA models.' *Ann Inst Stat Math,* 60, pp. 483–498

Spiegelhalter, D.J. *et al.* (2002). 'Bayesian measures of model complexity and fit.' *Journal of the Royal Statistical Society.* 64**,**4, pp. 583-616.

Spiegellhalter, D.J., Abrams, K.R. and Myles, J.P. (2004). *Bayesian approaches to clinical trials and health-care evaluation*. Chichester: Wiley

Stacey, R.D. (1996). *Strategic management and organizational dynamics.* London: Pitmann Publishing.

Stajduhar, I. and Dalbelo-Basic, B. (2010). 'Learning Bayesian networks from survival data using weighting censored instances.' *Journal of Biomedical Informatics*, 43, pp. 613-622.

Standing, C. *et al*. (2016). 'The paradoxes of telehealth: a review of the literature 2000-2015'. *Syst Res.* 12(1); doi: 10.1002/sres.2442

Star, S.L. (2002). 'Infrastructure and ethnographic practice: working on the fringes.' *Scand J Inform Syst.* 14(2), pp. 6.

Stefandis, D. *et al.* (2006). 'The current role of staging laparoscopy for adenocarcinoma of the pancrease: a review.' *Ann Oncol.* 17, pp. 189-199.

Steele, G.R. (2005). 'Critical thoughts about critical realism'. *Critical Review*: 17(1),pp. 133-154.

Stein, S.M. *et al.* (2016). 'Final analysis of a phase II study of modified FOLFIRINOX in locally advanced and metastatic pancreatic cancer.' *Br J Cancer*, 114(7), pp. 737-743

Stephenson, A.J. *et al*. (2006). 'Preoperative nomogram predicting the 10-year probability of prostate cancer recurrence after radical prostatectomy.' *J Natl Cancer Inst*, 98(10), pp. 715-7

Sterne, J.A.C. *et a*l.(2016). 'ROBINS-I: a tool for assessing risk of bias in non-randomized studies of interventions.' *BMJ.* **355,** i4919.

Stevenson, S.M. *et al.* (2014). 'Cost-effectiveness of neoadjuvant chemotherapy before radical cystectomy for muscle-invasive bladder cancer.' *Urology Oncology:* 32, pp. 1172-1177.

Steyerberg, E.W., Eijkemans, M.J. and Habbema, J.D. (1999). 'Stepwise selection in small data sets: a simulation study of bias in logistic regression analysis.' *J Clin Epidemiol,* 52, pp. 935–942.

Steyerberg, E.W. (2009). 'Clinical prediction models: a practical approach to model development, validation, and updating.' In: Gail, M. *et al* (eds). New York: Springer.

Steyerberg, E.W. *et al.* (2010). 'Assessing the performance of prediction models: a framework for traditional and novel measures.' *Epidemiology,* 21(1), pp. 128-38.

Strasberg, S.M., Linehan, D.C., Hawkins, W.G. (2009). 'The accordion severity grading system of surgical complications.' *Ann Surg*, 250(2), pp. 177-86

Strasberg, S.M. and Fields, R. (2012). 'Left-sided pancreatic cancer: distal pancreatectomy and its variants: radical antegrade modular pancreatosplenectomy and distal pancreatectomy with celiac axis resection.' *Cancer J*, 18, pp. 562–70.

Suker, M. *et al.* (2016). 'FOLFIRINOX for locally advanced pancreatic cancer: a systematic review and patient-level meta-analysis**.'** *Lancet Oncol,* 17(6), pp. 801-810.

Sun, L. and Shenoy, P.P. (2007). 'Using Bayesian networks for bankruptcy prediction: some methodological issues.' *European Journal of operational Research*. 180, pp. 738-753.

Sutton, A.J. and Abrams, K.R. (2001). 'Bayesian methods in meta-analysis and evidence synthesis.' *Stat Methods Med Res,* 10, pp. 277-303.

Sutton, R.S. and Barto, A.G. (1998). *Reinforcement learning: An Introduction. Vol. 1.* Cambridge: MIT Press.

Sutton, R. *et al*. (2002). 'The Surgical Risk Scale as an improved tool for risk-adjusted analysis in comparative surgical audit.' *Br J Surg*, 89(6), pp. 763-8

Tako, A.A. and Robinson, S. (2015). 'Is simulation in health different?' *J Oper Res Soc*. 66(4), pp. 602–14.

Talamonti, M. *et al*. (2006). 'A multi-institutional phase II trial of preoperative full-dose gemcitabine and concurrent radiation for patients with potentially resectable pancreatic carcinoma.' *Ann Surg Oncol*. 13, pp. 150–158.

Tam, V.C. *et al.* (2013). 'Cost-effectiveness of systemic therapies for metastatic pancreatic cancer'. *Curr Oncol*. 20:pp, 90-106.

Tang, Z. and McCabe, B. (2007). 'Developing complete conditional probability tables from fractional data for Bayesian belief networks.' *Journal of Computing in Civil Engineering*. 21(4), pp. 265-276

Tapper, E. *et al.* (2011). 'Staging laparoscopy for proximal pancreatic cancer in a magnetic resonance imaging-driven practice: what's it worth?' *HPB,* 13, pp. 732-737.

Tempero, M.A, Malafa, M.P. and Behrman, S.W. (2014). 'Pancreatic adenocarcinoma, version 2.2014: featured updates to the NCCN guidelines.' *J Natl Compr Canc Netw*. 12, pp. 1083–1093.

Tempero, M.A. *et al.* (2017). 'Pancreatic adenocarcinoma, version 2.2017, NCCN clinical practice guidelines in oncology.' *J Nati Compr Cancer Netw: JNCCN*, 15(8), pp. 1028-61

Tempero, M.A. *et al.* (2019). 'APACT: phase III, multicenter, international, open-label, randomized trial of adjuvant *nab*-paclitaxel plus gemcitabine (*nab*-P/G) vs gemcitabine (G) for surgically resected pancreatic adenocarcinoma.' *Journal of Clinical Oncology* 37:15

The Cancer Genome Atlas Research Network. (2017). 'Integrated genomic characterization of pancreatic ductal adenocarcinoma.' *Cancer Cell,* 32, pp. 185–203.

The Scottish Government. (2016*). Realistic Medicine: Chief Medical Officer for Scotland's Annual Report 2014-15*. Accessible at: https://www2.gov.scot/resource/0049/00492520.pdf (Downloaded 2nd May, 2019).

The Scottish Government. (2017). *Realising Realistic Medicine: Chief Medical Officer for Scotland's Annual Report 2015-16.* Acessible at: https://www.gov.scot/publications/chief-medical-officer-scotland-annual-report-2015-16-realising-realistic-9781786526731 (Downloaded 5th May, 2019).

The Scottish Government**.** (2019). *Personalising Realistic Medicine: Chief Medical Officer for Scotland's Annual Report 2017-18*. Accessible at: https://www.gov.scot/publications/personalising-realistic-medicine-chief-medical-officer-scotland-annual-report-2017-2018/pages/1/ (Downloaded 25th July, 2019)

Tilleman, E.H. *et al.* (2004). 'Limitation of diagnostic laparoscopy for patients with a periampullary carcinoma.' *Eur J Sur Oncol.* 30, pp. 658-662.

Tinchon, C. *et al.* (2013). 'Safety and efficacy of neoadjuvant FOLFIRINOX treatment in a series of patients with borderline resectable pancreatic ductal adenocarcinoma.' *Acta Oncologica.* 52:6, pp. 1231-1233.

Tingstedt, B. *et al*. (2011). 'Pancreatic cancer, healthcare cost, and loss of productivity: a register-based approach.' *World J Surg*, 35(10), pp. 2298-305.

Tol, J.A. *et al.* (2015). 'Metal or plastic stents for preoperative biliary drainage in resectable pancreatic cancer.' *Gut* doi:10.1136/gutjnl-2014-308762

Tonelli, M.R. and Shirts, B.H. (2017). 'Knowledge for Precision Medicine Mechanistic Reasoning and Methodological Pluralism.' *JAMA.*, 318: pp, 1649–1650.

Toomey, P. *et al.* (2009). 'Resection of portovenous structures to obtain microscopically negative margins during pancreaticoduodenectomy for pancreatic adenocarcinoma is worthwhile.' *Am Surg*, 75, pp. 804–9

Topal, B. *et al*. (2007). 'Hospital cost-categories of pancreaticoduodenectomy.' *Acta Chir Belg*, 107, pp. 373–377

Tran-Duy, A. *et al.* (2011). 'A discrete event modelling framework for simulation of long-term outcomes of sequential treatment strategies for ankylosing spondylitis.' *Ann Rheum Dis*, 70, pp. 2111-2118

Truett, J., Cornfield, J. and Kannel, W. (1967). 'A multivariate analysis of the risk of coronary heart disease in Framingham.' *J Chronic Dis* 20(7), pp. 511-24.

Tseng, J.F. *et al*. (2004). 'Pancreaticoduodenectomy with vascular resection (margin status and survival duration).' *J Gastrointest Surg*. 8: pp. 935–950

Tsoukas, H. (2017). 'Don't simplify, complexify: from disjunctive to conjunctive theorizing in organizational and management studies.' *J Manage Studies.* 54(2), pp. 132-53

Tunnicliffe-Wilson, J.C. (1981). 'Implementation of computer simulation projects in health care.' *J Operat Res Soc.* 32(9), 825–32.

Turrini, O. *et al.* (2009). 'Neoadjuvant 5 fluorouracil-cisplatin chemoradiation effect on survival in patients with resectable pancreatic head adenocarcinoma: a ten-year single institution experience.' *Oncology.* 76, 413–419.

Turrini, O. *et al*. (2013). 'Should the portal vein be routinely resected during pancreaticoduodenectomy for adenocarcinoma?' *Ann Surg*, 257, pp. 726–30.

Tzeng, C.W. *et al*. (2012). 'Yield of clinical and radiographic surveillance in patients with resected pancreatic adenocarcinoma following multimodal therapy.' *HPB.* 14, pp. 365–72.

Tzeng, C.W. *et al*. (2014). 'Treatment sequencing for resectable pancreatic cancer: influence of early metastases and surgical complications on multimodality therapy completion and survival.' *J Gastrointest Surg*. 18, pp. 16–24.

UK House of Lords. (2000). Science and Technology – Third Report. Accessible at: https://publications.parliament.uk/pa/ld199900/ldselect/ldsctech/38/3801.htm (Downloaded 4th April, 2018).

Ueno, H. *et al.* (2009). 'A randomised phase III trial comparing gemcitabine with surgery-only in patients with resected pancreatic cancer: Japanese Study Group of Adjuvant Therapy for Pancreatic Cancer.' *Br J Cancer*. 101, pp. 908–15.

Uesaka, K. *et al.* (2016). 'Adjuvant chemotherapy of S-1 versus gemcitabine for resected pancreatic cancer: a phase 3, open-label, randomised, non-inferiority trial (JASPAC 01).' *Lancet,* 388, pp. 248–257.

Ulrich, W. (1983). *Critical Heuristics of Social Planning: A New Approach to Practical Philosophy*. Berne, Switzerland: Paul Haupt.

Ulrich, W. (1987). 'Critical Heuristics of Social Systems Design.' *Eur. J. Oper. Res. 31*, pp. 276–283.

Ulrich, W. (2002). 'Critical Systems Heuristics.' In Daellenbach, H.G. and Flood, R.L. (eds). *The Informed Student Guide to Management Science*. London, UK: Thomson Learning, pp. 72–73

Ulrich, W. (2012). 'Operational research and critical systems thinking—An integrated perspective. Part 1: OR as applied systems thinking.' *J. Oper. Res. Soc. 63*, pp. 1228–1247

Ulrich, W. and Reynolds, M. (2010). 'Critical systems heuristics'. In: Reynolds, M and Holwell, S. (eds). *Systems Approaches to Managing Change: A Practical Guide*. London: Springer, pp.243–292
University College London (UCL). (2019). *ADEPTS study information.* Acessible at: https://www.ucl.ac.uk/surgery/research/situ-trials/adepts-study-information [Accessed June 1st, 2019]

Unno, M. *et al.* (2019). 'Randomized phase II/III trial of neoadjuvant chemotherapy with gemcitabine and S-1 versus upfront surgery for resectable pancreatic cancer (Pre-02/JSAP-05).' *Journal of Clinical Oncology.* 37(4), pp. 189-189

Van Buren, G. *et al.* (2013). 'Phase II study of induction fixed-dose rate gemcitabine and bevacizumab followed by 30 Gy radiotherapy as preoperative treatment for potentially resectable pancreatic adenocarcinoma.' *Ann Surg Oncol*. 20, pp. 3787–3793

Van der Brink, M. *et al.* (2004). 'Cost-utility analysis of preoperative radiotherapy in patients with rectal cancer undergoing total

Mesorectal excision: a study of the Dutch colorectal cancer group.' *Journal of Clinical Oncology,* 22(2), pp. 244-253.

Van Der Gaag, N.A. *et al.* (2010). 'Preoperative biliary drainage for cancer of the head of the pancreas.' *NEJM,* 362(2), pp. 129–137.

Van Geenen, R.C. *et al.* (2001). 'Segmental resection and wedge excision of the portal or superior mesenteric vein during pancreatoduodenectomy.' *Surgery,* 129, 158–63.

Van Houten, J.P, White, R.R. and Jackson, G.P. (2012). 'A Decision Model of Therapy for Potentially Resectable Pancreatic Cancer.' *The Journal of Surg Research*. 174, pp. 222-230.

Van Laethem, J.L. *et al*. (2010). 'Adjuvant gemcitabine alone versus gemcitabine-based chemoradiotherapy after curative resection for pancreatic cancer: a randomized EORTC-40013-22012/FFCD-9203/GERCOR phase II study.' *Journal of Clinical Oncology.* 28(29), pp. 4450-4456.

Van Limburg, M. *et al.* (2011). 'Why business modeling is crucial in the development of eHealth technologies.' *J Med Internet Res.*, 28;13(4):e124.

Van Lent, W.A., VanBerkel, P. and Van Harten, W.H. (2012). 'A review on the relation between simulation and improvement in hospitals.' *BMC Med Inform Decis Mak*. 12:18.

Van Tienhoven, G. (2018). 'Preoperative chemoradiotherapy versus immediate surgery for resectable and borderline resectable pancreatic cancer (PREOPANC-1) : A randomized, controlled, multicenter phase III trial.' *J Clin Oncol.* 36.

Varadhachary, G.R. *et al*. (2008). 'Preoperative gemcitabine and cisplatin followed by gemcitabine-based chemo-radiation for resectable adenocarcinoma of the pancreatic head.' *J Clin Oncol* 26, pp. 3487-3495.

Vargas, C. *et al.* (1995). 'Diagnostic laparoscopy: a 5year experience in a hepatology training programme.' *Am J Gastroenterol,* 90, pp. 1258-1262.

Velez-Serrano, J.F. *et al.* (2017). 'Prediction of in-hospital mortality after pancreatic resection in pancreatic cancer patients: A boosting approach via a population-based study using health administrative data.' *PLoS One.* 7:12(6):e0178757

Velikova, M. *et al.* (2014). 'Exploiting causal functional relationships in Bayesian network modeling for personalized healthcare.' *Int Journal of Approximate Reasoning*. 55, pp. 59-73.

Venkat, R. *et al.* (2012). 'Laparoscopic distal pancreatectomy is associated with significantly less overall morbidity compared to the open technique: a systematic review and meta-analysis.' *Ann Surg*. 255, pp. 1048–1059

Vento, P. *et al.* (2007). 'Impact of preoperative chemoradiotherapy on survival in patients with resectable pancreatic cancer.' *World Journal of Gastroenterology : WJG*. 13(21), pp. 2945-2951.

Verbeke, C. (2016). 'Morphological heterogeneity in ductal adenocarcinoma of the pancreas - Does it matter?' *Pancreatology*. 16, pp. 295–301.

Verduijn, M. *et al.* (2007). 'Prognostic Bayesian networks I: rationale, learning procedure, and clinical use.' *Journal of Biomedical Informatics'* pp. 609-618.

Verduijn, M. *et al.* (2017). 'Prognostic Bayesian networks I: rationale, learning procedure, and clinical use.' *Journal of Biomedical Informatics.* pp. 609-618.

Vergote, I. *et al.* (2010). 'Neoadjuvant chemotherapy or primary surgery in stage IIIC or IV ovarian cancer'. *NEJM* 363, pp. 943-53.

Versteijne, E. *et al*. (2018). 'Meta-analysis comparing upfront surgery with neoadjuvant treatment in patients with resectable or borderline resectable pancreatic cancer.' *The British Journal of Surgery*. 105(8), pp. 946-958.

Vincent, A. *et al.* (2011). 'Pancreatic cancer.' *Lancet.* 378, pp. 607-620

Vitale, A. *et al.* (2010). 'Use of Sorafenib in patients with hepatocellular carcinoma before liver transplantation: a cost-benefit analysis while awaiting data on Sorafenib safety.' *Hepatology:* 51(1), pp. 165-173.

Vittinghoff, E. and McCulloch, C.E. (2007). 'Relaxing the rule of ten events per variable in logistic and Cox regression'. *Am J Epidemiol* 165, pp. 710–718.

Von Neumann, J. and Morgenstern, O. (1953). *Theory of Games and Economic Behavior.* 3rd ed. Princeton: Princeton University Press.

Wade, V, Eliott, J.and Hiller, J. (2014). 'Clinician acceptance is the key factor for sustainable telehealth services.' *Qual Health Res.* 24(5), pp. 682–694.

Walczak, S. and Velanovich, V. (2017). 'An evaluation of Artificial Neural Networks in predicting pancreatic cancer survival.' *J Gastrointest Surg* 21, pp. 1606-1612

Wald, A. (1980). 'A Method of estimating plane vulnerability based on damage of survivors'. Centre for Naval Analysis.

Waldrop, M.W. (1992). *Complexity: the emerging science at the edge of order and chaos.* New York: Simon and Schuster

Walley, P. (1991). *Statistical Reasoning with Imprecise Probabilities.* London: Chapman and Hall.

Walley, P. (1996). 'Inferences from multinomial data: Learning about a bag of marbles' *J. Roy. Statist. Soc. Ser.* 58 ; pp. 3–57.

Wasserman, L. (2006). 'Frequentist Bayes is objective.' *Bayesian Analysis:* 1:451-456.

Wang,, P. (1993). 'Belief revision in probability theory.' In: Kaufmann, M. *Proceedings of the Ninth Conference on Uncertainty in Artificial Intelligence,* San Mateo, CA, pp. 519–526.

Wang, P. (2001). 'Confidence as higher-order uncertainty.' In: Proceedings of the Second International Symposium on Imprecise Probabilities and Their Applications, Ithaca, New York, pp. 352–361.

Wang, P. (2004). 'Limitations of Bayesianism.' *Artificial Intelligence, pp.* 158:97-106

Waters, J.A. *et al*. (2010). 'Robotic distal pancreatectomy: cost effective?' *Surgery.* 148, pp. 814–823

Weinstein, M.C. *et al.* (1980). *Clinical Decision Analysis.* Philadelphia: WB Saunders Company.

Weiser, M.R. *et al*. (2008). 'Individualized prediction of colon cancer recurrence using a nomogram.' *J Clin Oncol*. 26(3), pp. 380-5.

Wellner, U.F. *et al.* (2010). 'A simple scoring system based on clinical factors related to pancreatic texture predicts postoperative pancreatic fistula preoperatively.' *HPB*, 12(10), pp. 696-702.

Wilkowski, R. *et al.* (2009). 'Chemoradiotherapy with concurrent gemcitabine and cisplatin with or without sequential chemotherapy with gemcitabine/cisplatin vs chemoradiotherapy with concurrent 5-fluorouracil in patients with locally advanced pancreatic cancer--a multi-centre randomised phase II study.' *Br J Cancer*. 101, pp. 1853–1859.

Williamsson, C. *et al.* (2015). , 'Impact of a fast-track surgery programme for pancreaticoduodenectomy.' *Br. J. Surg.* 102(9), pp. 1133–1141.

Winter, J.M. *et al*. (2012). 'Survival after Resection of Pancreatic Adenocarcinoma: Results from a Single Institution over Three Decades.' *Ann Surg Oncol*, 19: 169.

Woermann, M. (2010). *A complex Ethics: Critical Complexity, Deconstruction and Implications for Business Ethics*.. Stellenbosch, South Africa: University of Stellenbosch Press.

Woermann, M. and Cilliers, P. (2012). 'The ethics of complexity and the complexity of ethics.' *S. Afr. J. Philos. 31*, pp. 447–463.

Wolff, R.F. *et al.* (2014). 'PROBAST: A Tool to Assess the Risk of Bias and Applicability of Prediction Model Studies.' *Ann Intern Med* 170, pp. 51–58.

Wolff, R.A. *et al*. (2017). 'Neoplasms of the exocrine pancreas.' In Bast R.C. et al. (eds). *Holland-Frei Cancer Medicine*. Hoboken, NJ: Wiley Blackwell; pp, 1129-1150

Wolpert, M. and Rutter, H. (2018). 'Using flawed, uncertain, proximate and sparse data in the context of complexity: learning from the case of child mental health'. *BMC Medicine* 16, 82

Wong, M.C.S. *et al*. (2017). 'Global temporal patterns of pancreatic cancer association with socioeconomic development.' *Sci Rep* 7:3165

Woods, B. *et al. NICE DSU technical support document 19. Partioned survival analysis for decision modelling in healthcare: a critical review.* Report by the Decision Support Unit. Available from: http://www.nicedsu.org.uk. [Accessed May 4, 2019].

World Cancer Research Fund/American Institute for Cancer Research. *Diet, Nutrition, Physical Activity and Cancer: a Global Perspective*. Continuous Update Project Expert Report 2018. Available at httP;www.dietandcancerreport.org (Accessed 7th June, 2019)

Wray, C.J. *et al.* (2005). 'Surgery for pancreatic cancer: recent controversies and current practice'. *Gastroenterology*, 128, pp. 1626-1641

Wu, B. *et al.* (2015). 'Treatment of moderate rheumatoid arthritis with different strategies in a health resource-limited setting: a cost-effectiveness analysis in the era of biosimilars**.'** *Clin Exp Rheumatol*, 33, pp. 20-26

Xu, C.P. *et al*. (2014). 'Effect of chemoradiotherapy and neoadjuvant chemoradiotherapy in resectable pancreatic cancer: a systematic review and meta-analysis.' *J Cancer Res Clin Oncol,* 140, pp. 549-59.

Xu, J. *et al.* (2017). 'A nomogram based on preoperative inflammatory markers predicting the overall survival of pancreatic ductal adenocarcinoma.' *Journal of Gastroenterology*, 32, pp. 1394-1402

Yang, Z.L., Wang, J., Bonsall, S. and Fang, Q.G. (2009). 'Use of fuzzy evidential reasoning in maritime security assessment.' *Risk Analysis*, 29(1), pp. 95-120

Yamamoto, K.N, *et al.* (2017). 'Personalized Management of Pancreatic Ductal Adenocarcinoma Patients through Computational Modeling.' *Cancer Res.* 15;77(12), pp. 3325-3335.

Yoshitomi, H. *et al.* (2008). 'A randomized phase II trial of adjuvant chemotherapy with uracil/tegafur and gemcitabine versus gemcitabine alone in patients with resected pancreatic cancer.' *Cancer.* 113, pp. 2448–56.

Yu, X.Z. *et al.* (2014). 'Benefit from synchronous portal-superior mesenteric vein resection during pancreaticoduodenectomy for cancer: a meta-analysis.' *Eur J Surg Oncol*. 40, pp. 371–378.

Zachariadis, M., Scott, S. and Barrett, M. (2010). *Exploring critical realism as the theoretical foundation of mixed-method research: evidence from the economics of IS innovations.* Cambridge: Cambridge Judge Business School.

Zadeh, L.A. (1962). 'From circuit theory to systems theory.' *Proc IRE*, 50, pp. 856-65.

Zadeh, L.A. (1965a). 'A Fuzzy sets'. *Inform Control*. 8, 338-53.

Zadeh, L.A. (1965b). 'Fuzzy sets and systems.' In: Fox J, (ed). *System theory.* Brooklyn (NY): Polytechnic Press, pp. 29-39.

Zadeh, L.A. (1969). 'Biological application of the theory of fuzzy sets and systems'. In: Proctor, L.D. (ed). *Proceedings of an International Symposium on Biocybernetics of the Central Nervous System*. Boston: Little Brown and Co., pp. 199-206.

Zadeh, L.A. (1996a). 'The birth and evolution of fuzzy logic, soft computing and computing with words: a personal perspective.' In: Klir, G.J. and Bo, Y. (eds). *Fuzzy sets, fuzzy logic, and fuzzy systems* Singapore: World Scientific, pp, 811-9.

Zadeh, L.A. (1996b). 'Fuzzy logic computing with words.' *IEEE Trans Fuzzy Syst,* 4(2), pp. 103-11

Zarin, W. *et al*. (2017). 'Characteristics and knowledge synthesis approach for 456 network meta-analyses: a scoping review.' *BMC Med.* 15, 3.

Zhan, H.X. *et al.* (2017). 'Neoadjuvant therapy in pancreatic cancer: a systematic review and meta-analysis of prospective studies.' *Cancer Med.* 6, pp. 1201–1219.

Zhang, Z. *et al.* (2007). 'Bayesian analysis of longitudinal data using growth curve models.' *International Journal of Behavioral Development* 31(4), pp. 374-383.

Zhang, X. *et al*. (2017). 'Big Data Science: Opportunities and Challenges to Address Minority Health and Health Disparities in the 21st Century.' *Ethnicity & Disease*. 27, pp. 95–106.

Zhang, X. (2018). 'Application of discrete event simulation in health care: a systematic review.' *BMC Health Services Research.* 18, pp. 687.

Appendix A

An Overview of the Current Genomic Understanding of Pancreatic Cancer

This appendix provides further detailed context for section 2.1.1 where the pathology of pancreatic cancer is discussed. The purpose of this appendix is to provide context for the later discussions surrounding precision medicine (gene targeted therapy) by demonstrating that in the near future a molecular taxonomy could help to partly inform clinical decision making (McGuigan *et al*., 2018; Collisson *et al*., 2019).

Mutations in the KRAS oncogene and telomere shortening have been identified in low grade pancreatic intraepithelial neoplasms suggesting that these are early changes within the pathway (Feldmann *et al*., 2007; Hruban *et al*., 2008) with mutation in p16, CDNK27, p53 and SMAD4 appearing later in both high grade pancreatic intraepithelial neoplasms and pancreatic adenocarcinoma where the rate of KRAS mutation was also been found to increase (Mohammed *et al*., 2014; Löhr *et al*., 2005). Additionally abnormalities in sonic hedgehog pathway and notch signaling have been implicated in the development of PDAC with 80% of these mutations thought to be sporadic (Midha *et al*., 2016; Vincent *et al*., 2011). Over recent years our understanding of the genomic aberrations characteristic of pancreatic cancer has expanded and is summarised in table Ai.

Table Ai: Summary of genomic aberrations characteristic of pancreatic cancer (Bailey *et al*., 2016).

Pathway / Process	Gene Mutated
SWI/SNF	ARID1A, SMARCA4, ARID1B, PBRM1
Chromatin	KDM6A, MLL3, MLL2, SETD2
WNT signalling	RNF43, TLE4, MARK2
NOTCH signalling	JAG1, BCORL1, NF2, FBXW7
TGFβ signalling	SMAD3, TGFBR1, ACVR1B, SMAD4, TGFBR2, ACVR2A
KRAS	KRAS, MAPK4
ROBO SLIT pathway	ROBO1, SLIT2, ROBO2, MYCBP2
RNA processing	RBM10, U2AF1, SF3B1
Cell cycle	CDKN2A, TP53BP2, TP53
DNA repair	BRCA1, ATM, ATF2, BRCA2, PALB2

Gene analysis has also revealed 32 genetic mutations associated with PDAC that resulted in the recognition of 4 subgroups: squamous, pancreatic progenitor, immunogenic and aberrantly differentiated endocrine exocrine (ADEX) (Bailey *et al*., 2016). Each sub-type was associated with different genomic signatures, histopathological findings and associated prognosis (Bailey *et al*., 2016). However, further large-scale analysis of the molecular characteristic of PDAC revealed that previously histopathologically indistinguishable tumours, and even their subtypes, harbored substantial molecular differences and overlaps that could have biological and clinical relevance (Collission *et al*., 2019) as summarised in table Aii.

Table Aii: Molecular subtyping of PDAC

Subtypes	Biological Insight	Clinical Relevance
Classical Quasi-mesenchymal Exocrine-like (Collisson *et al*., 2011; Collission *et al*., 2019)	ATCC PDAC cell lines have an absence of Exocrine-like subtype GATA6 and KRAS have a specific function in Classical subtype	Quasi-mesenchymal subtype have poorer survival compared to Classical subtype but are more sensitive to gemcitabine. Classical subtype is more sensitive to erlotinib
Epithelial Basal-like Classical Stromal subtypes: active and normal (Moffitt *et al*., 2015; Collission *et al*., 2019)	Subtype signature is maintained in metastases with Basal-like subtype in the majority of metastases while lung metastases are associated with Classical subtype	Basal-like subtype and active stroma in Classical type associated with poor survival. Basal-like subtype benefits from adjuvant chemotherapy.
Squamous Immunogenic Pancreatic Progenitor ADEX (Bailey *et al*., 2016; Collission *et al*., 2019)	Squamous subtype enriched for inflammation, cell proliferation, metabolic reprogramming and epigenic downregulation of endodermal genes. Squamous and Immunogenic subtypes both enriched for immune signalling Squamous subtype associated with adenosquamous histology Pancreatic Progenitor subtype associated	Poor survival in squamous subtype Subtype-specific therapeutic targets including: cell cycle and metabolic inhibitors and immunomodulation.

	with colloid in IPMN	
Multiple subtypes overlapping Integrated classifier (The Cancer Genome Atlas Research Network, 2017; Collission *et al*., 2019)	Overlap amongst previously defined subtypes mTOR signalling elevated in KRAS-wild type tumours Multiple different KRAS mutations identified within individual tumours	Treatment targets for KRAS-wild type tumours
Hedgehog (associated with: Quasi-mesenchymal, Basal and activated stromal or squamous) NOTCH (associated with Endocrine-like, normal stroma or ADEX subtypes) Cell Cycle (associated with Classical or Pancreatic Progenitor) (Sivakumar *et al*., 2017; de Santiago *et al.*, 2017; Collission *et al*., 2019)	Strong immunological differences between subtypes	Hedgehog associated with poor prognosis. Hedgehog and cell cycle subtype associated targets for immunotherapy Cell cycle subtype a potential target for metabolic therapy
Classical Quasi-mesenchymal Exocrine-like (Noll *et al*, 2016; Collission *et al*., 2019)	CYP450 gene expression in Exocrine-like tumours	Potential clinical relevance of gene CYP450
High cellularity (Pure Classical, Immune Classical, Pure Basal-like) All cellularities (Pure Classical, Immune Classical, Pure Basal-like, Stromal active and Desmoplastic)		Pure Basal-like associated with poor prognosis. Hypothetical subtype specific therapies that target immune avoidance.

(Puleo *et al*., 2018; Collission *et al*., 2019).		

However, as is discussed in section 2.1.1, this burgeoning deeper understanding of pancreatic cancer at a molecular level has not yet resulted in histopathological classifications to inform clinical decisions as they do in other cancer types (Collisson *et al*., 2019). Instead what this knowledge has highlighted is the complex and highly heterogeneous nature of this disease at molecular level (McGuigan *et al*., 2018; Collisson *et al*., 2019) that is intrinsically linked with uncertainty, ambiguity and complexity that surrounds the management of pancreatic cancer.

Appendix B

Overview and Critical Analysis of the Evidence Base for Surgical Practices

This appendix provides further detail to the discussion in section 2.1.3. Surgery is the only potentially curative treatment for pancreatic cancer and studies often report resection rates as well as rates of R0 resection. The purpose of this appendix is to add further context and strengthen the points being made in section 2.1.3 that even when the decision to operate is made, a sequence of subsequent intra-operative decisions are required within a context of uncertainty.

Arterial invasion is seen as a contraindication to surgery (Gall *et al*., 2015) but several observational studies have demonstrated that this is not only technically possible in cases of locally advanced disease but is also associated with better survival outcomes compared to those who did not undergo resection (Mollberg *et al*., 2015). However, a pancreaticoduodenectomy with arterial resection has greater than 5times the risk of perioperative mortality and poorer survival outcomes at 1 and 3years compared to pancreaticoduodenectomy without arterial resection (Yu *et al*., 2014; Mollberg *et al*., 2015). Therefore the International Study Group for Pancreatic Surgery (ISGPS) does not recommend arterial resection on a routine basis (Bockhorn *et al*., 2014) with others advocating surgical exploration where arterial involvement is suspected on imaging and palliative pathway where this is verified.

Venous resection of the Superior Mesenteric Vein-Portal Vein (SMV-PV) is now often performed in specialist high volume pancreatic centres where venous invasion is deemed probable (McGuigan *et al*., 2018). The original hypothesis underpinning this operative approach was that disease recurrence was caused by inadequate local therapy and that outcomes could be improved with better tumour clearance (Wray *et al*., 2005). What is contested is whether SMV-PV invasion, and poor outcomes, represents more aggressive tumour biology or, in light of the accepted fact that positive margins result in poor outcomes comparable to those treated non-operatively, SMV-PV invasion reflects location rather than biology (Wray *et al*., 2005). It follows that SMV-PV resection would improve survival outcomes.

No difference has been demonstrated in terms of operative morbidity and mortality when Pancreatoduodenectomy with venous resection is compared to standard Pancreatoduodenectomy (Fuhrman *et al*., 1996; Leach *et al*., 1998; Bachellier *et al*., 2001; Nakagohri *et al*., 2003; Howard *et al*., 2003; Nakano *et al*., 2002; Tseng *et al*., 2004; Oettle *et al*., 2013; Chakravarty *et al*., 2010; Murakami *et al*., 2013; Moldovan *et al*., 2012; Muller *et al*., 2009; Kaneoka *et al*., 2009; van Geenen *et al*., 2001; Riediger *et al*., 2006; Toomey *et al*., 2009; Chau *et al*., 2010). Although the former has been associated with prolonged operative time and blood loss (Oettle *et al*., 2013). However retrospective and observational studies have not concluded a survival advantage at 1 and 3years with venous resection although outcomes are no worse (Fuhrman *et al*., 1996; Leach *et al*., 1998; Bachellier *et al*., 2001; Nakagohri *et al*., 2003; Howard *et al*., 2003;

Nakano *et al*., 2002; Tseng *et al*., 2004; Oettle *et al*., 2013) whilst a more recent French study reported improved survival outcomes with venous resection (Turrini *et al*., 2013). The lack of RCTs mean that such studies are at risk of selection bias (McGuigan *et al*., 2018). What is not contested is that SMV-PV involvement is no longer a contraindication to resection and the ultimate goal of surgery is to achieve negative resection margins.

The benefits of modified surgical techniques are also ambiguous. Laparoscopic pancreatoduodenectomy and laparoscopic distal pancreatectomy have been shown to have comparable morbidity and mortality with open procedures but have not been proven to be conclusively superior (Asbun *et al*., 2012; Pericleous *et al*., 2012; Venkat *et al*., 2012). Meta-analysis have reported no difference in R0 resection rates between open and laparoscopic approach (Pericleous *et al*., 2012) although one meta-analysis has reported reduced blood loss and length of hospital stay (Pericleous *et al*., 2012). Once again however such observational studies are subject to selection bias (Gall et al., 2015). A modified technique for distal pancreatectomy to include radical antegrade modular pancreatosplenectomy reported R0 resection rates of 81% (Strasberg *et al*., 2012) and radical distal pancreatectomy with coeliac axis resection reported R0 resection rates up to 91% (Jing *et al*., 2013; Hirano *et al*., 2007). However this was associated with a 48% to 54% morbidity rate and median survival times reported with this approach ranged from 21months (Hirano *et al*., 2007) to only 9.25months (Jing *et al*., 2013).

Appendix C

Overview and Critical Analysis of the Evidence Base for Adjuvant Chemoradiotherapy

This section provides a more detailed critical analysis of the main studies underpinning the evidence base for adjuvant chemoradiotherapy as discussed in section 2.1.4.

5-FU was first established as the standard of adjuvant therapy in 1980s when the Gastrointestinal Tumour Study Group (GITSG) reported improved overall survival (21 *versus* 11 months; *p value* =0.03) and 2 year survival (43% *versus* 18%) for patients receiving adjuvant therapy compared to observation alone (Kalser & Ellenberg, 1985). However, the results of this small (n=43), unpowered study where 25% of participants did not receive treatment for 10 or more weeks post-operatively, were not replicated in the larger European Organisation for Research and Treatment of Cancer (EORTC) trial (n=218) which demonstrated no improvement in overall or 5-year survival (Klinkenbijl *et al*., 1999). However, both studies were criticised for using sub-optimal dose of radiotherapy with EORTC trial administering split dose radiotherapy, potentiating tumour regrowth (Saif, 2013; Neoptolemos *et al*., 2019). This trial offered no prospective assessment of surgical margins or maintenance dose of chemotherapy. Furthermore, 20% in the treatment arm did not receive treatment and the study population included peri-ampullary cancer with no subset analysis of PDAC (Wray *et al*., 2005; Saif, 2013).

ESPAC-1 (n=289) championed chemotherapy alone, demonstrating improved overall survival compared to no adjuvant therapy (20.1 v 15.5 months, p=0.009) but worse overall survival with chemoradiotherapy (15.9 *versus* 17.9 months, *p value*=0.05) (Neoptolemos *et al*., 2004; Regine *et al*., 2008). Again this trial has been criticised for sub-optimal dose and delivery of radiotherapy as 62% of patients experienced local recurrence with 35% of these experiencing local recurrence as the only site of initial failure (Wray *et al*., 2005). There has also been criticism of the methods of randomisation used, confusing a 2x2 fractional design, bias in longer time-to-treatment and the inclusion of R1 patients in the chemoradiotherapy group (Saif, 2013; Neoptolemos *et al*., 2019) with lack of quality assurance for radiotherapy planning, imaging and pathology assessment of resection margins (Wray *et al*., 2005). Despite this these findings were comparable with the Radiation Therapy Oncology Group (RTOG) 9704 trial which demonstrated the benefit of adding gemcitabine to 5-FU based chemoradiotherapy compared to 5-FU based chemoradiotherapy plus 5-FU-FA chemotherapy which gave a median survival of 20.5months and 16.8months respectively (Regine *et al*., 2008) but at a cost of more hematological grade 4 toxicities (Safi, 2013). This trial however lacked surgical and pathological standardisation.

Appendix D

Overview and Critical Analysis of the Evidence Base for Adjuvant Chemotherapy

This section provides a more detailed critical analysis of the main studies underpinning the evidence base for adjuvant chemotherapy as discussed in section 2.1.4.

The landmark CONKO-001 trial (n=368) demonstrated improved survival outcomes with adjuvant gemcitabine compared to observation alone following surgical resection (Neoptolemos *et al*., 2010). However, although disease free survival improved from 6.9 months to 13.4 months, overall survival only improved slightly from 20.2months to 22.8months with adjuvant gemcitabine (McGuigan *et al.*, 2018). The CONKO-005 trial compared adjuvant gemcitabine following R0 resection with and without the addition of erlotinib but this was not found to improve survival time further (Sinn *et al*., 2017).

The larger ESPAC-3 trial (Neoptolemos *et al*., 2010) showed that adjuvant gemcitabine was not superior to adjuvant 5-FU in terms of survival but gemcitabine had a lower toxicity profile (Neoptolemos *et al*., 2017). The superiority of gemcitabine was further challenged by the Japanese Adjuvant Study Group of Pancreatic Cancer (JASPAC-01) trial (n=385) which introduced S1 (a combination of tegafur, gimeracil and oteracil) as an alternative to gemcitabine. S1 demonstrated improved 2year overall and disease free survival (70%

versus 53%; 49% *versus* 29% respectively) (Uesaka *et al*., 2016). Although R0 resection was stated as the inclusion criteria, 13% of the included cases were R1 (Neoptolemos *et al*., 2019). The application of these findings to Caucasian patients has been questioned due to the possibility that polymorphisms in cytochrome CYP2A6 could result in higher plasma concentrations of 5-FU causing more severe gastrointestinal side effects (Saif *et al*., 2009).

Drawing on the success of dual therapy with gemcitabine and capecitabine in treating advanced and metastatic pancreatic cancer, the ESPAC-4 trial compared gemcitabine monotherapy to this dual-therapy combination in patients who had undergone R0 and R1 resection (Neoptolemos *et al*., 2017). This was an important trial as it was more inclusive with 60% of the study population having R1 resection, 80% of tumours were lymph node positive, and only 42% had a performance status of 0. Non post-operative CT scans and CA19-9 levels were also considered as part of the patient assessment. Results demonstrated that dual therapy had an improved overall median survival of 28months compared to 25.5months with 5year survival of 30%. Although grade3 or 4 neutropenia was more common in the dual therapy arm (38% *versus* 24%) there were fewer overall infective manifestations (3% *versus* 7%). Gemcitabine-capecitabine was therefore recommended over other adjuvant regimes in the 2017 American Society of Clinical Oncology (ASCO) Clinical Practice Guidelines for potentially curable pancreatic cancer (Khorana *et al.*, 2017).

Interim analysis has recently been reported of the APACT trial comparing adjuvant gemcitabine and adjuvant gemcitabine with nab-paclitaxel (Tempero *et al*., 2019). Although there was no difference in disease free survival, overall survival supported the latter but at the cost of a higher toxicity profile.

Appendix E

Overview and Critical Analysis of the Evidence Base for Neoadjuvant Therapy

This appendix supports section 2.1.5 by providing a more detailed critical analysis of existing prospective phase II neoadjuvant therapy drug trials for resectable, borderline resectable and locally advanced disease. Each disease stage has different anticipated outcomes that could affect treatment selection and decision making. The existing body of evidence will therefore be explored for each disease stage separately.

Evans *et al.* (1992) undertook one of the earliest feasibility studies of neoadjuvant approach using 50.4Gy of radiation with 5-fluorouracil (5-FU). They demonstrated that 11 of the 28 participants did not go on to have resection due to disease progression on re-staging. The Eastern Cooperative Oncology Group (ECOG) trial used mitomycin and 5-FU chemoradiation on 53 patients, only 24 of whom went on to have resection (Hoffman *et al.,* 1998). The median survival of those who underwent neoadjuvant approach followed by resection was 15.7 months and the median overall survival of all participants was 9.7 months (Hoffman *et al.,* 1998). Therefore, by intention to treat analysis, neoadjuvant approach had questionable benefit (Hoffman *et al.,* 1998; Royall *et al.,* 2015).

Between 1998 and 2001 a phase II study undertaken at the M.D. Anderson Cancer Centre used systemic gemcitabine and external-

beam radiation as neoadjuvant approach on 86 patients with potentially resectable pancreatic cancer (Evans *et al.,* 2008). 64 patients underwent surgical resection after neoadjuvant treatment with overall survival of 34 months and disease free survival of 28.6 months compared to 7.1 months overall survival and 13.2months disease free survival in those deemed unfit for surgery at re-staging (Evans *et al.,* 2008). A follow-up trial at the same institution using adjuvant gemcitabine and cisplatin, followed by gemcitabine and radiotherapy on 90 patients reported 28.3 month disease free survival and 35months overall survival in the 52 patients who underwent neoadjuvant therapy followed by surgery (Varadhachary *et al.,* 2008). This study demonstrated no increase in surgical complications in the neoadjuvant group and also highlighted that 19% of participants were spared costly and potentially morbid surgery due to identification of more aggressive tumour biology demonstrated by disease progression at restaging and/or declining performance status (Asare *et al.,* 2016; Royall *et al.,* 2015; Varadhachary *et al.,* 2008). However, others have contested this interpretation arguing instead that neoadjuvant approach delayed surgery resulting in loosing the window of resectability.

Estrella *et al.* (2012) summarised the M.D. Anderson experience between 1999-2007 and reported 33.5 month overall survival for patients who underwent neoadjuvant therapy prior to surgery compared to 26.5 month overall survival for those who underwent upfront surgery. However, this summary failed to incorporate patients who underwent neoadjuvant therapy but did not undergo resection in an intention-to-treat analysis (Royall *et al.,* 2015).

Furthermore resection rate varied from 57% to 74% reflecting a high drop-out rate (Royall *et al.,* 2012; Estrella *et al.,* 2012*)*. The time frame also covers a period where the neoadjuvant regime and protocol determining timing from diagnosis to preoperative restaging changed (Royall *et al.,* 2012). Therefore the summary study from which these conclusions are drawn covers a highly heterogeneous, non-randomised patient population (Royall *et al.,* 2012). Furthermore, a universal definition of borderline and locally advanced disease was only adopted in 2010, which impacts the interpretation of studies prior to this date (Tempero *et al*., 2017; Callery *et al*., 2009; Raufi *et al*., 2019). Whilst this has helped to improve reporting there still exists inter-institutional variation impacting decisions regarding resectability (Raufi *et al*., 2019) and many studies combine resectable, borderline and locally advanced cases within the neoadjuvant treatment arm.

Neoadjuvant therapy for resectable pancreatic cancer

Neoadjuvant therapy for resectable pancreatic cancer remains highly controversial (Raufi *et al*., 2019). No completed prospective phase III RCTs have yet compared neoadjuvant therapy to traditional upfront surgery and adjuvant therapy. However, preliminary results from the Prep-02/JSAP05 trial, which compared neoadjuvant gemcitabine and S1 (n=182) with upfront surgery followed by S1 (n=180), have recently been presented in abstract form and report a median overall survival of 36.7months and 26.6months respectively (P=0.015) (Unno *et al*., 2019). However grade 3 and 4 toxicities were higher in the neoadjuvant arm (78%) but there was no statistically significant

difference between resection, R0 or morbidity rates. As with the JASPAC-01 trial exploring S1 and gemcitabine monotherapy (Uesaka *et al*., 2016), considering the potentially increased risk of toxicity in Caucasians (Saif *et al*., 2009), the application of these findings to this ethnic group remains to be seen.

The NCCN recommends neoadjuvant chemotherapy for patients with resectable PDAC if they have high-risk features that include: elevated Ca19-9, large tumours, large regional nodes and disease-related symptoms that are rated as being severe (NCCN, 2018). ASCO guidelines recommend neoadjuvant therapy for patients with resectable PDAC only in the presence of reversible comorbidities that would delay surgery (Khorana et al., 2016). The main prospective phase II trials are outlined in table Ei. These small, mainly non-randomised trials seem to suggest a benefit with neoadjuvant strategy with improved overall median survival even where resection rates are reported as being lower when compared to upfront surgery strategy. A further unresolved question regarding neoadjuvant therapy for resectable pancreatic cancer is the role of chemoradiotherapy (Raufi *et al*., 2019). Thus far chemoradiotherapy does not appear to outperform chemotherapy (Raufi *et al*., 2019). However, as previously mentioned patient populations are small and non-comparable across these studies therefore much of the debate surrounding the role of neoadjuvant therapy for resectable pancreatic cancer remains unsettled.

Table Ei: Summary of Prospective Phase II Trials of Neoadjuvant Therapy for Resectable Pancreatic Cancer

Trial	n=	Single or Multicentre trial	Randomisation	Neoadjuvant Treatment Regime: CRT= chemoradiotherapy; CT= chemotherapy	Resection Rate (%)	R0 Resection Rate (%)	Median Overall Survival in months
Evans *et al.* (2008)	86	Single	No	CRT: 7 weekly intravenous infusions of gemcitabine 400 mg/m^2 plus radiation therapy (30 Gy in 10 fractions over 2 weeks).	64/86 (74)	60/64 (94)	22.7
Golcher *et al.* (2015)	66	Multicentre	Yes	CRT: 300 mg/m^2 gemcitabine and 30 mg/m^2 cisplatin on days 1, 8, 22, and 29 of radiotherapy (1.8 Gy to 55.8 Gy (tumor) or 50.4 Gy (regional lymph nodes), planning target volume ≤ 800 ml).	19/33 (58)	17/19(89)	17.4
				Upfront surgery. Adjuvant chemotherapy also given as per CONKO-001 study protocol.	23/33 (70)	16/23 (70)	14.4
Heinrich *et al.* (2008)	28	Single	No	CT: four biweekly cycles of gemcitabine 1,000 mg/m^2 and cisplatin 50 mg/m^2	25/28 (89)	20/25 (80)	26.5
Hong *et al.* (2014)	50	Multicentre	No	CRT: Proton beam therapy 240-MeV protons generated from a cyclotron delivered using 3D passively scattered protons. Most commonly, 3 fields were used, with 2 fields being treated per day. Capecitabine (1650 mg/m^2 divided twice daily) given Monday to Friday for 2 weeks for each dose level. Adjuvant gemcitabine chemotherapy for 6 months starting 4 to 10 weeks post surgery.	37/50 (74)	31/37 (84)	17
Joensuu *et al.* (2004)	33	Single	No	CRT: Gemcitabine intravenous infusion twice weekly was tested at 3 dose levels: 20, 50, and 100 mg/m^2. Radiation dose 50.4 Gy in 28 fractions.	21/28		25
LeScodan *et al.* (2009)	41	Multicentre	No	CRT: concurrent radiotherapy (50 Gy within 5 weeks) and chemotherapy 5-fluorouracil (300 mg/m^2/day, 5 days/week, weeks 1-5) and cisplatin (20 mg/m^2/day, days 1-5	26/41 (63)	21/26 (80.7)	9.4

				and 29-33).			
O'Reilly *et al.* (2014)	38	Single	No	CT: four cycles of intravenous infusion gemcitabine 1000mg/m^{2}and oxaliplatin 80 mg/m^2, every 2 weeks. Adjuvant gemcitabine intravenous infusion: 5 cycles (1000 mg/m^2 day 1, 8, 15 every 4 weeks).	27/38 (71)	20/27 (74%)	27.2
Palmer *et al.* (2007)	50	Single	Yes	CT: gemcitabine (1000 mg/m^2) every 7 days for 43 days (n=24)	9/24 (38)	6/9 (75%)	9.9
				CT: gemcitabine (1000 mg/m^2) and cisplatin (25 mg/m^2) (n=26)	18/26 (70)	12/18 (75)	16.6
Pister *et al.* (2002)	37	Single	No	CRT: 30 Gy external-beam radiation therapy and concomitant weekly 3-hour infusions of paclitaxel (60 mg/m^2)	20/37 (54)		12
Talamonti *et al.* (2006)	20	Multicentre	No	CRT: three cycles of gemcitabine (1000 mg/m^2 intravenously), with radiation during the second cycle (36 Gy in daily 2.4-Gy fractions)	17/20 (85)		24
Turrini *et al.* (2009)	34	Multicentre	No	CRT: radiation therapy (45 Gy) with continuous infusion of 5-fluorouracil accompanied by a cisplatin bolus.	17/34 (50)	17/17 (100)	15.5
Varadhachary *et al.* (2008)	90	Single	No	CT: gemcitabine (750 mg/m^2) and cisplatin (30 mg/m^2) every 2 weeks for 4 doses. CRT: 4 weekly infusions of gemcitabine (400 mg/m^2) combined with radiation therapy (30 Gy in 10 fractions administered over 2 weeks) delivered 5 days per week.	52/90 (69)	50/52 (96)	17.4
Vento *et al.* (2007)	22	Single	No	CRT: gemcitabine intravenous infusion twice weekly before irradiation at three dose levels, which were 20, 50 and 100 mg/m^2 for an average of 10 cycles. Tumour radiation dose 50.4 Gy given in 28 fractions of 1.8 Gy per day, five days per week.	15/22 (68)	8/15 (53)	27

Neoadjuvant therapy for borderline resectable pancreatic cancer

The role of neoadjuvant therapy in the treatment of borderline and locally advanced pancreatic cancer has been much more widely accepted and has been explored in several phase II trials (table Eii).

The majority of these trials are based on gemcitabine based chemotherapy regimes, with or without radiotherapy. In 2018 the role of gemcitabine based chemoradiotherapy was supported by the results of the phase III randomised PREOPANC trial presented at the 2018 ASCO meeting (Van Tienhoven *et al*., 2018). This trial compared combination gemcitabine and radiotherapy with upfront surgery for cases of resectable and borderline resectable disease and reported superior survival outcomes with neoadjuvant therapy (17.1 months *versus* 13.5 months; *p value* = 0.074) despite the resection rate being higher in the upfront surgery arm (60% *versus* 72%), although the R0 resection rate was higher in the neoadjuvant arm (63% *versus* 31%). This is corroborated by interim analysis published in abstract form of the NCT01458717 phase II/III RCT comparing gemcitabine based neoadjuvant chemoradiotherapy with adjuvant gemcitabine based chemoradiotherapy in borderline resectable cases (n=55) that reported a median overall survival of 23months *versus* 11months respectively (*p value* = 0.011) (Kwon *et al*., 2017). However, it is worth noting that the reported survival time in the upfront surgery group is much lower than expected and even lower than that reported for locally advanced and metastatic disease that received neoadjuvant therapy in other trials (Hackert *et al*., 2016).

Neoadjuvant FOLFIRINOX has also received much recent attention following its success in the palliative context of treating metastatic disease. The landmark trial by Conroy *et al*. (2011) found that in the setting of metastatic disease FOLFIRINOX had a median overall survival of 11.1 months compared to 6.8 months with gemcitabine but at the cost of higher rates of toxicity. This resulted in the

preferential use of mFOLFIRINOX regimes that contained altered dosing of components to reduce toxicity. However to date mFOLFIRINOX has not been directly compared to the original FOLFIRINOX in the setting of metastatic disease (Raufi *et al*., 2019; Stein *et al*., 2016; Hosein *et al*., 2012; Blazer *et al*., 2015). This is important as the aims of treatment in the metastatic setting, to slow and control disease, are different from those in the neoadjuvant setting, to shrink disease and eradicate micrometastases (Raufi *et al*., 2019). Therefore it cannot be assumed that mFOLFIRINOX in the neoadjuvant setting will have comparable success.

A single arm multicenter phase II trial by Katz *et al*. (2016) found that mFOLFIRINOX with radiotherapy for the treatment of borderline resectable disease, as defined by the NCCN guidelines, gave a median overall survival time of 21.7months with a resection rate of 68% and an R0 resection rate of 93%. However this was a small trial (n=22) and although the definition of borderline resectable was clear the study population contained some cases of resectable disease. Furthermore the rate of grade 3 or above toxicities was significant at 64%. A larger single arm phase II trial (n=48) showed an overall survival time of 37.7months with FOLFIRINOX and radiotherapy with a comparable resection rate (66%) and R0 resection rate (97%).

Whilst both these studies corroborate a growing optimism surrounding the use of FOLFIRINOX based regimes in the neoadjuvant setting questions still remain as to which specific regime combination provides optimal outcomes and whether mFOLFIRINOX sufficiently addressed concerns about toxicity levels and at what cost

to survival outcomes. Also the specific benefit of the addition of radiotherapy to neoadjuvant regimes, or whether radiotherapy should be reserved for specific cases, has been difficult to discern (Raufi *et al*., 2019). A follow-up Alliance A021501 study randomising 134 patients with borderline resectable disease to receive FOLFIRINOX with and without stereotactic body frame radiotherapy has been recently suspended following interim analysis (Raufi *et al*., 2019). The ESPAC-5FU trial is a four arm trial comparing upfront surgery and neoadjuvant chemoradiotherapy, gemcitabine-capecitabine and FOLFIRINOX for borderline resectable pancreatic cancer (ISCRTN registry, 2014) and may shed some light on some of these issues.

Table Eii: Summary of Prospective Phase II/III trials of Neoadjuvant Therapy for Borderline Resectable Pancreatic Cancer

Trial	n=	Single or Multicentre trial	Randomisation	Neoadjuvant Treatment Regime: CRT= chemoradiotherapy; CT= chemotherapy	Resection Rate (%)	R0 Resection Rate (%)	Median Overall Survival in months
Borderline Resectable Cases In Neoadjuvant Arm							
Chakraborty *et al.* (2014)	13	Single	No	CRT: 50Gy in 2.5 Gy fractions plus capecitabine 825mg/m^2 twice on radiation days	5/13 (38)	4/5 (80)	9.1
Murphy *et al.* (2018)	48	Single	No	FOLFIRINOX 8 cycles, re-staging and if vascular resolution radiotherapy 5Gy x5 with capecitabine. If persistent vascular involvement radiotherapy with 5-FU or capecitabine	32/48 (66)	31/32 (97)	37.7
Jang *et al.* (2018)	50	Multicentre	Yes	CRT: 45 Gy in 2.5 fractions and 9 Gy in 5 fractions 5 times a week for 6 weeks plus gemcitabine at 400mg/m^2. Adjuvant chemotherapy gemcitabine 100mg/m^2 every 4 weeks for 4 cycles (n=27)	24/27 (89)	14/24 (58)	21
				Upfront surgery plus adjuvant chemotherapy gemcitabine 100mg/m^2 every 4 weeks for 4 cycles (n=23)	23/23 (100)	6/23 (26)	12
Kim *et al.* (2013)	68	Multicentre	No	CRT: Gemcitabine 1g/m^2 on days 1,8, 15 + oxaliplatin 85mg/m^2 on days 1, 15 every 21 days for 2 cycles with 30 Gy radiation	43/68 (63)	36/43 (84)	18.2
Mixed Resectable and Borderline Resectable in the neoadjuvant arm							
PREOPANC (Van Tienhoven *et al.*, 2018)	246	Multicentre	Yes	Upfront surgery and adjuvant gemcitabine (n=127)	91/127	31% (only per cent given; unknown of calculated from study population or number resected)	13.5
				CRT: 15 x 2.4 Gy with gemcitabine 100mg/m^2 on days 1, 8, ad 15. Adjuvant gemcitabine (n=119)	74/119	65% (only per cent given; unknown of calculated from study population or number resected)	17.1

Motoi *et al.*(2013)	35	Multicentre	No	CT: Gemcitabine 1000 mg/m^2 days 1 and 8 with S-1 40mg/m^2 BID for 14 days every 21 days x 2 cycles	30/35 (86)	26/30 (87)	19.7
Katz *et al.*(2016)	22	Multicentre	No	CRT: Modified FOLFIRINOX plus 5.5 weeks of 50.4Gy radiotherapy with concurrent 825 mg/m^2 BID	15/22 (68)	14/15 (93)	21.7
Magnin *et al.*(2003)	32	Single	No	CRT: either split-course therapy (two courses of 15 Gy) or standard-fractionation therapy (45 Gy during 5 weeks) + concurrent 5-fluorouracil and a cisplatin bolus.	19/32 (59)	Not stated	16
Moutardier *et al.*(2002)	19	Single	No	CRT: 5-FU 650mg/m days 1-5 and days 21-25 + cisplatin 80mg/m bolus day 2 and day 22 with radiotherapy 30Gy split course or standard 45Gy	15/19 (79)	Not reported	20
Satio *et al.*(2009)	68	Single	No	CRT: radiotherapy plus 5-FU or gemcitabine (n=27)	18/27 (67)	9/18 (50)	24.5
				Upfront surgery plus adjuvant therapy (n=41)	30/41 (73)	7/30 (23)	18.5
Van Buren *et al.*(2013)	59	Single	No	CRT: fixed dose rate gemcitabine 1500 mg/m^2 + bevacizumab 10mg/kg for 3 cycles plus radiotherapy 30 gy in 10 fractions plus bevacizumab	43/59 (73)	38/41(93)	16.8
Mornex *et al.* (2006)	41	Multiple	No	CRT: radiotherapy 50Gy plus 5-FU (300 mg/m^2/day, 5 days/week, 5 consecutive weeks) and cisplatin (20 mg/m^2/day, Days 1-5 and 29-33)	26/41 (63)	Not stated	9.4
Casadei *et al.* (2015)	38	Single	Yes	Upfront surgery (n=20)	15/20 (75)	5/15 (33)	19.5
				CRT: gemcitabine 1g/m^2 days 1, 8, every 21 days for 2cucles followed by gemcitabine 50mg/m^2 twice weekly for 6 weeks (n=18)	11/18 (61)	7/11 (64)	27.5
Borderline and Locally Advanced Pancreatic Cancer in Neoadjuvant Arm							
Cardenes *et al.*(2011)	28	Single	No	CRT: gemcitabine 600mg/m^2 and radiotherapy 50.4Gy in 28 fractions 1.8Gy/day 5 days per week. If no progression weekly gemcitabine 100mg/m^2 on days 1, 8, 16 for 6 cycles	4/28 (14)	2/4 (50)	10.3
Esnaola *et al.*(2014)	37	Single	No	CRT: gemcitabine (1000 mg/m^2) + oxaliplatin (100 mg/m^2) repeated every 14 days for 6 cycles	11/37 (30)	11/11 (100)	11.8

				combined with weekly cetuximab (400 mg/m^2) on day 1 of week 1, followed by 11 weekly infusions of 250 mg/m^2 on day 1 of each subsequent week. Re-staged at 2-4 weeks. Patients with resectable disease went on to surgery. Patients with stable disease went on to chemoradiation with concurrent weekly capecitabine (800 mg/m^2 orally twice daily).			
Fiore *et al.* (2017)	34	Single	No	CRT: gemcitabine 1000 mg/m^2 and oxaliplatin 100 mg/m^2 every 14 days for four doses. For patients without disease progression radiation therapy and concurrent gemcitabine 600 mg/m^2 weekly.	19/34 (56)	15/19 (79)	19.2
Lee *et al.*(2012)	43	Single	No	CT: gemcitabine 1,250 mg/m^2 days 1 and 8, and capecitabine at 950 mg/m^2 b.i.d. days 1–14 every 3 weeks. Adjuvant CRT for R1 resection and unresectable disease post neoadjuvant therapy.	17/43 (40)	14/17 (82)	16.6
Leone *et al.* (2013)	39	Single	No	CRT: GEMOX (gemcitabine 1000mg/m^2 + oxaliplatin 100mg/m^2). If no disease progression gemcitabine 50mg/m^2 plus radiotherapy 50.4 Gy.	11/39 (28)	9/11 (82)	16.7
Marti *et al.*(2008)	26	Multiple	No	Gemcitabine (GEM) (1000 mg/m(2)) or CDDP (30 g/m(2)). Patients without progression of disease then underwent surgery or escalating GEM/CDDP doses combined with full-dose radiotherapy. If unresectable disease at restaging patients had further GEM/CDDP.	4/26 (15)	3/4(75)	13
Massucco *et al.* (2006)	72	Single	No	CRT (n=28): Radiotherapy 45 Gy. Chemotherapy: gemcitabine twice weekly for 5 weeks at dose 100 mg/m^2 twice weekly (n=15) or Dose reduced to 50 mg/m^2 twice weekly (n=8) or two courses of induction gemcitabine and oxaliplatin-based	8/28 (29)	7/8 (88)	15.4

				chemotherapy before chemoradiation (n=5).			
				Upfront surgery with or without gemcitabine based adjuvant chemotherapy or chemoradiotherapy (n=44)	44/44 (100)	30/44 (68)	14
Small *et al.* (2011)	32	Single	No	CRT: gemcitabine, 1,000 mg/m2, every 1 to 2 weeks + bevacizumab, 10 mg/kg every 2 weeks, and 36 Gy of radiotherapy (2.4-Gy fractions during cycle two).	10/29 (34)	Not stated	11.07

Neoadjuvant therapy for locally advanced pancreatic cancer

Cases of locally advanced pancreatic cancer have, by definition, no evidence of metastatic spread but their conversion to resectability has traditionally been a rare occurrence. However, a growing number of prospective phase II trials have reported some promising results with neoadjuvant chemotherapy and chemoradiotherapy (table Eiii). The majority of these neoadjuvant regimes are gemcitabine based but it is also worth noting that there is a growing body of observational studies that are reporting similarly positive findings with mFOLFIRINOX. In 2012 Hosein *et al*. published their retrospective series evaluating neoadjuvant FOLFIRINOX for locally advanced PDAC and reported that 7 out of 18 patients (39%) were converted to resectable disease. However, this was a small retrospective study that actually also included some borderline cases of disease by NCCN definition. A larger retrospective study of 43 patients that included 25 cases of locally advanced disease, as defined by NCCN definition, reported that mFOLFIRINOX resulted in 11/25 (44%) undergoing resection with 10/11 (91%) achieving R0 resection with median overall survival not reached at time of reporting (Blazer *et al.*, 2015). These findings were corroborated by a

further retrospective study of 29 patients with local advanced PDAC treated with neoadjuvant FOLFIRINOX which resulted in 41.3% undergoing resection and of those resected 83% had R0 resection with an overall cohort median survival time of 18.6months (Nanda *et al*., 2015). Interim analysis of the randomised phase III trial CONKO-007 that explores induction chemotherapy of FOLFIRINOX or gemcitabine with and without the addition of radiotherapy of locally advanced pancreatic cancer has also reported positive preliminary results (Lee *et al.,* 2016, Brunner *et al*., 2019). Out of 126 patients 36 underwent surgery. Of the 25 who had R0 resection overall survival was 26.5months with non-R0 surgery giving an overall survival of 16.9months, which was comparable to the overall survival in the non-operative group of 16.5months (Brunner *et al*., 2019). This trial is expected to conclude in 2020 and will hopefully help to clarify the role of the addition of radiotherapy to neoadjuvant regimes.

In summary, whilst the current body of evidence surrounding the treatment of locally advanced PDAC is plagued by the same issues of mainly small, underpowered studies with a high degree of heterogeneity due to variations in definitions of staging and treatment regimes, it does appear to show that whilst not all cases will be converted to resectability, when conversion is achieved such a response to neoadjuvant therapy can have a profound impact on survival time as evidenced by the R0 resection rates amongst resected cases (Raufi *et al*., 2019). However, as table Eiv shows the optimal combination and dosing regimes of mFOLFIRINOX still has to be established with variation between studies making comparison difficult.

Table Eiii: Summary of Trials of Neoadjuvant Therapy for Locally Advanced Pancreatic Cancer

Trial	n=	Single or Multicentre trial	Randomisation	Neoadjuvant Treatment Regime: CRT= chemoradiotherapy; CT= chemotherapy	Resection Rate (%)	R0 Resection Rate (%)	Medina Overall Survival in months
Jensen *et al.* (2014)	23	Single	No	CRT: Virginia Mason Protocol: 5-FU, cisplatin, interferon-alpha and radiotherapy	7/23 (30)	6/7 (86)	11.5
Landry *et al.* (2010)	21	Multiple	Yes	CRT: gemcitabine 500mg/m^2 for 6 weeks with radiotherapy (n=10)	3/10 (30)	Not reported	19.4
				CRT: gemcitabine 175 mg/m^2 days 1, 5, 29 and 33 plus cisplatin 20mg/m^2 on days 1-5 and 29-32, plus 5-FU 600 mg/m^2 on days 1-5 and 29-32 plus radiation with 5-FU 225 mg/m^2 for 6 weeks (n=	2/11 (18)	Not reported	13.4
Herman *et al.* (2015)	49	Multiple	No	CRT: gemcitabine 1000mg/m^2 x 3 doses plus radiotherapy 33.0 Gy	4/49 (8)	4/4 (100)	13.9
Ikeda *et al.*(2013)	60	Multiple	No	CRT: S-1 80mg/m^2 BID plus radiotherapy 50.4 Gy	2/60 (3)	Not reported	16.2
Sherman *et al.*(2015)	45	Single	No	CT (CRT in arterial involvement): GTX capecitabine (1500 mg/m^2 days 1-14, gemcitabine 750mg/m^2 days 4 and 11, docetaxel 30 mg/m^2 days 4 and11.	40/45 (89)	28/40 (70)	29
Crane *et al.*(2011)	69	Single	No	CRT: gemcitabine (1,000 mg/m^2) and oxaliplatin (100 mg/m^2) every 2 weeks for four doses + radiation (50.4 Gy to the gross tumor only) with concurrent capecitabine (825 mg/m^2 twice daily on radiation treatment days). Cetuximab (500 mg/m^2) was on day 1 of chemotherapy and continued every 2 weeks during chemotherapy and chemoradiotherapy.	9/69 (13)	9/9 (100)	19.2
Laurent *et*	22	Multiple	No	CRT: Gemcitabine/oxalipla	4/22 (18)	3/4(75)	17

al. (2009)				tin 2 cycles followed by 5 weeks of radiotherapy plus a weekly fixed dose gemcitabine and an escalating dose of oxaliplatin from 40 up to 70 mg/m(2).			
Lin *et al.*(2005)	42	Single	No	CRT: 3x 6 week courses gemcitabine 1000 mg/m^2 gemcitabine once weekly x 2 weeks; 1 week break; and radiotherapy total dose Gy.	6/42 (14)	3/6 (50)	10.3
Lind *et al.* (2008)	17	Single	No	CRT: 2 x courses Xelox (oxaliplatin 130 mg/m^2 day 1; capecitabine 2000 mg/m^2 day 1-14) plus 3-D conformal radiotherapy (50.4 Gy; 1.8 Gy fractions) with reduced Xelox (days 1-5).	8/17 (47)	8/8 (100)	19
Magnino *et al.* (2005)	23	Single	No	CRT: gemcitabine 100mg/m^2 (n=15) or 50mg/m^2 (n=8) plus radiotherapy 45 Gy in 1.8 Gy fractions.	6/23 (26)	5/6 (83)	14
Mattiucci *et al.* (2010)	40	Single	No	CRT: Weekly gemcitabine (100 mg/m^2) and radiotherapy 50.4 Gy. Then 5 cycles gemcitabine (1000 mg/m^2.	4/40 (10)	4/4 (100)	15.5
Sahora *et al.* (2011)	25	Single	No	CT: gemcitabine (900 mg/m^2) and docetaxel (35 mg/m^2) on days 1, 8, and 15 of a 28-day cycle.	12/33 (36)	9/12 (75)	16
Wilkowski *et al.*(2009)	95	Multiple	Yes	CRT: 5-fluorouracil 350 mg m(-2) per day + radiotherapy 50.4 Gy (n=31)	4/31 (13)	8/18 (44)	9.6
				CRT: gemcitabine 300 mg m(-2), and cisplatin (30 mg m(-2) + radiotherapy 50.4 Gy (n=32)	8/32 (25)		9.3
				CRT: as above then followed by gemcitabine 1000 mg m(-2) and cisplatin 50 mg m(-2) every 2 weeks (n=31)	6/31 (19)		7.3
Maximous *et al.* (2009)	25	Single	No	CRT: Gemcitabine (300 mg/m2) plus radiotherapy 50.4 Gy	6/25 (24)	Not stated	12
Al-Sakhun *et al.* (2003)	20	Single	No	CRT: PACE (cisplatin 100 mg/m^2 day 1, cytarabine 2 g/m^2 12 hours x 2 doses, and caffeine 400 mg/m^2 each after cytarabine dose; and days 3 to 21, 5-FU 250 mg/m^2. Followed by radiotherapy 50.4 Gy.	3/20	Not stated	13.4

Table Eiv: Summary of mFOLFIRINOX studies

Study	Type of Study	Study Population Size	Stage of Pancreatic Cancer	Modified FOLFIRINOX dosing regime	Resection Rate (%)	R0 resection Rate (%)	Median Overall Survival in months
Mahaseth *et al.* (2013)	Retrospective	24	Borderline resectable/ Locally Advanced and Metastatic	Oxaliplatin (85mg/m^2) Irinotecan (180mg/m^2) 5-FU infusion (2,400 mg/m^2)	42	83	17.8
Marthey *et al.* (2015)	Retrospective	77	Locally advanced	Oxaliplatin (85mg/m^2) Irinotecan (180mg/m^2) 5-FU bolus (400 mg/m^2) 5-FU infusion (2,400 mg/m^2)	36	89	22
Nada *et al.* (2015)	Retrospective	29	Locally advanced	Oxaliplatin (85mg/m^2) Irinotecan (180mg/m^2) 5-FU infusion (2,400 mg/m^2)	41.3	83	18.6
Sadot *et al.* (2015)	Retrospective	101	Locally advanced	Oxaliplatin (68mg/m^2) Irinotecan (144mg/m^2) 5-FU bolus (320 mg/m^2) 5-FU infusion (1,920 mg/m^2)	31	55	25
Blazer *et al.* (2015)	Retrospective	39	Locally advanced	Oxaliplatin (85mg/m^2) Irinotecan (165mg/m^2) 5-FU infusion (2,400 mg/m^2)	51.1	86	18
Stein *et al.* (2016)	Retrospective	31	Borderline resectable / Locally advanced	Oxaliplatin (85mg/m^2) Irinotecan (135mg/m^2) 5-FU bolus (300 mg/m^2) 5-FU infusion (2,400 mg/m^2)	41.9	100	26.6

Appendix F

Critical Analysis of Diagnostic Laparoscopy Cost Effectiveness Studies

This appendix supports section 2.2.1. The reason for focusing on this aspect of the pancreatic cancer management pathway is because the evidence base surrounding diagnostic laparoscopy is debated and clinical practice varies yet several cost effectiveness and cost analysis studies exist. The purpose of this appendix is to explore the quality of methods used to assess the cost effectiveness of diagnostic laparoscopy for staging pancreatic cancer to ascertain whether lessons can be learned regarding modeling under uncertainty.

The role of diagnostic laparoscopy is more ambiguous and controversial. As the NICE guidelines (2018) state, laparoscopy with laparoscopic ultrasound should be performed where resectional surgery is considered to be a possibility but small volume peritoneal and/or liver metastases are suspected. Given the current understanding of pancreatic cancer as a systemic disease, some would argue that micrometastatic disease should be suspected in all cases of potentially resectable PDAC. Diagnostic laparoscopy has a mortality rate less than 0.1% and has been championed as a safe means of detecting small-volume metastatic disease, not visible on pre-operative scanning, prior to proceeding with exploratory laparotomy (Boyd & Nord *et al*., 2000; Garcea *et al*., 2012). However, with advances in imaging techniques, the role of diagnostic laparoscopy has become controversial with some arguing that this

should no longer have a place in routine practice (Evans, 2018). Single-center studies have reported a reduced diagnostic yield in avoiding unnecessary exploratory laparotomy from 19% to 9.5% (Tilleman *et al*., 2004) with others reporting the numbers benefiting from diagnostic laparoscopy over the past decade falling to between 6% and16% (Garcea *et al*., 2012). Conversely estimates of patients who undergo exploratory laparotomy but are found to have non-resectable disease have been reported at 25% to 40% (Mayo *et al*., 2009), despite advances in preoperative imaging techniques (Morris et al., 2015). Detection of metastatic disease is essential in preventing patients from undergoing unnecessary exploratory laparotomy, with associated morbidity, longer hospital stay and delay in commencing palliative chemotherapy (Jayakrisshnan *et al*., 2015). Furthermore, costs associated with unnecessary surgery are significant. Resection with and without complication costs £12006 and £7083 respectively (Morris *et al*., 2015; NHS, 2012). Exploratory laparotomy without resection costs £5378 and £4487 with and without complication respectively (Morris *et al*., 2015; NHS, 2012). Diagnostic laparoscopy, including histological testing of tissue, costs £955 (Morris *et al*., 2015; NHS, 2012). In a setting of increasingly limited resources, optimisation of treatment must be considered in both terms of patient benefit and cost-effectiveness (Jayakrisshnan *et al*., 2015).

A recent Cochrane review, whilst concluding that diagnostic laparoscopy could reduce unnecessary exploratory laparotomy from 40% to 18%, called for further research into cost-effectiveness of staging diagnostic laparoscopy (Allen *et al*., 2013). Only one of the studies in this review had a low risk of bias and overall the quality of

the studies was low, covering a time frame when advances were being made in the quality of CT imaging (Allen *et al*., 2013). None of the studies included patients who had received neoadjuvant therapy. One of the postulated benefits of neoadjuvant therapy is that it eliminates micrometastesis (Asare *et al*., 2016; Lee *et al*., 2016). This raises the question as to whether diagnostic laparoscopy still has a role in the neoadjuvant context. Recently there has been an interest in identifying patients whose radiology or serology places them at higher risk of developing occult metastatic disease (Garcea *et al*., 2012). Increased diagnostic yield from diagnostic laparoscopy has been associated with: tumor size and location (Callery *et al*., 2009; Stefandis *et al*., 2006), high Ca 19-9 levels (Alexakis *et al*., 2015) and pro-inflammatory markers (C-reactive protein, neutrophil-lymphocyte ratio, platelet-lymphocyte ratio, hypoalbuminaemia) (Garcea *et al*., 2012; Smith *et al*., 2008).

A total of four papers exploring cost-effectives of diagnostic laparoscopy for staging pancreatic cancer prior to exploratory laparotomy were identified, one of which explored the subject matter in relation to patients receiving neoadjuvant therapy (table Fi).

Table Fi: Summary of Cost-Effectiveness Studies for Staging Diagnostic Laparoscopy

Author/ Country	Study Population	Costing	Utility and time horizon for outcomes	Payments Excluded	Methodology
Jayakrisshnan *et al.* (2015) USA	Included patients treated in both upfront surgery and neoadjuvant pathways Single Centre	Third party payer's perspective based on Medicare payments using ICD-9 codes	Cost per QALMs. 6 months	Readmission post surgery, complications chemo and radiotherapy, end-of-life care	Decision tree. ICER calculated per QALM. WtP set at US$4166 per QALM. Sensitivity analysis
Morris *et al.* (2015) UK	Upfront surgery pathway only: resectable PDAC or peri ampullary cancer on CT scan Single Centre	From National Schedule of Reference Costs 2011-12- NHS trusts and NHS foundation trusts: NHS own costs.	Cost per QALYs. 2 weeks, 3 months and 6 months	End-of-life-care, re-admission	Decision tree. MNBs: mean QALY per patent treated x decision maker's maximum WtP – net cost per patient of treatment. WtP set at £20 000 – £30 000 per QALYs. Sensitivity analysis: PSA.
Enestvedt *et al.* (2008) USA	Upfront surgery pathway: included only resectable PDAC defined by CT scan. Data drawn from OSCaR.	OHSU billing data, fiscal year 2005-6 using CPT codes. Physician charges from Medicare and Medicaid Services fees using CPT code.	Mean charge per patient per ultization strategy for DL. Retrospective: 1996-2003	Subsequent PC-related procedures: re-operation, endoscopic interventions. End-of-life-care, complications.	Analytic charge measure per patient per diagnostic laparoscopy utilization strategy of 1) routine 2) case-specific 3) no utilization. No sensitivity analysis performed.
Tapper *et al.* (2011) USA	Resectable or borderline resectable disease on MRI scan. Single Centre	Cost incurred by hospital based on ICD-9 codes and Emroy University Hospital financial department data.	Cost of procedure / total cost of hospitalization US$ Retrospective: 2004-2008	Assumed no false negatives or complications from DL. Out patient end-of-life care, readmission/ complications.	Formulae to determine number of metastases required for diagnostic laparoscopy to be cost-effective. No sensitivity analysis performed.

QALMs = quality adjusted life months, QALYs = quality adjusted life years, ICER = incremental cost-effectiveness ratio, WtP= Willingness to pay, MNBs = monetary net benefits, PSA= Probabilistic analysis, OSCaR= Oregon State Cancer Registry, OHSU = Oregon Health and Science University

Three of the four studies recommended routine diagnostic laparoscopy prior to exploratory laparoscopy (Morris *et al*., 2015; Jayakrisshnan *et al*., 2015; Enestvedt *et al.* (2008). One study concluded that diagnostic yield from diagnostic laparoscopy was marginal and cost-effectiveness was poor (Tapper *et al*., 2011). However these conclusions must be interpreted with some caution.

Whist Jayakrisshann *et al.* (2015) reported savings of $10695 and $4158 per quality adjusted life months (QALM) in upfront surgery and neoadjuvant groups respectively, this was based on the probability that 30-60% of patients were found to have un-resectable disease at diagnostic laparoscopy (Jayakrisshnan *et al*., 2015). However, the payer perspective used in this study fails to capture other costs and factors such as operating room time and equipment with added human resources and disruption to efficacy of process in the operating room with routine setting up of laparoscopic equipment prior to exploratory laparotomy (Jayakrisshnan *et al*., 2015). Furthermore, when sensitivity analysis was performed diagnostic laparoscopy did not prove cost effective at a diagnostic yield of less than16% (Jayakrisshnan *et al*., 2015). Juxtapose this with advances in pre-operative imaging and neoadjuvant therapy and the likelihood of advanced disease being detected at diagnostic laparoscopy seems more likely to continue to fall.

This point is also valid when interpreting the results of Tapper *et al.* (2011) who, having identified that palliative procedures were less costly (open palliative bypass: $21957.18; endoscopic palliative procedure: $11304.00) than pancreaticoduodenectomy ($26122.43), found that routine diagnostic laparoscopy increased costs by 3.6% hence cost-effectiveness of this approach depended on the number of patients converting to endoscopic palliation (Tapper *et al*., 2011). However, this study did not perform any sensitivity analysis and failed to set a figure at which diagnostic laparoscopy would become cost-effective. Instead this study calculated the number of metastases that would need to be identified at diagnostic laparoscopy to justify

costs. The benefit of this approach in relation to cost-effectiveness is questionable given the role of vascular invasion and metastatic spread, as opposed to a set number of metastases, determining resectability. Although this study did conclude that diagnostic laparoscopy held minimal diagnostic yield and marginal cost-effectives (Tapper *et al*., 2011), any presumed cost-effectives of diagnostic laparoscopy drawn from these results would be questioned as analysis was biased towards over estimating cost-effectiveness of diagnostic laparoscopy. Several arguably incorrect assumptions were made. Firstly that diagnostic laparoscopy detected all metastatic disease when 7%-35% of negative diagnostic laparoscopies have been reported to have metastatic disease at exploratory laparotomy (Tapper *et al*., 2011; Vargas *et al*., 1995). Secondly, diagnostic laparoscopy was assumed to have a complication rate of 0% and did not prolong hospital stay when minor and major complication rates have been reported at 1.7%-5.1% and 6.7% and 2.3% respectively (Tapper *et al*., 2011; Vargas *et al*., 1995). Further potential bias lies in the assumption that all non-diagnostic laparoscopy patients who received palliative endoscopic intervention did so as inpatient whilst all patients who underwent diagnostic laparoscopy received palliative interventions as outpatients (Tapper *et al*., 2011). Finally the generalisability of findings from this retrospective study are questionable as, although this was population based, costs were based on single center data (Tapper *et al*., 2011).

Morris *et al.* (2015) reported that although diagnostic laparoscopy and direct exploratory laparotomy had similar costs (£7470 *versus*

£7480), the cost of diagnostic laparoscopy (£995) was offset by avoiding negative exploratory laparotomy. Diagnostic laparoscopy also had improved quality-adjusted-life-years (QALYs) (0.346 *versus* 0.337) with 63-66% cost-effectiveness with willingness-to-pay set at between £20000 and £30000 per QALY (Morris *et al*., 2015). However, this could be misinterpreted as the mean QALY gain (0.346-0.337=0.0009), although statistically significant, is small and less than the minimal clinically important difference which, in health state utility values, ranges from 0.010 to 0.048 (Morris *et al*., 2015). Furthermore, when diagnostic laparoscopy took place on the same admission as planned exploratory laparotomy, the cost increased to £8224 as, although cost of hospital stay was avoided if un-resectable disease was found, the cost associated with booked theatre time was not (Morris *et al*., 2015).

Enestvedt *et al.* (2008) supported either case-specific or routine diagnostic laparoscopy as they found it did not add significant expense based on resource utilisation. However, the generalisability of these findings are limited by costs drawn from a single center that do not represent either incurred or reimbursed costs (Enestvedt *et al.*, 2008). Furthermore no sensitivity analysis was performed to prove their cost-effectiveness argument considering that whilst diagnostic laparoscopy avoided exploratory laparotomy in 28% of patients, only 8% of those who had exploratory laparotomy, and none who had diagnostic laparoscopy, had un-resectable disease, whilst 26% of cases deemed resectable at diagnostic laparoscopy were found to have metastatic disease at exploratory laparotomy (Enestvedt *et al.*, 2008). This questions both diagnostic yield and

cost-effectiveness of diagnostic laparoscopy in this study population. Furthermore additional costs such as human resources, additional theatre time etcetera were not factored into their analysis.

Appendix G

Critical Analysis of Cost-Effectiveness Studies of Neoadjuvant Therapy for Solid Organ Malignancies.

This appendix supports section 2.2.2 and provides a detailed critical analysis of each cost-effectiveness analysis study for neoadjuvant therapy for solid organ malignancies. As the existing body of evidence for the management of pancreatic cancer is so limited this broader critical analysis was undertaken to assess whether the quality of modeling improved where the evidence base surrounding neoadjuvant therapy for other malignancies was better established.

Gordon et al. (2012). Modeling the Cost-Effectiveness of Strategies for treating Oesophageal Adenocarcinoma and High-grade Dysplasia.

Gordon *et al.* (2012) aimed to establish cost-effectiveness of current strategies for treating oesophageal adenocarcinoma and high-grade dysplasia using a decision-analytic model with results reported in QALYs. For T2-T4 cancers decision arms included no surgery, surgery without neoadjuvant therapy and surgery after neoadjuvant therapy. The numbers in the non-surgery group who received chemoradiotherpy or no treatment were not made clear. The neoadjuvant therapy and chemoradiotherapy regimes used were also not clearly defined. Adjuvant therapy post surgery was also not included as a separate treatment arm.

Data used to populate the model was based on Australian Cancer Study Clinical Follow-up Study (ACS) and prospectively collected databases from major oesophageal cancer units on Adelaide and Brisbane. Overall 2000 patients were included with 1000 having undergone oesophagectomy. However ACS only included patients with invasive tumors, potentially biasing data (Gordon *et al.,* 2012). Treatment probabilities were calculated from literature and Australian all cause mortality data. However, literature search criteria and inclusion criteria for studies on which these were based were not stated. Remaining gaps in evidence were populated by expert opinion from five surgeons therefore introducing potential bias.

Costs were calculated from patient-level resource data and priced using national price schedules and public hospital clinical costings. Costs included all follow-up costs and mean cost of oesophagectomy included intensive care unit admission and in hospital adverse events. Costs excluded were not stated but readmission and adverse events associated with chemo/radiotherapy and palliative care and indirect costs were not presented in the results. Utility scores were obtained from literature review and used to adjust estimates. Sensitivity analysis was then undertaken to address potential uncertainty in these estimates. Incremental net benefit was calculated over a 5year period for each alternative management scenario, one of which related to neoadjuvant therapy with FDG-PET/CT to assess response to neoadjuvant therapy.

Results regarding the role of neoadjuvant therapy compared to surgery first were ambiguous. Net benefit of adding FDG-PET/CT is highlighted at $805; 95% UI $59-$1,596. However, the focus was on benefit of earlier detection with no assessment of costs associated with this strategy. There was no assessment of the possibility of neoadjuvant therapy avoiding unnecessary surgery or impacting on adverse events post operatively et cetera. No conclusion can therefore be drawn from this paper on cost-effectiveness analysis of neoadjuvant therapy.

Hultman et al. (2012). Costs and clinical outcome of neoadjuvant systemic chemotherapy followed by cytoreductive surgery and hyperthermic intraperitoneal chemotherapy in peritoneal carcinomatosis from gastric cancer.

This study set out to evaluate cost and clinical-effectiveness of neoadjuvant therapy prior to cytoreductive surgery followed by hyperthermic intraperitoneal chemotherapy and/or early postoperative intraperitoneal chemotherapy compared to palliative systemic chemotherapy alone. Survival was calculated by Kaplan-Meir method with costs estimated with Bootstrap resampling method and presented in US $ for QALYs gained. All interventions were comprehensively described.

Neoadjuvant therapy study population was small (n=10) and drawn from a single center without randomisation or blinding. Systemic chemotherapy group (n=10) were selected as matched control patients from an ongoing randomised control trial (GATAC trial).

Selection was blinded to response and survival time. This meant that within this group patients were randomized to one of two systemic chemotherapy regimes under GATAC trial protocol, which could potentially skew data of this study. Both groups were matched for age, gender, performance status, tumor extent and ASA grade.

Costs were based on Uppsala University Hospital data and Swedish National Pharmacy 2008 pricelist and a detailed list of all pre-treatment, treatment, and post-treatment costs were taken from retrospective review of medical records for each patient. No details were given about costs excluded but costs of palliative/ hospice care and indirect costs are not detailed. No discounting of costs was undertaken due to short survival times of these patients.

No data was available on quality-of-life or health utility, therefore Health Utility Weights (HUW) were estimated on World Health Organisation (WHO) performance status or Karnofsky Performance Score (KPS). No quality of life data was collected during this study.

Results from this study reported costs per life-year-gained $166,716 and QALY gained $ 175,164. However, sensitivity analysis showed that if complication rates from surgery fell by 50% then cost of treatment in neoadjuvant therapy group would fall to $12400 (Hultman *et al.,* 2012). This highlights a fundamental limitation of this study in its small sample size as one patient alone in the neoadjuvant therapy group had very high treatment costs ($487, 756) due to complications hence skewing results (Hultman *et al.,* 2012).

Vitale et al. (2010). Use of Sorafenib in Patients with Hepatocellular Carcinoma Before Liver Transplantation: a cost-benefit analysis while awaiting data on Sorafenib safety.

This study used Markov decision model to assess cost-effectiveness of Sorafenib prior to liver transplant for patients with Hepatocellular Carcinoma (HCC) compared to no bridging therapy. Endpoints were: delay in HCC progression as hazard ration (HR), survival in quality-adjusted-life-days (QALDs), transplant probability, cost utility ratio, willingness-to-pay and incremental net health benefit (INHB).

The model was populated with data from major studies and randomised controlled trials although search and inclusion criteria were not given. Costs, in euros, were taken from current payments within Italian public healthcare system. Indirect costs were reported as being excluded. Given that this study was undertaken whilst awaiting data on Sorafenib safety potential costs of adverse events associated with this drug were not considered. Equally the impact of recurrence rates post liver transplant was not considered. Where assumptions were made in the model they were all justified on best available evidence in the literature. No robust data was available on tumor stage of waiting list patients at time of drop out. Assumptions were therefore made based on current guidelines. This, and all other assumptions made, were included in the Monte Carlo probabilistic sensitivity analysis.

Results reported cost-utility ratio €197 per QALD with INHB 37 QALDs, assuming willingness-to-pay (WtP) was €346 per QALD (Vitale *et al.,* 2010). INHB continued to rise most significantly over the first 6 months on waiting list. The report therefore concluded that neoadjuvant therapy is cost-effective compared to no treatment for T2-HCC patients awaiting LT for up to 6 months (Vitale *et al.,* 2010). However, in some ways this study is premature in its conclusions. Benefits of neoadjuvant therapy with Sorafenib could be underestimated as more contemporary studies have shown declining HR on time to progression therefore the HR in this study may be lower than assumed (Vitale *et al.,* 2010; Gubanski *et al.,* 2010). This model did not consider potentially beneficial effects of neoadjuvant therapy on biological tumor aggressiveness thus post LT reoccurrence rate. Conversely as little was know about the negative impact of Sorafenib on postoperative complications or its toxic effects, particularly pertinent in transplant patients who will be immunocompromised, costs could have been underestimated and benefits over estimated (Vitale *et al.,* 2010).

Ercolani et al. (2011). Effectiveness and cost-effectiveness of peri-operative versus post-operative chemotherapy for resectable colorectal liver metastases.

A Markov decision model, populated with data from a detailed literature search with specified inclusion criteria spanning 10years, was used to determine cost-effectiveness of neoadjuvant therapy compared to upfront surgery followed by adjuvant chemotherapy for resectable colorectal liver metastases. Interventions were clearly

detailed and outcomes reported in quality-adjusted-life-months (QALMs), incremental cost-effectiveness ratio (ICER), HR of reoccurrence, willingness-to-pay, and recurrence-free-survival (RFS).

Costs were in euros and taken from the Italian public healthcare system 2010 and drug costs were based on wholesale costs as recommended by pharmoco-economic analysis. Costs excluded were not stated but indirect costs were not included in analyses. However, potential complications of both treatment approaches and palliative treatments were included. In particular literature has highlighted increased post-operative complications with neoadjuvant therapy and this was factored into the model in the form of morbidity occurrence and post-operative hospital stay for the neoadjuvant therapy group (Ercolani *et al.,* 2011).

Univariant sensitivity analysis showed that 3-year RFS and cost of hepatectomy were main determinants of cost-effectiveness of neoadjuvant therapy. 3-year RFS of 36.8% = ICER €448.1/QALM and 3-year RFS of 31.5% = ICER €8075/QALM (Ercolani *et al.,* 2011). Two-way sensitivity analysis showed that neoadjuvant therapy was cost-effective if RFS was equal to or less than 25% (Ercolani *et al.,* 2011). Overall neoadjuvant therapy was deemed cost-effective as, although increased life expectancy was small, costs of neoadjuvant therapy were also small with ICER below WtP. Furthermore the costs of neoadjuvant therapy were offset by savings from patients becoming unresectable hence avoiding costs of hepatectomy (Ercolani *et al.,* 2011). However in subgroups with favorable tumor behaviour savings were minimal which highlights an area were

further research is needed (Ercolani *et al.,* 2011). As with all modeling studies results are limited by the quality of literature available.

Poston et al. (2001). Costs of neoadjuvant chemotherapy and surgery in patients with liver metastases from advanced colorectal cancer.

This study used a simple decision model to compare a theoretical cohort of 2000 patients with advanced colorectal cancer and unresectable liver metastases being treated with oxaliplatin + 5-FU/FA versus 5-FU/FA alone. Outcomes measure included: mean overall survival estimates, drug acquisition costs, costs associated with surgery and health-benefits measured as LYG.

Resectability rates were taken from the de Garmont RCT and set at 11.4% and 4.1% for the oxaliplatin + 5-FU/FA and 5-FU/FA arms respectively. Mean post-surgical survival was estimated from Kaplan-Meier survival curves taken from a single retrospective study on reduction of tumour size in patients treated with oxaliplatin + 5-FU/FA. Patients who were still alive at the end of the follow-up period were given an over all survival assumed to be the same as that of age matched normal population. Whilst this may represent an overestimate in survival, given recurrence rates, this assumption was tested in sensitivity analysis by reducing length of survival by 25% (Poston *et al.,* 2001). Again however, this was not calculated based on best available evidence on reoccurrence rates.

Costs were gathered from a variety of sources. 6-month chemotherapy costs were taken from unpublished economic evaluation of the de Garmont data. Costs of surgery came from The Royal Liverpool University Hospital. Costs of surgery included: pre-operative assessment, surgery, post-operative intensive care, inpatient hospital stay, and out-patient clinical appointments. Costs excluded were: drug administration costs, all other post-operative costs including further chemotherapy/surgery, palliative care and indirect costs. Costs were not discounted in cost-effectiveness analysis. However, given that mean survival was estimated at 9.0 and 1.7 years for each arm, yet the study used a 6-month time horizon, overall cost for each arm could be over estimated (Poston *et al.,* 2001).

Results from this study showed that ICER of oxaliplatin + 5-FU/FA compared to 5-FU/FA alone of £11,985 with sensitivity analysis giving an ICER of £5489-£15,624 per LYG with combination treatment increasing resectability to 7.3% to 17.5% (Poston *et al.,* 2001).

Van der Brink et al. (2004). Cost-Utility Analysis of Preoperative Radiotherapy in Patients with Rectal Cancer Undergoing Total Mesorectal Excision (TME): a study of the Dutch Colorectal Cancer Group.

Markov model populated with data obtained from multicentre RCT (Kapiteljn *et al.,* 1999) was used to compare cost and QALY of patients undergoing TME with and without neoadjuvant therapy.

Outcomes were: life expectancy, lifetime costs per patient, QALY and ICER.

Probabilities and transition rates for the model were established using Gompertz distribution to allow increases and decreases with time. For mortality rates proportional hazards for neoadjuvant therapy and age were considered. Non-significant proportional hazards (PH) were only excluded if *P value* > 0.05 and sensitivity analysis was performed on all included and excluded estimated PH and on PH of neoadjuvant therapy for local reoccurrence rate. Sensitivity analysis, covering a wide range of possible outcomes, was also performed by R-status to investigate benefit of improved diagnostics.

One potential source of bias was the fact that patients with co-morbidities were excluded from the RCT (van der Brink *et al.,* 2004; Kapiteljn *et al.,* 1999). This was reflected in mortality in the study population being lower than average Dutch mortality (van der Brink *et al.,* 2004). However, to counter this, long-term mortality rates were taken from Dutch-life tables. Reoccurrence rates were also varied over both groups to assess all potential outcomes in QALY and cost-effectiveness.

This was the only study that used quality of life data collected from the study population at 3,6,9, 12, 18 and 24 months post surgery by questionnaire as well as additional qualitative interviews with 112 patients in the cost-utility analysis (CUA) group, although selection to this group was not detailed.

Costs were taken from price index rate for Dutch healthcare sector and included, from health care perspective: primary treatment, continued care and recurrence treatment. For societal perspective they included: informal care costs, travel costs, time costs and out-of-pocket costs calculated from diaries from the CUA population. Time and travel costs were based on average duration and distances for different types of healthcare in Netherlands. Hospital time costs were based on 8hours/day and valuation of time was assessed in CUA sample using WtP model.

Results showed that loss of quality-of-life due to NAT was outweighed by increased life expectancy (0.67 years, 0.39 QALYs costing $9800 per patient) (van der Brink *et al.,* 2004).

The strengths of this study are that it collects prospective data alongside a RCT with a large sample population, and sensitivity analysis covers a wide range of probabilities and possible outcomes. However, generalisation, as with all studies is a potential issue. Also cost-effectiveness was measured within clinical trial and so may be different if applied routinely, considering co-morbidities were excluded from the RCT (van der Brink *et al.,* 2004). Whilst exclusion of non-significant PH avoided a type 1 error, this could have potentiated a type 2 error (van der Brink *et al.,* 2004). However, as all PH were included in sensitivity analysis ICER remained within acceptable limits.

Rocconi et al. (2005). Management strategies for stage IB2 cervical cancer: a cost-effectiveness analysis.

Decision analysis model, populated by data from literature review, was created to compare cost-effectiveness of three management strategies of stage 1B2 cervical cancer: 1) radical hysterectomy with pelvic and para-aortic lymphadenectomy plus tailored adjuvant chemoradiotherapy (RHYST) 2) primary chemoradiotherapy for all patients (CTRT) and 3) neoadjuvant therapy followed by radical hysterectomy and tailored adjuvant chemoradiation (NAC). Outcomes were measured in 5year disease free survival and cost per cure/ survivor. Chemotherapy regimes were not clearly specified but presumed to be the same in this theoretical model.

Literature review search terms and inclusion criteria were not stated but where possible phase III and II trials were used. However, estimates from literature varied widely and it was not made clear how this study decided upon estimates used given such variations. For example, it was estimated that 40% needed adjuvant therapy post surgery when literature reports this percentage to range from 34 to 84% (Rocconi *et al.,* 2005; Namkoong *et al.,* 1995; Landoni *et al.,* 1997; Delgado *et al.,* 1989; Rettenmaier *et al.,* 1989). In particular 5year DFS was used, yet in the literature for NAC, few studies have follow-up beyond 2 years (Rocconi *et al.,* 2005; Landoni *et al.,* 1997; Serur *et al.,*1997; Eddy *et al.,* 1995; Kim *et al.,* 1988). 5year disease free survival was estimated at 70% on the basis that existing literature reported 5year disease free survival to as good as RHYST and CTRT which was therefore also estimated at 70% to avoid bias

(Rocconi *et al.,* 2005). Similarly it was estimated that 25% of NAC received adjuvant therapy, based on the assumption that NAC downstages tumors, and that 65% in NAC would need no further treatment. However no reference is given to justify these estimates. Also to avoid bias, probability of aborted surgery and positive para-aortic nodes were considered equal to RHYST strategy but these estimates were not references to best available data. These estimates were however included in one-way sensitivity analysis. Only grade 3 and 4 complications were included in this model and costs related to treatment-associated complications were excluded from analysis.

Laboratory and procedure costs were derived from consultation with University of Alabama at Birmingham Pharmacy department. All other charges were estimates from adjusting local charges using cost-to-charge ration of 60%. Excluded costs were indirect costs and reimbursements. Analysis was from a third party payer perspective.

This study concluded that cost per cure was reasonable but policy makers would need accept WtP $500,000 per survivor for NAC or $2.2million per survivor in CTRT. Calculations of WtP were not given in this study but these figures are higher than widely reported limits of WtP.

This model was narrow in its scope and cost to diagnose and treat complications were not included (Rocconi *et al.,* 2005). Each cohort was assumed to have node-negative disease as all patients had pre-operative CT scans which neglects estimates on accuracy of CT scanning. Also patients with intermediate risk factors were not

included as the role of chemoradiation in these patients has been debated (Rocconi *et al.,* 2005). Quality of life was also not considered in this analysis.

Attard et al. (2015). Cost-effectiveness analysis of neoadjuvant pertuzumab and trastuzumab therapy for locally advanced, inflammatory, or early HER2-positive breast cancer in Canada.

A three health state Markov model (event-free, relapse free and dead) was used to conduct a cost-utility analysis of data separately from the Neosphere trail (Glannl *et al.,* 2012) and TRYPHAENA trial (Schneeweiss *et al.,* 2013) to assess cost-effectiveness of neoadjuvant pertuzumab and trastuzumab for locally advanced, inflammatory or early HER2-positive breast cancer. All interventions were described in full.

The model was populated with data from NeoSphere and TRYPHAENA (Schneeweiss *et al.,* 2013) trials. However, the Neosphere trial was conducted to isolate the treatment effect of pertuzumab as pre-operative agent. Yet in the FEC arm of this trail patients were meant to receive this agent post-operatively but in clinical practice received the agent pre-operatively. This questions applicability of findings to clinical practice. Also the TRYPHAENA data is limited by the lack of a comparator arm without pertuzumab (Attard *et al.,* 2015). Data was not available on event-free-survival and overall for those who did not achieve complete pathological response with neoadjuvant pertuzumab, for the entirety of the 28-year time horizon (Attard *et al.,* 2015). This data was taken from a

single retrospective study analysing HER-2 patients treated with neoadjuvant trastuzumab and chemotherapy. Hence the model assumed that for this group, events free and overall survival were dependent on achieving complete pathological response or not (Attard *et al.,* 2015). Survival beyond a 10year period was extrapolated from survival study published by Kim *et al.* (1988) Once again reliance on a single study on which to base model estimates introduced potential bias. Quality of life data used adjusted health utilities from breast cancer specific systematic review but utility decrements for adverse events were not included (Attard *et al.,* 2015).

Costs included direct medical costs: all drugs, administration, adverse event management, supportive care and subsequent therapy. Costs were taken from published literature and Ontario cost database and were inflated to 2014 Canadian dollars. Costs were discounted at a rate of 5% per anum. Analysis was from the payer's perspective.

Results from the NeoSphere trial showed the addition of neoadjuvant pertuzumab had an incremental cost $23,658 per LYG and $25,388 per QALY while TRYPHAENA analysis predicted incremental cost of $43,047 per LYG and $46,196 per QALY (Attard *et al.,* 2015).

Poonawalla et al. (2015). Cost-effectiveness of neoadjuvant chemotherapy versus primary surgery in elderly patients with advanced ovarian cancer.

This retrospective cohort study analysed data from the Surveillance, Epidemiology and End Results (SEER)-Medicare linked database to establish cost-effectiveness of NAT versus primary debulking surgery for patients aged over 65 years with advanced stage III/IV epithelial ovarian cancer. Analysis was from payer perspective using Medicare costs. Total healthcare costs were included. Excluded costs were: indirect costs and out-of-pocket costs (deductibles or co-payments) hence this study lacked an overall societal evaluation. A phase-of-care approach was used to estimate cumulative treatment costs, and effectiveness was measured as years of survival, ICER and propensity-score-adjusted net monetary benefit regression to estimate cost-effectiveness per life year gained.

Within the study population 12% (n=591) received neoadjuvant therapy and 88% (n=4252) received primary debulking surgery with or without adjuvant therapy. There was no randomisation between treatment groups. Although propensity score was used to adjust for baseline characteristics differences between the two groups (Poonawalla *et al.,* 2015), there was clearly a significant variation in the population size. Neoadjuvant therapy regimes were not stated and the study's definition of primary debulking surgery was not made clear.

Analysis did not further divide the upfront surgery group into those receiving adjuvant therapy. This introduces bias considering years of survival was a key outcome and the potential impact complications of adjuvant treatment might have had on cumulative costs. Survival analysis was used to reweight estimates for mean healthcare costs.

This could have introduced bias in the presence of censored data (Poonawalla *et al.,* 2015). Interestingly mean costs in both groups increased when weighting with survival, indicating observed costs were underestimated (Poonawalla *et al.,* 2015). Specifically there was an increase in continuing-phase costs, which could be explained by high recurrence rates (Poonawalla *et al.,* 2015).

Results showed that neoadjuvant therapy had an ICER of $174,173 for 0.1 LYG (Poonawalla *et al.,* 2015). In high-risk patients ICER for neoadjuvant therapy was $42,987 and 0.8 LYG. No adjustment for quality-of-life was made. Results are limited in their generalisability based on the study inclusion criteria (Poonawalla *et al.,* 2015). The study also carries limitation from the database such as: missing claims and incomplete data reporting (Poonawalla *et al.,* 2015). Key information missing from the database included tumor distribution and extent of resection, which has a significant impact on prognosis.

Rowland et al. (2015). Cost-utility comparison of neoadjuvant chemotherapy versus primary debulking surgery for treatment of advanced-stage ovarian cancer in patients 65 years old or older.

A 5-year Markov Model was populated with data from randomised controlled trials, EORTC 55971(Vergote *et al.,* 2010) and GOG 152 (Ozols *et al.,* 2003) trials, to evaluate cost-effectiveness of NAT relative to SF for stage III/IV ovarian cancer in patients aged 65 years and over. Outcome measures included: overall survival, surgical complications, probability of treatment cost and quality-of-life. Treatment costs were taken from Medicare with hospital costs

estimated from Agency for Healthcare Quality and Research's Healthcare Cost and Utilisation Project. CEA was from healthcare system perspective. Excluded costs included surveillance costs, chronic complications and indirect costs. All interventions were clearly detailed.

Complication estimates rates were based on Medicare data and not on RCTs. This introduces bias as complication rates of surgery could be underestimated and patients with more co-morbidities or disease burden may have preferentially received neoadjuvant therapy. Chronic complications were assumed to be equal across both groups therefore were excluded. No evidence was referenced to support this assumption. This introduces further bias as cost associated with increased chronic complications being more prevalent could affect overall results. Therefore overall cost assumptions are likely to underestimate cost difference and underestimate neoadjuvant therapy savings (Rowland *et al.,* 2015). However, 1 and 2-way sensitivity analysis in addition to probabilistic sensitivity analysis using Monte Carlo simulation was performed to address uncertainty associated with all model parameters.

Results reported a cost saving of $5616 per patient treated with neoadjuvant therapy, assuming equal survival (Rowland *et al.,* 2015). However, upfront surgery improved overall survival by 1.5 months or 3.2 months or longer, this would be cost-effective at $100,000/QALY and $50,000/QALY threshold respectively (Rowland *et al.,* 2015). This highlights some key limitations of this study. The model was based on one RCT. Overall survival has the strongest

influence on variability of model outcomes yet overall survival was the aspect of the RCT that was most heavily criticized with a subsequent study questioning the overall survival reported in the EORTC trial (Dewdney *et al.,* 2010). Reliance on a single trial also questions the generalisability of findings to clinical setting as alternative treatment options, such as intraperitoneal chemotherapy, was not included in the model (Rowland *et al.,* 2015).

Stevenson et al. (2014). Cost-effectiveness of neoadjuvant chemotherapy before radical cystectomy for muscle-invasive bladder cancer.

This study compared cost of treatment, duration of survival and adjusted quality-of-life survival for patients with muscle-invasive bladder cancer treated with radical cystectomy (RC) or with neoadjuvant therapy. Outcomes were measured cost per QALY, and all interventions were clearly defined in this retrospective review.

Costs were obtained from nationwide 2009 Healthcare Costs and Utilisation Project. Chemotherapy costs were established from international sources. Indirect costs were not included. Quality-of-life literature was lacking for bladder cancer so adjustment for QALYs was based on studies involving similar conditions.

Differences in total mean costs were analysed using in dependent variable *t* tests. QALYs were assessed using Kaplan-Meier analyses and compared using stratified log-rank test with two-tailed *P* values also used. Costs per QALY were calculated using the formulae: (mean

cost of treatment for NAT cohort – mean cost of treatment for upfront surgery cohort)/ (median survival for neoadjuvant therapy cohort – media survival for upfront surgery cohort). Ratios derived were compared using *t* test.

Assessment in this way did not allow sensitivity analysis given that estimates of quality-of-life and complications were made from limited available research (Stevenson *et al.,* 2014). It also did not allow for assessment of all possible scenarios in each cohort with corresponding sensitivity analysis. No multivariate analysis was performed to validate differences in survival between the two cohorts. Also patients who received neoadjuvant therapy but not surgery were included in an intention to treat analysis as part of the neoadjuvant therapy cohort. This could have reduced mean total cost for neoadjuvant therapy as they did not incur cost of radical cystectomy and also may have increased QALY survival as quality of life would not have been affected by surgery (Stevenson *et al.,* 2014). Data was also not available on randomisation of patients to upfront surgery or neoadjuvant therapy group and there were significant differences between the two cohorts (age and race), which were not considered in the analysis (Stevenson *et al.,* 2014). This study was also limited by being retrospective and cohorts were drawn from a single center. Also some costs were estimated from published sources, others were based in billing, and chemotherapy prices came from international literature, which could limit accuracy of total costs (Stevenson *et al.,* 2014).

Appendix H

A Critical Analysis of Existing Risk Stratification and Predictive Models for Pancreatic Cancer Surgery

This appendix supports section 2.3.1 by providing a historical background to predictive modeling and risk stratification in medicine. This helps to establish good modeling practice, which is then used as a basis to critically evaluate existing predictive models applied to pancreatic cancer surgical patients.

One of the earliest surgical risk stratification tools was the American Society of Anesthesiologists (ASA) grading system (Saklad, 1941) which offered a subjective pre-operative assessment of physical status without any statistically rigorous underpinnings. During the 1960s and 70s however advancements in fields of computing and statistics led to multivariate log regression analysis taking a leading role in predictive medicine (Lewis & Volmer, 2012). Its application to the Framington Heart Study data (Truett *et al.,* 1967), although not the first of its kind, proved to be a seminal paper resulting the increased application of log regression analysis to medical problems (Hosmer &Lemeshow, 1989). This, juxtaposed with advances in cardiac surgery (Lewis & Volmer, 2012), led Goldmann *et al.* (1977) to produce the cardiac risk index (CRI) to assess the risk of cardiac events from non-cardiac surgery. This marked an increased focus on preoperative risk assessment (Lewis & Volmer, 2012). In the same decade Cox developed his Cox Regression model enabling

proportional hazard modeling to analyze censored data and allow multivariate modeling of survival data (Cox, 1972).

Throughout the 1970s and 80s there was a focus on using pre-operative nutritional indexes to assess operative risk to general surgical patients (Buzby *et al.,* 1980; Harvery *et al.,* 1981). The development of the Fong Score (Fong *et al.,* 1999) to predict recurrence after liver resection for metastatic colorectal cancer however demonstrated that risk scores and predictive modeling could be applied to highly targeted clinical questions which marked a focus on organ/disease specific predictive modeling (Lewis & Volmer, 2012) as exemplified by the Model for End-Stage Liver Disease used to determine organ allocation in transplant surgery (Malinchoc *et al.,* 2000; Kamath *et al.,* 2007). There then followed two key developments: a growth in cancer specific prognostic models (Kattan *et al.,* 2001; Peeters *et al.,* 2005; Bochner *et al.,* 2006; Brennan *et al.,* 2004; Stephenson *et al.,* 2006; Weiser *et al.,* 2008; Prediction Tools, 2012) and standardisation of definition of post-operative complications (Lewis & Volmer, 2012). This resulted in a focus initially on predicting post-operative morbidity and mortality, which has more recently expanded to also include the readmission and resource utilisation with subsequent cost-effectiveness analysis (Lewis & Volmer, 2012). In the United States large administrative databases, such as the National Surgical Quality Improvement Program (NSQIP) (Khuri, 2005), are maturing and becoming increasingly accessible. This should enable modeling of more specific and rare disease and surgical outcomes (Lewis & Volmer, 2012). This also raises some pertinent questions about the current and future

applications of predictive modeling to pancreatic surgery that will now be explored.

In its most simplistic form regression analysis predicts an outcome (Y) from line-of-best-fit to input data. Therefore 'a' estimates outcome (Y) and 'b' estimates the alteration in Y associated with an alteration in measurement X:

$$Y=a+bX$$

Applied to multiple variables the co-efficients (β_{0-3}) estimate the increase in outcome 'Y' corresponding to per unit increase in input variables (X,W,Z). Hence larger coefficients result in greater impact on outcome:

$$Y= \beta_0+ \beta_1X+\beta_2W+\beta_3Z$$

This technique may be adequate to model continuous variables but not dichotomous outcomes, therefore are unable to provide surgeons with the type of yes/no answers they often seek: will my patient survive this operation, will they develop a post-operative complication, will they be alive 1year, 2years or 5years post operatively? (Lewis & Volmer, 2012)

In univariate cases linear modeling is simply inadequate. To illustrate, using linear regression to model CA 19-9 levels as a predictor of malignancy would provide the nonsensical probability of malignancy as greater than 1 for any CA 19-9 level above 700 (Lewis

& Volmer, 2012). One method of addressing this is to model data to a logistic curve by grouping CA 19-9 levels so that the probability of malignancy lies between 0% and 100%, providing more information from input data to show that as CA 19-9 levels increase, so does the likelihood of malignancy (Lewis & Volmer, 2012). This approach has the benefit of being adaptable to multivariate analysis and can fit data to a function:

$f(z)=e^2 / e^z+1 = 1/1+e^{-z}$
where $z= \beta_0+ \beta_1X+\beta_2W+\beta_3Z....$

An advantage of logistic regression is that it can utilize the intuitive interpretation of odds ratios (OR) as coefficients can easily be converted: $OR =e^{\beta}$ (Hosmer & Lemeshow, 1989). To illustrate, the risk of a smoker undergoing pancreaticoduodenectomy procedure developing a post operative myocardial infarction could be represented with an OR greater than 1, hence increased risk. Equally continuous variables can be modeled in a similar way so that smokers undergoing pancreaticoduodenectomy could have their increase in risk of post-operative myocardial infarction predicted per unit increase in age (Lewis & Volmer, 2012).

Cox Proportional Hazard Models are widely used to model censored survival and disease free survival in time series data (Cox *et al.,* 1984) by plotting survival as well as hazard, with hazard ratios (HR) derived from βcoefficients of regression (Lewis & Volmer 2012). Therefore this method of modeling may be able to predict, for example, that smokers are three times more likely to die within

30days pancreaticoduodenectomy (HR=3.00) (Lewis & Volmer, 2012).

Linear regression is therefore good for predicting one-time events such as death at 30 days post operatively, surgical mortality, readmission etcetera, and is used to predict near time surgical outcomes (Lewis & Volmer, 2012). Cox proportional hazard models predict survival at any desired length of time (Lewis & Volmer 2012).

Models, particularly those used in decision making must report decision analytic measures (Steyerberg *et al.,* 2010; Lewis & Volmer, 2012). This involves both discrimination, measure of sensitivity and specificity graphically represented by receiver operating characteristic (ROC) curve, and calibration, the observed-to-expected (O/E) ratio (Steyerberg *et al.,* 2010; Lewis & Volmer, 2012). Regarding the latter, concordance statistic (c-statistic) is important in indicating the level of discrimination with a value of 1 depicting perfection but greater than 0.7 deemed reasonable. The c-statistic is important where models are found to have errors in calibration, such as overestimating mortality, as a high c-statistic means that the model can still be useful in stratifying patients at higher risk of mortality (Lewis & Volmer, 2012). Such an error in calibration may be depicted with a slope of less than 1, indicating that the model is over fitted (Lewis & Volmer, 2012) hence requiring correction through shrinkage of the regression coefficient (Miller *et al.,* 1993). Alternatively the Hosiner-Lemeshow test is a more rigorous test of calibration, comparing O/E rates deciles and uses X^2 distribution to

test the null hypothesis that O and E are equivocal, hence a *P value* <0.005 would indicate poor calibration (Lewis & Volmer, 2012). Pearson's R statistic is better applied to continuous variables and, similar to the Brier Score, measures the percentage variation between observed and expected that can be explained by the model (Lewis & Volmer, 2012; Bland *et al.,* 1995). Decision-analysis techniques are pertinent in models used for decision making to assess the impact of false positives and false negatives with decision curve analysis (Bland *et al.,* 1995) used to assess the model's clinical impact (Lewis & Volmer, 2012).

Validation confirms the predictive performance of the model as well as its generalisability (Miller *et al.,* 1993). It also identifies over fitting, a risk with a high numbers of variables, as well as deficiencies due to inappropriate study techniques (Altman *et al.,* 2000). This involves testing discrimination and calibration on novel data. Such novel data may be acquired through split sample validation (Lewis & Volmer, 2012). However, this approach can destabilise the model by reducing the amount of data it is built on whilst also meaning that the model is being internally validated on a data set that is too small hence inadequate assessment. Bootstrapping, where the model is developed from all data and assessed on bootstrap samples of the original data with replacement, is instead recommended (Steyerberg *et al.,* 2009). This has the advantage of assessing discrimination and calibration as well as confidence intervals (Steyerberg *et al.,* 2009) with the element of randomness in repeated bootstrapping overcoming overoptimistic conclusions (Lewis & Volmer, 2012). Temporal validation also allows the model to be constructed from all

available data and validated against novel data prospectively collected within the same centre (Lewis & Volmer, 2012). This however fails to properly assess generalisability of the model (Altman *et al.,* 2000) therefore external validation on novel data from a different centre is the most stringent method of validation (Altman *et al.,* 2000) although issues with case-mix and timing must be considered as models developed on retrospective data may not apply to more contemporary, novel or refined treatment pathways (Lewis & Volmer, 2012).

A good model therefore will have been recently externally validated as well as internally validated and demonstrate excellent discrimination, segregating patients according to risk, as well as calibration. The model will also only focus on variables that impact on outcomes. With this framework in mind a critical analysis of existing predictive models in pancreatic surgery is presented.

Predicting Mortality and Morbidity following Pancreatic Cancer Surgery

Nebraska Nomogram

In 2009 the Nebraska Nomogram (Are *et al.,* 2009), using a conceptual framework based on the MSKCC survival nomogram (Brennan *et al.,* 2004), aimed to predict post-operative mortality based on pre-operative factors only. Data was taken from the National Inpatient Sample (NIS) (n= 4482). 15 variables were assigned a point value from coefficients of a multivariate regression

model. Whilst it could be argued that the number of variables is excessive, a strength of this model is its use of ICD-9 procedure and diagnostic codes and the Elixhauer system of comorbidity classification which, in conjunction with the use of a national database, enhances generalisability of the model. However, this model does not account for pathological or intraoperative data. Whilst this can be justified on the basis that it is a preoperative predictive tool, the influence of intra and post-operative care can influence in-hospital mortality. The true clinical value of this model is also yet to be established as the model has not yet been externally validated by other studies (Lewis & Volmer, 2012).

Surgical Outcomes Analysis and Research Pancreatic Resection Mortality Score (SOAR)

Similar to the Nebraska Nomogram, the SOAR model aims to predict post-operative inpatient mortality based on pre-operative factors only (Hill *et al.,* 2010). Like the Nebraska Nomogram its use of the NIS database enhances generalisability and its availability online facilitates wider clinical use. However, it too awaits external validation through further studies and does not account for the influence of factors such as intra-operative and post-operative care on mortality. This model uses log regression methods to create an integer risk score to identify low, moderate and high-risk patients. A strength of this model over the Nebraska model is that this risk score is then further adjusted for hospital specific risk therefore accounting for the positive influence of high-volume specialist centres on better surgical outcomes (Lewis & Volmer, 2012).

Johns Hopkins Pancreaticoduodenectomy Mortality Model

This model specifically predicted 30 and 90-day mortality post pancreaticoduodenectomy (Venkat *et al.,* 2011). The specialist institution's database underwent multivariate logistic regression to identify predictive factors that were converted to integer score for mortality. The strength of this model lies in the large, detailed database form an institution with considerable reputation in pancreatic surgery. However, this model lacks external validation and has limited generalisability as distal pancreatectomies were excluded as authors felt mortality from this type of surgery was so low in their experience that it would not allow adequate statistical power. The use of single-centre data to build the model also potentially limits the generalisability as it does not allow adjustment for institution associated risk factors such as high versus low volume settings (Venkat *et al.,* 2011).

Wisconsin NSQIP Mortality and Morbidity Calculators for Pancreaticoduodenectomy and Distal Pancreatectomy

NSQIP database underwent multivariate regression analysis to identify variables influencing outcome (Greenblatt *et al.,* 2011). The variables did not include peri or post-operative factors that might have influenced morbidity and mortality outcomes. There was also no adjustment for hospital specific risk factors although large sample size and use of national database does give good generalisability. This

model also lacks external validation. However, recognising that type of pancreaticoduodenectomy may affect outcomes, the authors used the same methods to construct a second model specifically for distal pancreatectomy using NSQIP database (n=1797), which also underwent internal validation (c-statistic=0.79) but again lacks external validation.

HPB Risk Calculator

Log regression analysis of ASC-NSQIP database of all patients undergoing pancreatectomy was used to identify variables affecting mortality, serious morbidity and overall morbidity (Parikh *et al.,* 2010). Again the use of a large national database enhances generalisability although external validation has not taken place and there was no adjustment of risk of hospital specific risk factors. Although all types of pancreatectomy were included, outcomes did not give frequency and morbidity related to operation specific complications. However, further prospective validation and work on updating this calculator is ongoing.

Physiological and Operative Severity Score for the enumeration of Mortality and Morbidity (POSSUM)

Although widely used in clinical settings the POSSUM model was based on a single institution's retrospective database with the number of pancreatic resections included being small, although the model has undergone external validation. The model also

incorporates both pre-operative patient factors and operative factors. Whilst this is clinically relevant this might also make it difficult to identify the impact of poor surgical performance (Lewis & Volmer, 2012). To illustrate, high blood loss and multiple procedures predict increased risk of morbidity and mortality but this could be a result of poor surgical technique (Sutton *et al.,* 2002). The finding that POSSUM and P-POSSUM (Goffi *et al.,* 1999; Prytherch *et al.,* 1998) has been shown to predict mortality but not morbidity accurately (Pratt *et al.,* 2008) was based on a study evaluating the model's calibration not discrimination. However subsequent similar studies have shown conflicting results with one reporting poor calibration (n=652; O/E = 0.88) (de Castro *et al.,* 2009) and another reporting good calibration (n=265; O/E = 0.90). This highlights the need for standardised reporting systems and coding of relevant factors such as post-operative complications (Strasberg *et al.,* 2009; Dindo *et al.,* 2004; Lewis & Volmer, 2012). Furthermore univariate and multivariate analysis of the variables included in the model has shown that over half are not significant predictors of morbidity which suggests that a future direction of the model would also be in its simplification for pancreatic resections (Lewis & Volmer, 2012). However, this model has the advantage of including both pre-operative patient data as well as operative factors (Goffi *et al.,* 1999;Lewis & Volmer, 2012) and has been shown to accurately predict economic outcomes with higher POSSUM scores associated with increased costs (Pratt *et al.,* 2008).

Surgical Apgar Score

The Surgical Apgar Score (Gawande *et al.,* 2007) was designed as a simple risk score for post-operative complications. Based on a national database the model had potential for good generalisability but lacked external validation. It was later adapted to predict perioperative morbidity, as defined by the Clavien-Dindo scale, for pancreaticoduodenectomy patients. It was fond that this model could predict grade 2 or above complications but not necessarily mortality (Assifi *et al.,* 2012). Using only intraoperative data the proposed use of this model was in deciding whether patients should go to intensive care unit or a step-down unit post operatively if they were deemed to be at lower risk of complications (Assifi *et al.,* 2012).

Milan Pancreatic Morbidity Score

This model aimed to predict major post-operative morbidity, as defined by Clavien-Dindo classification (Dindo *et al.,* 2004), post pancreaticoduodenectomy (Braga *et al.,* 2011). The model is based on operative factors and therefore is not intended for preoperative use (Lewis & Volmer, 2012). Generalisability of the model is limited by use of single institutional data and the lack of external validation but a strength of this model is the use of standardised coding and classification of complications although use of ASA grading potentiates a degree of subjectivity in this variable (Lewis & Volmer, 2012).

Balzano et al. (2017): Preoperative score to predict early death after pancreatic resection.

Based on a comparatively smaller single centre database, this model identified preoperative factors predicting early mortality through log regression, multi and univariate analysis. There is the potential in small databases to miss additional significant risk factors. Although this limits generalisability, the model did undergo external validation (n=182). Discrimination was quantified by area under the curve (AUC) of the receiver operating characteristic (ROC) curve of 88.7% and calibration tested with Hosmer-Lemeshow test (*P* value=0.043). Standardised score was used for variables such as Geriatric Nutritional Risk Index although ASA score introduces an element of subjectivity. There is also selection bias in the database as this only included patients deemed well enough to be admitted for surgery (Balzano *et al.,* 2017). Furthermore the definition of non-metastatic liver disease was ever/never and only included cirrhosis and chronic hepatitis when type and severity of liver disease can potentially impact mortality (Balzano *et al.,* 2017).

Velez-Serrano et al. (2017): Prediction of in-hospital mortality after pancreatic resection for pancreatic cancer

The approach in this model is unique amongst other models in employing machine learning through boosting method which the authors state allows them to broaden their framework and include hundreds of variables. Whilst this means that satisfactory AUC and Brier scores are reported even when only preoperative variables are included, inferring better classification power, research into optimal tree depth is limited (Velez-Serrano *et al*., 2017). Data mining of large

databases and inclusion of many variables also does not equate with a more accurate model. The large national database on which this model is based, although enhancing generalisability and including adjustment for hospital volume, did not include all important clinical information such as stage of metastases or treatment with chemotherapy or radiotherapy. Furthermore this model lacks external validation.

Table Hi: Summary of models predicting postoperative morbidity and mortality

Model	Database	Outcome Measures	Variables	Validation	Strengths & Weaknesses
Nebraska Nomogram	National Inpatient Sample (n=4482): all patients undergoing pancreatic resection for pancreatic malignancy	Post operative hospital mortality	Renal failure, neurological disorders, hypothyroidism, heart failure, liver disease, hypertension, cardiac arrhythmias, diabetes, chronic obstructive pulmonary disease (COPD), degree of resection, age, admission type, hospital size.	Internal, temporal (n=999) (c= 0.76)	• Generalisability • Use of standardized coding • Arguably too many variables • Negates influence of intra and post operative care mortality • Lacks external validation • No adjustment for hospital specific risk factors
Surgical Outcome Analysis and Research (SOAR) Pancreatic Resection Mortality Score 2010	National Inpatient Sample (n= 5715)	Post-operative hospital mortality	Modified Charlson Score, gender, age, degree of resection, hospital volume	Internal split-sample n=1428) (c=0.71)	• Generalisability • Risk of mortality adjusted for hospital specific risk • Negates influence of intra and post operative care mortality • Lacks external validation
John Hopkins Pancreaticoduodenectomy Mortality Model 2011	Institutional database (n=1383)	30 and 90 day mortality post pancreaticoduodenectomy for total	Age, gender, albumin, tumor size, procedure type	Internal-split sample (n=593; 30 day c=0.74; 90 day c=0.73). Hosme-	• Detailed database from high volume specialist

		pancreatectomy		Lemeshow test, P=0.36 and 0.09	centre • Integer score • Online availability • Limited generalisability • No external validation
Wisconsin NSQIP Mortality and Morbidity Calculators for Pancreaticoduodenectomy and Distal Pancreatectomy 2011	NSQIP (n= 4945)	30 day morbidity and mortality after pancreaticoduodenectomy	COPD, hypertension, neoadjuvant radio therapy, serum creatinine and hypoalbuminemia	Internal-split sample (n=1254, c=0.69)	• Generalizability • Negates influence of intra and post operative care mortality • No adjustment for hospital specific risk factors • Online availability • No external validation
HPB Risk Calculator 2010	ACS-NSQIP (n=13558)	30-day mortality from all major HPB surgical procedures	Age, BMI, ASA status, ffccardiac disease, bleeding disorder, shortness of breath, ventilator dependence, ascites, steroid use, sepsis, emergency admission, procedure type	Internal validation (n=13558, c=0.75). Hosmer-Lemeshow test P=NS	• Generalisability • Use as comparison tool between centres • No external validation • No adjustment for hospital specific risk factors
Possum 1991	Single institution database (n=1372 all procedures; pancreaticoduodenectomyies n=326)	Morbidity	Age, cardiovascular disease, vital signs, Glasgow coma score, haemoglobing, WCC, BUN, serum sodium and potassium, operative mode and severity, number of procedures, blood loss, peritoneal	External validation for post pancreatic resection morbidity (n=326; O/E = 0.96)	• External validation and calibration • Wide spread clinical use • Used in economic evaluation • Includes preoperative and operative factors

			contamination, malignancy.		• Data points not easily accrued • Single institutional data, number of pancreatic resections small.
Surgical Apgar Score 2007	ACS-NSQIP: general and vascular surgical patients (n=303)	Morbidity and mortality: major post operative complications and 30-day mortality	Blood loss, lowest mean arterial pressure, lowest pulse	External validation for Clavien grade 2+ complications post pancreaticoduodenectomy	• Generalisability due to national database • Integer score • Only based on operative factors therefore provides immediate post-operative assessment of risk
Milan Pancreatic Morbidity Risk Score	Single institution database (all pancreatic resections; n=469)	Major morbidity	Pancreatic texture, pancreatic duct diameter, blood loss, ASA score	Internal split-sample validation (n=231; c=0.743; O/E = 1.11; Hosmer-Lemeshow test, P=0.937)	• Based on operative factors therefore provides immediate post-operative assessment of risk • Uses defined method of grading complications • ASA introduces source of subjectivity • No external validation
Balzan *et al.* (2017)	Single institution: all pancreatic ductal adenocarcinoma resections (n=296)	One year mortality post resection	Nutritional status, American Society of Anaesthesiologists' score, pain at presentation, non-metastatic liver disease,	External validation (n=182; OR 7.1; p<0.0001; Hosmer-Lemeshow test, P=0.403)	• Single centre limits generalisability but externally validated • No adjustment for hospital specific risk

					• factors • ASA score introduces subjectivity
Velez-Serrano et al. (2017)	Minimum Basic Data Set of the National Surveillance System for Hospital Data in Spain (n=4,088)	In-hospital mortality	Age, location of tumour, type of resection, hospital surgical volume, comorbid conditions: hypertension, cardiac disease, vascular disease, COPD, dementia, connective tissue disease, peptic ulcer disease, liver disease, diabetes, renal disease, HIV, metastatic carcinoma	Internal validation, sensitivity analysis (AUC 0.91; Brier Score 0.09).	• Generalisability given national database • Large number of variables but potential for relevant variables to be excluded due to limitations of dtabase • Use of standardized coding systems • Adjustment for hospital specific risk factors • More research needed on optimal tree depth

Models Predicting Specific Complications from Pancreatic Surgery

The International Study Group for Pancreatic Fistula (Bassi *et al.,* 2005) standardised the definition of post-operative pancreatic fistulae (POPF), resulting in the advent of models predicting this specific complication that occurs in 15%-30% of all proximal and distal pancreatic resections (Pratt *et al.,* 2007; Callery *et al.,* 2009b).

The Freibug Fistula Model

The Freibug Fistula Model (Wellner *et al.,* 2010) was based on 62 consecutive pancreaticoduodenectomies from a single institution. The aim of the model was to predict risk of POPF in the pre-operative phase so that treatment strategies could be adjusted for high-risk patients. The small dataset and single centre reduces generalisability and does not account for hospital or surgeon associated risk factors. Furthermore the model lacks external validation. The importance however was to highlight a move towards specific predictive modeling to target therapy more effectively to individual patients which could have a cost-effectiveness impact. To illustrate, high-risk patients would undergo more prophylactic measures whereas low risk patients could forego such additional treatments.

Fistula Risk Score

This model sought to go further and incorporate impact of operative factors combined with pre-operative findings (Callery *et al.,* 2009b). In this was this model predicts POPF risk pre-operatively and also supports the surgeon's decision making for high-risk patients intra-operatively (Lewis & Volmer, 2012). However this model is based on a small database from a single institution. It so far lacks external validation and does not adjust for hospital or surgeon associated risk factors. It does however highlight that as databases mature and definitions of complications become more rigorously defined, complication specific predictive models should become more sophisticated (Lewis & Volmer, 2012).

Korean Model for Predicting Pancreatic Leakage (Kim *et al.,* 2013b)

Bivariate and univariate logistic regression analysis on a single institution database of only 100 patients defined variables for inclusion in this model which lacks external validation. Generalisability of this model is therefore limited particularly as all operations were conducted by a single surgeon and POPF rates are higher than reported in the wider literature. Furthermore the model does not predict grade A POPF therefore overall risk prediction of POPF may be underestimated.

Clinical Risk Score to Predict Pancreatic Fistula after Pancreatoduodenectomy (Callery *et al.,* 2013)

Multivariable analysis of a single institution database identified variables for inclusion of this model that was internally, prospectively validated. However the model lacks external validation and the single institution database limits generalisability considering potential impact of hospital and surgeon associate risk factors on outcome. Furthermore during validation the operating surgeons were aware of the model which cold have influenced their practice (Callery *et al.,* 2013). A key variable, pancreatic texture, was defined subjectively and not cross-referenced with histopathology reports. However this is the case with other existing models predicting POPF also. This model went further in predicting the economic impact of POPF, which is an important step in fully realising the potential of

such predictive models although a full economic analysis of the impact of implementing such a model is awaited.

Predictive Model for Pancreatic Fistula Based on Amylase value in drains (AVD-based model).

Partelli *et al.* (2014) took a different approach to predicting POPF based on amylase value in drains (AVD) on day one post surgery. This identified patients who require closer monitoring in the immediate postoperative period. This interesting model was based on single institution data and lacks external validation. Also low specificity is reported. Furthermore findings could have been influenced by the lack of protocol determining removal of postoperative drain, potentiating the occurrence of POPF due to pressure gradient of the drain across the anastomoses where drains were left for a prolonged period. Further work is also needed to determine the value of AVD in predicting POPF for specific types of resection and in the context of other pre and intra-operative findings to determine AVD level cut-off for making predictions (Hackert *et al.,* 2014).

Table Hii: Summary of Models predicting POPF

Model	Database	Outcome Measures	Variables	Validation	Strengths & Weaknesses
Freiburg Fistula Model	Single centre (n=62)	POPF	Age, preoperative diagnosis other than pancreatic carcinoma/ chronic pancreatitis, smoking history, weight loss, history acute pancreatitis, high/ medium and low risk defined by pancreatic texture	Internally validated	• Not externally validated • Based on single centre data and small sample size • Limited generalisability • No adjustment for hospital and surgeon associated risk
Fistula Risk Score	Single centre (n=233)	POPF	Texture, disease pathology, small pancreatic duct diameter, blood loss,	Internally validated (n=212)	• Combines pre-operative and intraoperative findings therefore also providing decision support during high-risk intra operative scenarios • Single institution retrospective data • Small number in dataset • Limited generalisability • No adjustment for hospital and surgeon associated risk
Korean Pancreatic Leakage Model	Single centre database (n=100)	POPF	Age, gender, operation, texture, pancreatic duct size, combined superior mesenteric and portal vein resection, pathology of origin, blood loss	Internal validation (n=29) AUC= 0.714 (95% CI 0.517 to 0.865)	• Small single centre database • Limited generalisability • Surgery performed by a single surgeon • High institutional leakage rate (41%) • Does not predict grade A POPF therefore does not accurately predict overall risk of POPF • No external validation
Clinical Risk Score for Pancreatic Fistula after	Single institution	POPF	Pancreatic duct size,	Internal Validation	• Evaluates economic

Pancreatoduodenectomy	(n=233)		pancreatic texture, high-risk pathology, excessive blood loss)	(prospectively) (n=212) (AUC=0.942)	impact of model • Single centre database and sample size limits generalisability • Pancreatic texture subjective judgment not correlated with histopathology • No external validation • Surgeons in validation study were aware of model being validated so may have influenced practice
AVD-based Model	Single Centre database (n=231)	POPF	Amylase value in drains (AVD) day 1 post-operatively, operation, pancreatic texture	Internally validated. AVD day 1 post-operation >5000: Sensitivity 71% and specificity 90%, day 5 post operation AVD >200: 93% and 83% Day 1 AVD; AUC 0.876 (p<0.00001)	• Single centre • No external validation • Low specificity reported • No protocol for standardizing drain removal potentiating some POPF resulting from drain pressure gradient across anastomosis.

Models Predicting Long-Term Survival

In pancreatic cancer long-term survival is poor but risk of operative morbidity and mortality remains relatively high. Therefore although, the majority of predictive models focus on immediate surgical outcomes, long-term survival predictions are key in patient counseling and clinical decision making (Lewis & Volmer, 2012).

Memorial Sloan-Kettering Cancer Centre (MSKCC) Nomogram for pancreatic adenocarcinoma survival

Brennan *et al.* (2004) created their model based on single centre data of resected pancreatic ductal adenocarcinoma (PDAC) (n=555) using Cox regression analysis. Although use of single centre data and relatively small patient numbers could limit generalisability, the model was externally validated (Ferrone *et al.,* 2005) and combined demographic, pathological and operative data meaning that the model can support individualised patient counseling by both surgeons and oncologists (Lewis & Volmer, 2012). Variables found not to be significant in the Cox regression analysis were however still included in the final model although this enhanced predictive ability of the model (Lewis & Volmer, 2012). The database pre dated neoadjuvant therapy and adjuvant treatment was not included as a variable.

Pancreatoduodenectomy Prognostic Index (PPI)

Dasari *et al.* (2016) performed univariate and multivariate Cox regression analysis on a single institution database of patients who had undergone pancreaticoduodenectomy for malignancy between 2004 and 2014. This analysis revealed tumour site, stage and lymph node ratio as being significant for predicting 1 year and 3 year survival. The generalisability of this model is limited by the small, retrospective, single institution database and lack of external validation. In particular the number of duodenal carcinomas was sub-optimal and further validation of the model for this particular sub-group is required. Also adjuvant therapy was not included in the final model as it was found to be significant at 1 year but not 3year survival. This however could be due to inconsistencies in the

indications and type of adjuvant therapy offered over the study time period and between subtypes of tumour. Timing and number of cycles of adjuvant therapy received also varied introducing further inconsistencies. To assume that adjuvant therapy therefore does not have a bearing on prognosis may therefore not be an accurate assumption.

Interactive Bayesian SEER model for predicting lymph node ratio and survival

Smith *et al.* (2014) used a novel interactive Bayesian approach to model true but unobservable lymph node ratio (LNR) and overall survival for patients with resected PDAC. The use of a large national database and rigorous coding enhances generalisability. Although the model underwent extensive satisfactory split internal evaluation and sensitivity analysis it has not yet undergone external validation. The clinical impact of this model is primarily in guiding pathological examination of resected specimens to determine how many lymph nodes should be examined before determining lymph node involvement. Although the model may predict survival in light of LNR, it does not consider impact of adjuvant treatment on survival therefore has limited clinical application to patient counseling. However, this model is important in its novel method of predictive model particularly when considering its future application to modeling individual patient genomic data.

Table Hiii: Summary of models predicting long-term survival

Model	Database	Outcome Measures	Variables	Validation	Strengths & Weaknesses
MSKCC Nomogram	Single centre (n=555)	1,2,3 year mortality for resected PDAC	Tumour size, tumour location, posterior margin, T-stage, positive lymph node count, splenectomy, differentiation, portal vein resection, gender, margin resection, age, weight loss, back pain	Externally validated	• Combines preoperative, tumour factors and operative factors • Can be used by surgeons and oncologists • Externally validated • Based on single centre data and relatively small sample size
PPI	Single centre (n=567)	1 and 3 year survival	Tumour site, stage and lymph node ratio	Validated on institutional prospective database (n=194; AUROC score: 0.74)	• Single institution retrospective data • Small number in dataset • Limited generalisability particularly in institutions with differing adjuvant therapy protocols
SEER Model	NCI SEER cancer registry (n=6400)	Lymph node ratio and overall survival for PDAC following resection	Age, gender, marital status, grade, histology, T and M stages, tumour size, radiation therapy	Internal split-sample validation (n=2133; concordance index=0.65 (95% CI 0.63-0.66); posterior p values (lymph node ration: p=0.3300; survival: p=0.4847)	• Large, national database enhances generalisability • Limited patient specific information available on SEER database • No external validation • Focus on guiding pathological assessment of resected specimen • Use in post operative prediction of survival but as does not include adjuvant treatment has limited use in patient counseling

Appendix I

A Review and Critical Analysis of Methodological Quality of Prognostic Development Studies for Resectable Pancreatic Cancer

The aim of this appendix is to support the discussion in section 2.3.2 regarding the methodological quality of prognostic development studies for resectable pancreatic cancer. The body of this appendix and the discussion in section 2.3.2 has been published in the review article written by the autor of this thesis:

Bradley, A., Van Der Meer, R. and McKay, C.J. (2019) 'A systematic review of methodological quality of model development studies predicting prognostic outcome for resectable pancreatic cancer'. *BMJ Open,* 9**:**e027192. doi: 10.1136/bmjopen-2018-027192

The purpose is to describe and assess the methodological quality of prediction research pertaining to model development studies that predict post resection prognosis of PDAC published since 2000. As no prognostic model has yet been established for use in routine clinical practice, this date restriction was intended to capture the possibility of newer models incorporating latest developments in the management of potentially resectable pancreatic cancer including the use of neoadjuvant therapy. Prognostic modelling studies in this context included prognostic model development studies with and without external validation and external validation studies with model updating. This included only prognostic multivariable

prediction studies where the aim was to identify a relationship between two or more independent variables and the outcome of interest to predict prognosis. Predictor finding studies and studies that investigated a single predictor, test, or marker were not included. Studies investigating only causality between variables and an outcome were excluded. Model impact studies and external validation studies without model updating were excluded as the focus was on assessing the methodological quality of prognostic model development.

All methodological issues that are considered to be important in prediction research were critically analysed according to the Checklist for critical Appraisal and data extraction for systematic Reviews of prediction Modeling Studies (CHARMS) checklist (Moons *et al*., 2014). This checklist is designed for appraisal of all types of primary prediction modeling studies including emerging methods of neural network and vector machine learning (Moons *et al*., 2014). Data pertaining to the domains outlined in the CHARMS checklist are analysed and presented. These domains include: data sources, sample size, missing data, candidate predictors, and model development, performance and evaluation (Table Ii). Risk of bias assessment of included studies was performed according to the Prediction model Risk of Bias Assessment Tool (PROBAST)(Wolff et al., 2014).

Table Ii: Summary of Classification of Domains from CHARMS Checklist (Moons *et al*., 2014).

Domain	Key Information
Data Source	Registry data, randomized-controlled-trial data, case-control data, cohort data
Participant Selection	Participant eligibility (inclusion/ exclusion criteria, description of participants, treatment received)
	Recruitment methods (setting, location, number of centres, consecutive participants, study dates)
Model Outcomes	Definition of outcomes: type (single or combined endpoints), was the same definition used in all participants?
	Definition of methods for measuring outcomes: same in all participants, blinding, were candidate predictors part of the outcome?
	Duration of follow-up or time of outcome occurrence reported
Candidate Predictors	Number, type, definition, method and timing of measurement, was assessment blinded, how were candidate predictors handled within the model?
Sample Size	Number of participants and number of outcomes or events. Event per variable (number of outcomes / number of candidate predictors)
Missing Data	Number of participants with any missing data, number of participants with missing data for each predictor variable, methods for handling missing data
Model Development	Modelling methods, methods for selecting predictors to include in multivariable analysis, methods and criteria for selection of predictors during multivariable

	analysis, shrinking of predictors or regression co-efficients
Performance and Evaluation	Calibration and discrimination with confidence intervals, classification measures (sensitivity, specificity, predictive value etc.), methods for testing performance, comparison of data distribution of predictors for development and validation datasets, in poor validation was model updating performed, alternative presentations of the model (nomogram, calculator, score etc.)
Presentation of Results and Discussion	Comparison with other studies, generalizability, strengths and limitations

This review included a total of 15 model development studies, based on a total of 20,510 patients, published between 2004 and 2018. A full summary of included studies is provided in Table Iii with risk of bias assessment provided in Table Iiii.

Table Iii: Summary of included model development studies

Study	Data Source and Study Population Size (n=20510)	Model Outcome	Candidate Variables	Method of Predictor Selection	Included Predictor Variables	Missing Data	Modeling Method	Model Performance and Validation	Presentation
Brennan *et al.* (2004)	Single institution database (n=555)	1,2,3 year survival	n=14	Cox multivaraite analysis; p-value but non significant variables included	age, sex, portal vein inclusion, splenectomy, margin, location, differentiation, posterior margin, nodes positive, nodes negative, back pain, T stage, weight loss, maximum pathological axis (n=14)	Predicted using regression models	Multivaraite Cox proportional hazards regression	Calibration plot: good fit Discrimination: *c*-statistic 0.64 External Evaluation: absent; internal validation by bootstrap method	Nomogram
Kanda *et al.* (2014b)	Single institution database (n=324)	Death within 12 months of surgery	n=19	Pre-selection by Univaraiate analysis then Cox multivaraite analysis. P-value < 0.1	CEA, CA19-9 (n=2)	Complete case analysis	Multivaraite Cox proportional hazards regression	Calibration: absent Discrimination: AUC 0.702 External Evaluation: absent	Index
Miura *et al.* (2014)	Single institution database (n=50)	1,3,5 year survival	n=50	Pre-selection by Univaraiate analysis then Cox multivaraite analysis. P-value <0.05	Platelet count, CRP, CA19-9 (n=3)	Complete case analysis	Multivaraite Cox proportional hazards regression	Calibration: none Discrimination: none External Evaluation: absent	Score
Shen *et al.* (2018)	Multi-centre databases (n=239)	6,12,18 month survival	n=17	Pre-selection by Univaraiate analysis then Cox multivaraite analysis. P-value <0.05	age, length of tumour contact, peripancreatic venous abnormality, lymph node staging (n=4)		Multivaraite Cox proportional hazards regression	Calibration plot: good fit Discrimination: *c*-statistic 0.824 External Evaluation: performed	Nomogram
Xu *et al.* (2017)	Single institution database (n=265)	1,3,5 year survival	n=14	Pre-selection by Univaraiate analysis then Cox	Tumour grade, pathological stage, neural invasion, vascular invasion,		Multivaraite Cox proportional hazards regression	Calibration: calibration curve Discrimination: 1yr: c-index 0.86,	Nomogram

				multivariate analysis. P-value <0.05	Neutrophil Lymphocyte Ratio, Platelet to Lymphocyte Ratio, Albumin Globulin Ratio (n=7)			AUC: 0.938. 3yr: c-index 0.837, AUC 0.844, 5yr: c-index: 0.809, AUC 0.884 External Evaluation: absent	
Walczak & Velanovich (2017)	Single institution database (n=219)	7 month survival post surgery	n=7	Single hidden layer back propogation trained ANN.	age, sex, stage, survival time, quality of life, adjuvant therapy, resection details (n=7)	Complete case analysis	Artificial Neural Network	Calibration: absent Discrimination: AUC: 0.6576, sensitivity 91.30%, specificity 38.27% External Validation: absent; internal validation by random split method	Calculator
Hsu *et al.* (2012)	Single institution database (n=740)	Death at 9 and 12 months	n=15	Pre-selection by Univariate analysis then Cox multivariate analysis. P-value <0.05	age, tumour size, comorbidities, tumour grade (n=4)	Complete case analysis	Multivariate Cox proportional hazards regression	Calibration: none Discrimination: none External Validation: absent; internal validation with p-value <0.05	Score
Botsis *et al.* (2009)	Single institution database (n=218)	Survival time	n=26	Pre-selection by Univariate analysis then Cox multivariate analysis. P-value <0.05	Age, differentiation, tumour size, Alk Phos, Albumin, Ca19-9 (n=6)	MICE presuming data missing at random	Multivariate Cox proportional hazards regression	Calibration: absent Discrimination: *c*-statistic 0.73 External Validation: absent; internal validation by bootstrap method	Score
Liu *et al.* (2018)	Multi-centre databases (n=1223)	Survival time	n=10	Pre-selection by Univariate analysis then Cox multivariate analysis. P-value <0.05	American Joint Commission on Cancer stage, tumour grade, post-operative Ca19-9 (n=3)		Multivariate Cox proportional hazards regression	Calibration: absent Discrimination: *c*-statistic 0.70, AIC: 2406.37 External Validation: absent; internal validation by random split method	Score
Balzano *et al.* (2017)	Single institution database	1 year mortality	n=56	Pre-selection by Univariate	American Society of Anaesthesiologists' score, Geriatric		Multivariate Cox proportional hazards	Calibration: Hosmer-Lemeshow 0.403. Discrimina	Score

	(n=296)			analysis then Cox multivariate analysis. P-value <0.2 for univariate analysis and <0.1 for multivariate analysis	Nutritional Risk Index, abdominal/back pain, non metastatic liver disease or insulin resistance (n=4)		regression	tion: R^2 53.5%, AUC: 88.7%. External Validation performed	
Smith & Mezhir (2014)	National Registry Database (n=6400)	6 months, 1,3,5 year survival	n=12	Backward stepdown selection process	Survival: age, gender, marital status, race, grade, histology, T&M, size, radiation, Lymph Node Ratio (n=11). Lymph node ratio: grade T&M stage, size (n=4)	Complete case analysis	Bayesian Model	Calibration curve: goodness-of-fit statistic 0.4847 Discrimination: *c*-statistic: 0.65 External Validation: absent; internal validation by random split method	Calculator
Pu *et al.* (2017)	Single institution database (n=220)	1,2,3 year survival	n=20	Pre-selection by Univariate analysis then Cox multivariate analysis. P-value <0.05	Differentiation, American Joint Commission on Cancer stage, Alkaline Phosphate to Albumin Ratio (n=3)	Complete case analysis	Multivariate Cox proportional hazards regression	Calibration curve: optimal consistency Discrimination: training: 0.673 validation: 0.693 External validation: absent; internal validation by random split method	Nomogram
Dasari *et al.* (2015)	Single institution database (n=567)	1,3 year survival	n=13	Pre-selection by Univariate analysis then Cox multivariate analysis. P-value <0.05	Tumour site, T stage, Lymph Node Ratio (n=3)	Complete case analysis	Multivariate Cox proportional hazards regression	Calibration: none Discrimination: AUC 1yr &3yr: 0.66 & 0.74 External Validation performed	Index
Pu *et al.* (2018)	National Registry Database (n=3458)	1,3,5 year survival	n=12	Pre-selection by Univariate analysis then Cox multivariate analysis.	Age, grade, T stage (n=3)	Complete case analysis	Multivariate Cox proportional hazards regression	Calibration curve: optimal fit Discrimination: *c*-statistic 0.63 External and internal	Nomogram

				P-value <0.05				validation performed using bootstrap method	
Katz *et al.* (2012b)	National Registry Database (n=5736)	3 year survival	n=7	Pre-selection by Univaraiate analysis then Cox multivaraite analysis. P-value <0.05	Age, gender, race, site, grade, stage, radiotherapy (n=7)	Complete case analysis	Multivaraite Cox proportional hazards regression	Calibration curve: results not reported Discrimination: absent External Validation: absent	Calculator

Table Iiii: Assessment of Risk of Bias of Included Studies Using Prediction model Risk of Bias Assessment Tool (PROBAST) (Wolff et al., 2014)

Study	Risk of Bias				Applicability			Overall	
	Participants	Predictors*	Outcome*	Analysis	Participants	Predictors	Outcome	Risk of Bias	Applicability
Brennan *et al.* (2004)	+	-	-	+	?	+	+	+	+
Kanda *et al.* (2014b)	?	-	-	-	?	+	+	-	+
Miura *et al.* (2014)	?	-	-	-	?	+	+	-	+
Shen *et al.* (2018)	?	-	-	-	+	+	+	-	+
Xu *et al.* (2017)	?	-	-	-	?	+	+	-	+
Walczak & Velanovich (2017)	?	?	-	-	?	+	+	?	+
Hsu *et al.* (2012)	+	-	-	?	?	+	+	+/?	+
Botsis *et al.* (2009)	?	-	-	-	?	+	+	-	+
Liu *et al.* (2018)	+	?	-	?	+	+	+	?	+
Balzano *et al.* (2017)	+	-	-	-	?	+	+	-	+
Smith & Mezhir (2014)	+	-	-	-	+	+	+	-	+
Pu *et al.* (2017)	?	-	-	-	?	+	+	-	+
Dasari *et al.* (2015)	+	-	-	-	?	+	+	-	+
Pu *et al.* (2018)	+	-	-	?	+	+	+	+	+
Katz *et al.* (2012b)	+	?	-	?	+	+	+	+/?	+

+ = low risk of bias/ low concern regarding applicability; - = high risk of bias/ high level of concern regarding applicability; ? = unclear level of risk of bias / concern regarding applicability; *blinding was an issue across all studies even where risk and concern were low in other areas

The number of model development studies, with (n=3) (Shen *et al*., 2018; Balzano *et al*., 2017; Dasari *et al*., 2016) and without (n=12) external validation (Pu *et al*., 2018; Brennan *et al*., 2004; Kanda *et al*., 2014b; Miura *et al*., 2014; Xu *et al*., 2017; Walczak & Velanovich, 2012; Hsu *et al*., 2012; Botsis *et al*., 2009; Liu *et al*., 2018; Smith & Mezhir, 2014; Pu *et al*., 2017; Katz *et al*., 2012b), increased sharply in recent years. Multivariable Cox regression proportional hazard regression was the most commonly employed modeling method (n=13) with 2 studies employing alternative techniques (Bayesian model: n=1(Smith & Mezhir, 2014); Artificial Neural Network (ANN): n=1 (Walczak & Velanovich, 2012) (Table Iii). 6 models could be applied preoperatively (Balzano *et al*., 2017; Kanda *et al*., 2014b; Miura *et al*., 2014; Hsu *et al*., 2012; Botsis *et al*., 2009; Pu *et al*., 2017). 5 studies focused on predicting poor prognosis (survival time under 7 months n=1, under 9 months n=1, under 12 months n=2, 6,12 and 18months survival n= 1) and one model predicted prognosis of 3 years or more (Table Iii). 7 models predicted prognosis at set time intervals (6months, 1, 3 and 5 years n=1; 1,2,3 years n=2; 1,3 years n=1; and 1,3 and 5 years n=3) and 2 studies did not categories survival time (Table Iii).

Source of data, participant selection and follow-up

A cohort design, commonly recommended for prognostic model development (Moons *et al*., 2009), was used across all 15 models. 5 studies used data from prospectively maintained databases (Balzano *et al*., 2017; Dasari *et al*., 2016; Brennan *et al*., 2004; Hsu *et al*., 2012; Liu *et al*., 2018), with 1 of these studies collecting data prospectively

alongside clinical trials (Liu *et al*., 2018). 7 studies used retrospective data (Shen *et al*., 2018; Kanda *et al*., 2014b; Miura *et al*., 2014; Xu *et al*., 2017; Walczak & Velancovich, 2012; Botsis *et al*., 2009; Pu *et al*., 2017). 3 studies used data from the cancer data registry (Pu *et al*., 2018; Smith & Mezhir, 2014; Katz *et al*., 2012b). Prospective cohort designed is recommended as it enables optimal measurement of predictors and outcome (Bouwmeester *et al*., 2012). Retrospective cohorts are thought to yield poorer quality data (Moons *et al*., 2009) but do enable longer follow-up time (Bouwmeester *et al*., 2012).

Participant recruitment was well described with inclusion criteria and description of cohort characteristics as well as study dates reported in all 15 studies. Length of follow-up time was clear in 14 studies (Shen *et al*., 2018; Balzano *et al*., 2017; Dasari *et al*., 2016; Brennan *et al*. 2004; Kanda *et al*., 2014b; Miura *et al*., 2014; Xu *et al*., 2017; Walczak & Velanovich 2012; Hsu *et al*., 2012; Botsis *et al*., 2009; Liu *et al*., 2018; Smith & Mezhir, 2014; Pu *et al*., 2017; Katz *et al*., 2012b). Consecutive sampling was reported in 3 studies (Shen et al., 2018; Miura et al., 2014; Walczak & Velanovich, 2012) but whether all consecutive participants were included, or number of participants who refused to participate, could not be evaluated as this was rarely reported across all studies. Non-consecutive sampling can introduce a risk of bias (Altman, 2001; Altman *et al*., 2009; Altman *et al*., 2001). The majority of models were developed using single centre databases (n=10) (Balzano *et al*., 2017; Dasari *et al*., 2016; Brennan *et al*., 2004; Kanda *et al*., 2014b; Miura *et al*., 2014; Xu *et al*., 2017; Walczak & Velanovich, 2012; Hsu *et al*., 2012; Botsis *et al*., 2009; Pu *et al*., 2017) which can limit the generalisability of the model. This was followed by use of cancer registry database (n=3)

(Pu *et al*., 2018; Smith & Mezhir, 2014; Katz *et al*., 2012b) and multicentre databases (n=2) (Shen *et al*., 2018; Liu *et al*., 2018).

Model Outcomes

In all 15 studies outcomes were clearly defined with the same outcome definition and method of measurement applied at all patients. However none of the studies reported blinding the outcome measurement for predictor values. Best practice dictates that assessor of the outcome occurrence should be blinded to ascertainment of the predictor (Moons *et al*., 2009; Laupacis *et al*., 1997) so as not to bias estimation of predictor effects for the outcome (Moons *et al*., 2009; Bouwmeester *et al*., 2012). Although such a bias would not be a major factor in prediction of all cause mortality (Bouwmeester *et al*., 2012; Moons *et al*., 2014), the majority of studies predicted disease-specific prognosis, whereby bias could come into play in variables requiring subjective interpretation, such as results from imaging (Moons *et al*., 2014).

Candidate Predictors

A variety of candidate predictors were considered across all 15 model development studies (Table Iiv). The mean number of candidate predictors was 19.47 (range 7 to 50). The definition, method and timing of measurement of candidate predictors were clear across all 15 studies although, as previously discussed lack of blinding was an issue. 3 studies reported categorisation of candidate predictor variables prior to model development (Shen *et al*., 2018;

Dasari *et al*., 2016; Brennan *et al*., 2004). 10 studies specifically detailed how categorical data was analysed as non-binary (Shen *et al*., 2018; Balzano *et al*., 2017; Brennan *et al*., 2004; Kanda *et al*., 2014b; Miura *et al*., 2014; Xu *et al*., 2017; Hsu *et al*., 2012; Botsis *et al*., 2009; Smith & Mezhir, 2014; Pu *et al*., 2017). 13 studies detailed how time-to-event data was analysed as non-binary (Shen *et al*., 2018; Balzano *et al*., 2017; Dasari *et al*., 2016; Pu *et al*., 2018; Kanda *et al*., 2014b; Miura *et al*., 2014; Xu *et al*., 2017; Hsu *et al*., 2012; Botsis *et al*., 2009; Liu *et al*., 2018; Smith & Mezhir, 2014; Pu *et al*., 2017; Katz *et al*., 2012b). Handling such data as binary is not recommended practice as this can result in less accurate predictions, as with dichotomizing predictor variables (Royston *et al*., 2006).

Table Iiv: Summary of frequency of included variables in prognostic model development studies

Variable	Number of Models Variable is Included in	Combined study population of all models in which the variable is included
Tumour grade	9	18,815
Age	8	17,565
Tumour Stage	7	14,630
Tumour size	5	8,154
Gender	4	12,910
Ca19-9	4	1,815
Vascular Involvement	3	1,059
Tumour location	3	6,858
T stage	3	10,433
Margin status	2	774
Lymph node involvement	2	794
Back pain	2	851
CEA	2	374
Lymph node ratio	2	6,967
Co-morbidities	2	1,036
Race	2	12,136
Splenectomy	1	555
Posterior margin positive	1	555
Weight loss	1	555
Platelet count	1	50
Neural Involvement	1	265
Neutrophil Lymphocyte Ratio	1	265
Platelet to lymphocyte ratio	1	265
Albumin to globulin ratio	1	265
Quality of Life	1	219
Adjuvant therapy	1	219
Radiotherapy	1	12,136

Alkaline Phosphate	1	218
Albumin	1	218
Alkaline phosphate to albumin ratio	1	220
Geriatric Nutritional Index	1	296
Non metastatic liver disease or insulin resistance	1	296
Marital status	1	6400

Statistical Power: sample size and missing data

Mean sample size was 1367 (range 50-6400). Event per variable (EPV) is the number of predictors assessed compared to the number of events. Statistical power of 10 studies (Shen *et al*., 2018; Balzano *et al*., 2017; Pu *et al*., 2018; Brennan *et al*., 2004; Miura *et al*., 2014; Walczak & Velanovich, 2012; Hsu *et al*., 2012; Liu *et al*., 2018; Pu *et al*., 2017; Katz *et al*., 2012b) could be assessed using the recommended EPV rule of statistical power for Cox regression models of 10 events per candidate predictor, as determined by the smallest group (Harrell, 2001; Peduzzi *et al*., 1996; Peduzzi *et al*., 1995; Steyerberg *et al*., 1999; Vittinghoff & mcCulloch, 2007). Of these studies 5 did not achieve statistical power according to this rule (Shen *et al*., 2018; Balzano *et al*., 2017; Miura *et al*., 2014; Walczak & Velanovich, 2012; Pu *et al*., 2017). Recently an EPV of 10 has been criticised as being too simplistic for calculating minimum sample size required for models predicting binary and time-to-event outcomes (Riley et al., 2019). Instead there is a move toward applying the following three criteria to determine the minimum

sample size required for such models: (i) predictor effect estimates defined by a global shrinkage factor of ≥0.9, (ii) small absolute difference in the model's apparent and adjusted Nagelkerke's R^2 (≤ 0.05), and (iii) precise estimation of the overall risk in the population. Initial testing of this approach suggests that it will minimise overfitting and ensure precise estimates of overall risk (Riley *et al*., 2019).

Most studies (n=9) used complete case analysis (Balzano *et al*., 2017; Dasari *et al*., 2016; Pu *et al*., 2018; Kanda *et al*., 2014b; Miura *et al*., 2014; Walczak & Velanovich, 2012; Hsu *et al*., 2012; Botsis *et al*., 2009; Liu *et al*., 2018; Smith & Mezhir, 2014; Pu *et al*., 2017; Katz *et al*., 2012b). This approach results in loss of statistical power and can introduce bias as missing data rarely occurs randomly and often pertains to participant or disease characteristics (Bouwmeester *et al*., 2012). 2 studies reported missing data per candidate variable (Brennan *et al*., 2004; Botsis *et al*., 2009). 1 of these studies handled missing data by predicting input using regression modelling (Brennan *et al*., 2004). The other study handled missing data by applying the Multivariate Imputation by Chained Equations (MICE) method assuming data were missing at random (MAR) (Botsis *et al*., 2009). Imputation, particularly multiple imputation, of missing data is advocated to reduce bias and maintain statistical power (Harrell *et al*., 1996; Donders *et al*., 2006; Marshall *et al*., 2010). 4 studies did not give details of missing data (Shen *et al*., 2018; Balzano *et al*., 2017; Xu *et al*., 2017; Liu *et al*., 2018).

Model development

All 15 studies detailed how many candidate predictors were considered but none of the studies detailed how candidate predictors were selected with prior expert knowledge of disease inferred. 1 study selected predictors on multivariable analysis (Brennan *et al*., 2004). Most studies (n=12) employed pre-selection by univariable analysis of predictors for inclusion in multivariable analysis (Shen *et al*., 2018; Balzano *et al*., 2017; Dasari *et al*., 2016; Pu *et al*., 2018; Kanda *et al*., 2014b; Miura *et al*., 2014; Xu *et al*., 2017; Hsu *et al*., 2012; Botsis *et al*., 2009; Liu *et al*., 2018; Pu *et al*., 2017; Katz *et al*., 2012b). Although this method is commonly employed it is not recommended as it carries a greater risk of predictor selection bias, particularly in smaller sample sizes (Collins *et al*., 2015). Predictors not significant in univariable analysis may become significantly associated with outcome following adjustment for other predictors (Moons *et al*., 2014). Predictors pre-selected due to large but spurious association with outcome can result in increased risk of overfitting (Moons *et al*., 2014). Furthermore multivariable analysis for predictor selection can result in overfitting and unstable models (Bouwmeester *et al*., 2012). This is a particular risk when outcomes are few but many predictors are analysed (Bouwmeester *et al*., 2012). None of the studies described shrinkage technique as a method for addressing possible overfitting (Moons *et al*., 2014). In the case of low EPV, shrinkage methods could not account for all bias (Moons *et al*., 2014).

14 studies used backward elimination methods (Shen *et al*., 2018; Balzano *et al*., 2017; Dasari *et al*., 2016; Pu *et al*., 2018; Kanda *et al*., 2014b; Miura *et al*., 2014; Xu *et al*., 2017; Walczak & Velanovich 2012; Hsu *et al*., 2012; Botsis *et al*., 2009; Liu *et al*., 2018; Smith & Mezhir, 2014; Pu *et al*., 2017; Katz *et al*., 2012b). This included an ANN that used single hidden layer back propagation to train the model (Walczak & Velanovich, 2012), and a Bayesian model that employed backward step down selection process (Smith & Mezhir, 2014). Of the remaining 12 studies employing this method nominal *P*-value was used as the criteria for predictor inclusion. 10 of these studies used *P*-value <0.05 (Shen *et al*., Dasari *et al*., 2016; Pu *et al*., 2018; Miura *et al*., 2014; Xu *et al*., 2017; Hsu *et al*., 2012; Botsis *et al*., 2009; Liu *et al*., 2018; Pu *et al*., 2017; Katz *et al*., 2012b), 2 of which also reported additionally using Akaike Information Criteria (Botsis *et al*., 2009; Liu *et al*., 2018).1 study used *P*-value < 0.1 (Kanda *et al*., 2014b), and 1 study used *P*-value < 0.2 for univariate analysis and *P*-value < 0.1 for multivariate analysis (Balzano *et al*., 2017). The use of a small *P*-value has the benefit generating a model from fewer predictors but carries the risk of missing potentially important variables whilst the use of larger *P*-values potentiates inclusion of predictors of less importance (Moons *et al*., 2014). 1 study reported using multivariable analysis for predictor selection determined by *P*-value but then included non-significant factors in the final model to include all 7 candidate variables, therefore effectively employing full model approach (Brennan *et al*., 2004). Whilst full model approach can avoid selection bias (Moons *et al*., 2014), the potential for selection bias still remained in this study, as details were not given on how candidate predictors were decided.

Predictor selection can also incur bias when continuous predictors are categorised (Moons *et al*., 2014). 12 studies reported categorisation (Shen *et al*., 2018; Balzano *et al*., 2017; Dasari *et al*., 2016; Pu *et al*., 2018; Brennan *et al*., 2004; Kanda *et al*., 2014b; Miura *et al*., 2014; Xu *et al*., 2017; Hsu *et al*., 2012; Botsis *et al*., 2009; Smith & Mezhir, 2014; Pu *et al*., 2017). 3 studies specifically stated that categorization was performed prior to modelling (Shen *et al*., 2018; Dasari *et al*., 2016; Brennan *et al*., 2004). All 12 studies described appropriate statistical techniques for handling continuous variables.

Model performance and evaluation

8 studies reported calibration of their model (Shen *et al*., 2018; Balzano *et al*., 2017; Pu *et al*., 2018; Brennan *et al*., 2004; Xu *et al*., 2017; Smith & Mezhir, 2014; Pu *et al*., 2017; Katz *et al*., 2012b), most commonly presented as calibration curve. 1 study reported Hosmer-Lemeshow test (Balzano *et al*., 2017), a test sometimes criticised for limited statistical power to assess poor calibration and failure to indicate magnitude or direction of miscalibration (Moons *et al*., 2014). 12 studies reported discrimination measured as either *c*-statistic (n=4) (Shen *et al*., 2018; Brennan *et al*., 2004; Botsis *et al*., 2009; Smith & Mezhir,2014) or area-under-the-curve (AUC) of the receiver operated curve (n=4) (Balzano *et al*., 2017; Dasari *et al*., 2016; Kanda *et al*., 2014b; Walczak & Velanovich, 2012) or both (n=4) (Pu *et al*., 2018; Xu *et al*., 2017; Liu *et al*., 2018; Pu *et al*., 2017). Although commonly used, the *c*-statistic can be influenced by predictor value distribution and be insensitive to inclusion of additional predictors (Moons *et al*., 2014). 9 studies reported

confidence intervals with discrimination measures (Brennan *et al*., 2004; Shen *et al*., 2018; Balzano *et al*., 2017; Dasari *et al*., 2016; Pu *et al*., 2018; Xu *et al*., 2017; Botsis *et al*., 2009; Smith & Mezhir, 2014; Pu *et al*., 2017). R^2 was reported in 1 study (Balzano *et al*., 2017). Sensitivity and specificity were also poorly reported (n=2) (Walczak & Velanovich, 2012; Liu *et al*., 2018). Internal validation was rarely performed. 3 studies used bootstrapping (Pu *et al*., 2018; Brennan *et al*., 2004; Botsis *et al*., 2009) and 4 studies used random split method (Walczak & Velanovich, 2012; Liu *et al*., 2018; Smith & Mezhir, 2014; Pu *et al*., 2017). 3 studies included external validation as part of model development (Shen *et al*., 2018; Balzano *et al*., 2017; Dasari *et al*., 2016). However, the external validation datasets were small. Shen et al. (2018) used 17 variables and the external validation dataset contained only 61 patients. Balzano *et al*. (2017) used 56 variables, using univariable analysis to select for multivariable analysis, but the derivation set had only 78 patients and the external validation dataset had only 43 patients. In one of these studies it was unclear how many events occurred in the external validation cohort (Dasari *et al*., 2016). None of the studies described external validation of their models separate to the derivation authors and none of the studies described impact analysis of their models.

Presentation of Results and Discussion

12 studies presented both unadjusted and adjusted results of the full model with all candidate predictors considered (Shen *et al*., 2018; Balzano *et al*., 2017; Dasari *et al*., 2016; Pu *et al*., 2018; Kanda *et al*.,

2014b; Miura *et al*., 2014; Xu *et al*., 2017; Hsu *et al*., 2012; Botsis *et al*., 2009; Liu *et al*., 2018; Pu *et al*., 2017; Katz *et al*., 2012b) and 1 study presented adjusted results only (Brennan *et al*., 2004). All 15 studies offered alternative presentation of the model. The most common form of presentation of prognostic models was nomograms (n=5) and prognostic scores (n=5) followed by prognostic calculators (n=3) and prognostic index (n=2) (Table 20). All 15 studies reported interpretation of models as being for application to clinical practice and all studies discussed comparison, generalisability, strengths and weaknesses of their model as recommended by several guidelines including PRISMA statement (Moher *et al*., 2009).

The results of this critical review and discussion of the overall findings are further discussed in section 2.3.2.

Appendix J

A Review and Critical Analysis of Machine Learning Methods Applied to Decision Making in Pancreatic Cancer Management

The body of this appendix and the discussion in section 2.3.4 has been published in the review article written by the autor of this thesis:

Bradley, A., Van der Meer, R. and McKay, C. (2019) 'Personalized pancreatic cancer management: a systematic review of how machine learning is supporting decision-making'*. Pancreas,*48 (5). pp. 598-604.

The aim of this appendix is to support the discussion in section 2.3.4 regarding the current use of machine learning techniques to support clinical decision making in pancreatic cancer management. Machine learning has grown in popularilty in recent years as a means of modeling complex adaptive systems containing a large number of variables with a high degree of uncertainty. Some of the most commonly used machine learning methods are summarised in table Ji.

Table Ji: Summary of Common Methods of Machine Learning			
Method	Application	Strengths	Limitations
Bayesian Network (BN)	Decision support Risk Assessment Prediction (Abbod *et al*., 2014; Velikova *et al*., 2014; Verduijn *et al*., 2017; Lucas *et al*., 2004)	Allows for incorporation of individual patient data, disease progression and impact of different treatment options on the predicted outcome (Velikova *et al*., 2014; Verduijn *et al*., 2017). Facilitates prognosis updating and scenario testing (Verduijn *et al*., 2017). Provides information on process and outcome variables therefore predict outcomes pertaining to quality and not just amount of survival time (Verduijn *et al*., 2017; Lucas *et al*., 2004). Uses probabilistic inference when data is limited and can still make predictions based on global averages of the patient population (Verduijn *et al*., 2017; Lucas *et al*., 2004).	Accurate use of data in elicitation of priors is an area of ongoing investigation and debate (Lucas *et al*., School *et al*., 2013; Hampson *et al*., 2014; Johnson *et al*., 2010). An over reliance on machine-learned network structures, could mean fundamental causal relationships well established in medical knowledge are lost hence limiting the applicability (Lucas *et al*., School *et al*., 2013; Hampson *et al*., 2014; Johnson *et al*., 2010). Can only model linear dependencies (Abbod *et al*., 2014).
Artificial Neural Network (ANN)	Modeling Prediction Image interpretation Classification(Abbod *et al*., 2014)	Models non-linearity and complex relationships (Abbod *et al*., 2014; Bartosch-Härlid *et al*., 2008). Handles high-dimension problems(Abbod *et al*., 2014; Bartosch-Härlid *et al*., 2008). Can generalize (Bartosch-Härlid *et al*., 2008).	Heavy data requirements with long training times requiring many design decisions (Abbod *et al*., 2014; Bartosch-Härlid *et al*., 2008). May not generalize well to other data sets (Abbod *et al*., 2014; Bartosch-Härlid *et al*., 2008). Lacks transparency(Abbod

		Does not impose any restrictions on the input variables (Abbod *et al*., 2014; Bartosch-Härlid *et al*., 2008).	*et al*., 2014; Bartosch-Härlid *et al*., 2008).
Fuzzy Logic (FL)	Modeling Prediction Classification (Abbod *et al*., 2014)	Models non-linearity (Abbod *et al*., 2014). Handles uncertainty and complexity (Johnson *et al*., 2010; Gursel, 2016; Barro & Marín, 2002; Dweiri & Kablan, 2006). Enables prediction to move from probability to plausibility (Grossi, 2015). Transition to a contiguous value is gradual rather than abrupt reflecting human decision-making processes (Gursel, 2016; McNeill &Thro, 1994; Bouchon-Meunier & Zadeh, 1995). Can assess more observed variables yet fewer values are required (McNeill &Thro, 1994; Bouchon-Meunier & Zadeh, 1995). Transparent (Abbod *et al*., 2014).	Extensive expert knowledge of the system to be modeled is required (Roychowdhury *et al*., 2004). Requires more fine-tuning and simulation prior to being operational (Pratihar *et al*., 1999). Cannot model high-dimension problems (Abbod *et al*., 2014).

Following a comprehensive search of MEDLINE, Embase, PubMed and Cochrane databases only 6 studies were identified that met the inclusion criteria of machine learning methods applied to predictive modeling and decision-analysis related to pancreatic cancer management. Three studies were Markov decision-analysis models

comparing two competing treatment options: neoadjuvant therapy *versus* upfront surgery (deGus *et al*., 2016; Sharma *et al*., 2015; Van Houten *et al*., 2012). Three studies focused on predicting survival time (Smith & Mezhir, 2014; Hayward *et al*., 2010; Walczak & Velanovich, 2017). One of these studies also predicted lymph node ratio (Smith & Mezhir, 2014). One of these studies additionally explored prediction of Eastern Cooperative Oncology Group (ECOG) quality-of-life scores, surgical outcomes and tumour characteristics (Hayward *et al*., 2010). One study performed direct comparison between predictive accuracy of machine learning techniques and linear and logistic regression (Hayward *et al*., 2010).

Three studies used Marko decision tree models (deGeus *et al*., 2016; Sharma *et al*., 2015; Van Houten *et al*., 2012), 1 study used Bayesian modeling (Smith & Mezhir, 2014) 1 study used ANN (Walczak & Velanovich, 2017) and 1 study explored machine learning algorithms including: BN, decision trees, *k*-nearest neighbor, and ANN (Hayward *et al*., 2010)(Table Jii).

Table Jii: Summary of Included Machine Learning Studies				
Study	**Participant Population**	**Method**	**Outcome Measure**	**Main Limitations**
deGeus *et al.* (2016)	Synthesised data from phase II trials and cohort studies	Markov decision-analysis	Survival in months and quality adjusted life months for upfront surgery versus neoadjuvant therapy	Use of single electronic database of journals Synthesised small underpowered studies with high level if heterogeneity Relied heavily on retrospective cohort studies
Sharma et al. (2015)	Synthesised data from phase II trials	Markov decision-analysis	Survival in months and quality adjusted life months for upfront surgery versus neoadjuvant therapy	Use of single electronic database of journals Synthesised small underpowered studies with high level if heterogeneity
Van Houten et al. (2012)	Synthesised data from phase II trials and cohort studies	Markov decision-analysis	Survival in months and quality adjusted life months for upfront surgery versus neoadjuvant therapy	Use of single electronic database of journals Synthesised small underpowered studies with high level if heterogeneity Included borderline resectable cases in neoadjuvant cohort Relied heavily on retrospective cohort studies
Smith & Mezhir (2014)	Cancer Registry (n = 6400)	Interactive Bayesian Model	Survival at 6 months, 1,3 and 5-year survival	Follow-up time unclear Unclear if consecutive sampling used Selection method of candidate predictors not clear Complete base analysis used No external validation

Walczak &Velanovich (2017)	Retrospective single institution database (n = 219)	Artificial Neural Network	Death at 7 months post resection	Consecutive sampling used but unclear all consecutive participants included Selection method of candidate predictors not clear Complete base analysis used No external validation No calibration
Hayward et al. (2010)	Retrospective single institution database (n = 91)	Machine learning algorithms including: Bayesian Network, decision trees, *k*-nearest neighbor, and ANN	Survival as time dependent event, Eastern Cooperative Oncology Group (ECOG) quality-of-life scores measured at 6 months	Unclear if consecutive sampling used Complete base analysis used No external validation No calibration

Decision-analysis Models

Three studies attempted to employ Markov decision analysis to compare upfront surgery and neoadjuvant approach (deGeus *et al*., 2016; Sharma *et al*., 2015; Van Houten *et al*., 2012). Sharma *et al*. (2015) used data drawn from 21 prospective phase II and III trials. De Gus *et al.* (2016) also included data from retrospective studies compiled from a literature search from a single search engine. Both these studies, although reportedly analysing strategies for resectable pancreatic cancer used studies that included borderline resectable and locally advanced pancreatic cancer in an intention-to-treat analysis (deGus *et al*., 2016; Sharma *et al*., 2015). All 3 studies used an intention-to-treat approach to analysis and, although they reported a slight benefit from neoadjuvant approach, neither strategy was conclusively superior (deGus *et al*., 2016; Sharma *et al*., 2015;

Van Houten *et al*., 2012). All 3 existing studies were solely based on synthesised evidence from published trials therefore share the limitations of the existing body of evidence mainly: heterogeneity and small underpowered sample size.

Prediction Models

A cohort design, commonly recommended for prognostic model development (Moons *et al*., 2009), was used for all three predictive models (Smith & Mezhir, 2014; Hayward et al., 2010; Walczak & Velanovich, 2017). Two studies used retrospective single centre databases (ANN n = 219 (Walczak & Velanovich, 2017); comparison study n = 91 Hayward et al., 2010)), which can limit generalisability, and 1 study used cancer data registry (BN n = 6,400) (Smith & Mezhir, 2014). Prospective cohort designed is recommended as it enables optimal measurement of predictors and outcome (Bouwneester *et al*., 2012). Retrospective cohorts are thought to yield poorer quality data but do enable longer follow-up time (Moons *et al*., 2009).

Participant recruitment with inclusion criteria and description of cohort characteristics were well reported, as were study dates in all 3 studies (Smith & Mezhir, 2014; Hayward et al., 2010; Walczak & Velanovich, 2017). Length of follow-up time was clear in 2 studies (Smith & Mezhir, 2014; Walczak & Velanovich, 2017). Consecutive sampling was reported in 1 study (Walczak & Velanovich, 2017) but whether all consecutive participants were included, or number of participants who refused to participate, could not be evaluated in any of the 3 studies (Smith & Mezhir, 2014; Hayward et al., 2010; Walczak & Velanovich, 2017). Non-consecutive sampling can

introduce a risk of bias (Altman, 2001; Altman *et al*., 2009; Altman *et al*., 2001).

In all 3 studies outcomes were clearly defined with the same outcome definition and method of measurement applied to all patients (Smith & Mezhir, 2014; Hayward et al., 2010; Walczak & Velanovich, 2017). The interactive Bayesian model predicted 6month, 1,3 and 5year survival post resection and lymph node ratio (Smith & Mezhir, 2014). The ANN predicted 7-month mortality after resection (Walczak & Velanovich, 2017). Hayward *et al.* (2010) focused on data mining techniques but treated survival outcome as a time-dependent-event for resected and un-resected patients, with ECOG measured at 6 months post-resection. Number of candidate variables ranged from 7 to 19. The definition, method and timing of measurement of candidate predictors were clear in all 3 studies (Smith & Mezhir, 2014; Hayward et al., 2010; Walczak & Velanovich, 2017). How candidate predictors were selected were not made clear in 2 studies (Smith & Mezhir, 2014; Walczak & Velanovich, 2017) which may be illustrative of the non-transparent 'black-box' analysis sometimes employed by forms of artificial intelligence (AI). One study extensively explored algorithms for data mining and categorisation of the datasets (Hayward et al., 2010). The other 2 studies used backward elimination methods (Smith & Mezhir, 2014; Walczak & Velanovich, 2017). The ANN used single hidden layer back propagation to train the model (Walczak & Velanovich, 2017), and the Bayesian model employed backward step down selection process (Smith & Mezhir, 2014). All 3 studies used complete case analysis (Smith & Mezhir, 2014; Hayward et al., 2010; Walczak & Velanovich,

2017). This approach results in loss of statistical power and can introduce bias as missing data rarely occurs randomly and often pertains to participant or disease characteristics (Bouwneester et al., 2012).

None of the studies underwent external validation. The interactive Bayesian model (Smith & Mezhir, 2014) and ANN (Walczak & Velanovich, 2017) employed random split technique between training and validation datasets. This points to a potential key weakness in the application of machine learning techniques as random split technique can result in over and under fitting of the model, particularly as details of cross validation were not given (Reitermanov´a, 2010). Techniques of data splitting are poorly described and can result in a high degree of variance of model performance (Reitermanov´a, 2010). More sophisticated techniques of data splitting that exploit the structure of the data exist and provide more confident results, but at higher computational cost (Reitermanov´a, 2010). Only the interactive Bayesian model reported calibration with goodness-of-fit statistic (P = 0.300 for prediction of lymph-node-ratio; P = 0.4847 for survival prediction) (Smith & Mezhir, 2014). The ANN reported discrimination as area under curve (AUC) of the receiver operated curve (ROC) (AUC, 0.6576; sensitivity, 91.30%; specificity, 38.27%) (Walczak & Velanovich, 2017). The interactive Bayesian model reported discrimination as *c*-statistic (0.65; 95% CI, 0.63-0.66) (Smith & Mezhir, 2014). Although commonly used, the *c*-statistic can be influenced by predictor value distribution and be insensitive to inclusion of additional predictors (Moons *et al*., 2014). The study by Hayward *et al.* (2010) compared

machine learning to log regression and found that for survival prediction Bayesian modeling out performed log regression (accuracy 0.60 *versus* 0.42). Furthermore in predicting outcome for ECOG at 6 months post-resection log regression performance improved from *r*-squared value, 0.26 to 0.32 when modified with machine learning algorithm 'linear regression with bagging'(Hayward *et al.*, 2010).

Conclusion

Of the 6 existing studies that apply machine learning techniques to support clinical decision making in the management of pancreatic cancer 3 studies used Markov decision tree models to perform decision analysis (deGus *et al*., 2016; Sharma *et al*., 2015; Van Houten *et al*., 2012). Three studies used machine learning methods for predictive modeling: 1 study used Bayesian modeling (Smith & Mezhir, 2014), 1 study used ANN (Walczak & Velanovich, 2017) and 1 study explored machine learning algorithms including: BN, decision trees, *k*-nearest neighbor, and ANN (Hayward *et al*., 2010).

The main issues identified with decision-analysis studies were reliance on data from a single database search and the quality of the existing studies pertaining to the treatment of potentially resectable pancreatic cancer being mainly small and underpowered with a high degree of heterogeneity (Tempero *et al*., 2014; Asare *et al*., 2016). The issues identified with the predictive models were overreliance on single institution retrospective databases, which could affect generalisability. There was also a lack of clarity as to whether

consecutive sampling was employed and how candidate predictors were selected. A major issue identified was the lack of external validation across all 3 predictive models. Although 2 studies used random-split technique, details of cross-validation were not provided which potentiates issues of over or under fitting. Only one study reported calibration of their model.

Appendix K

Bayesian Network Meta-analysis: Potentially Resectable Pancreatic Cancer

**SF= surgery first pathway; NAT = neoadjuvant therapy; SF+adj = surgery first plus adjuvant therapy; surgery only= surgical resection no adjuvant therapy*

Resection Rates: Phase II/III Studies

Figure Ki: Results of fixed effects and random effects (vague prior) models

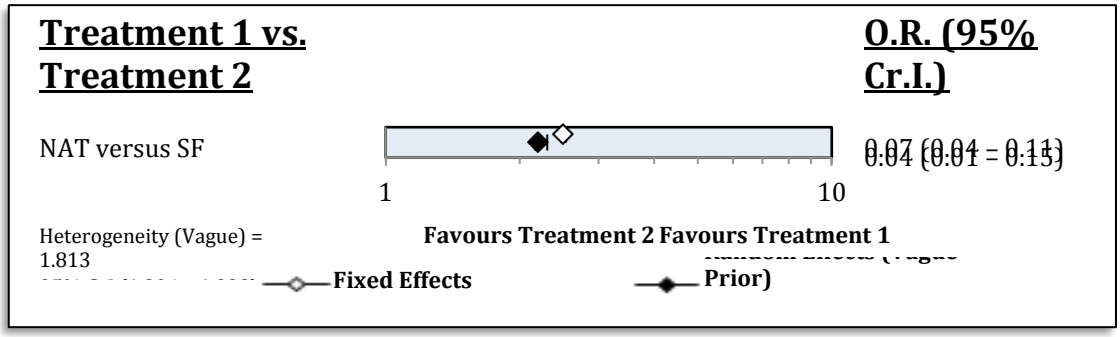
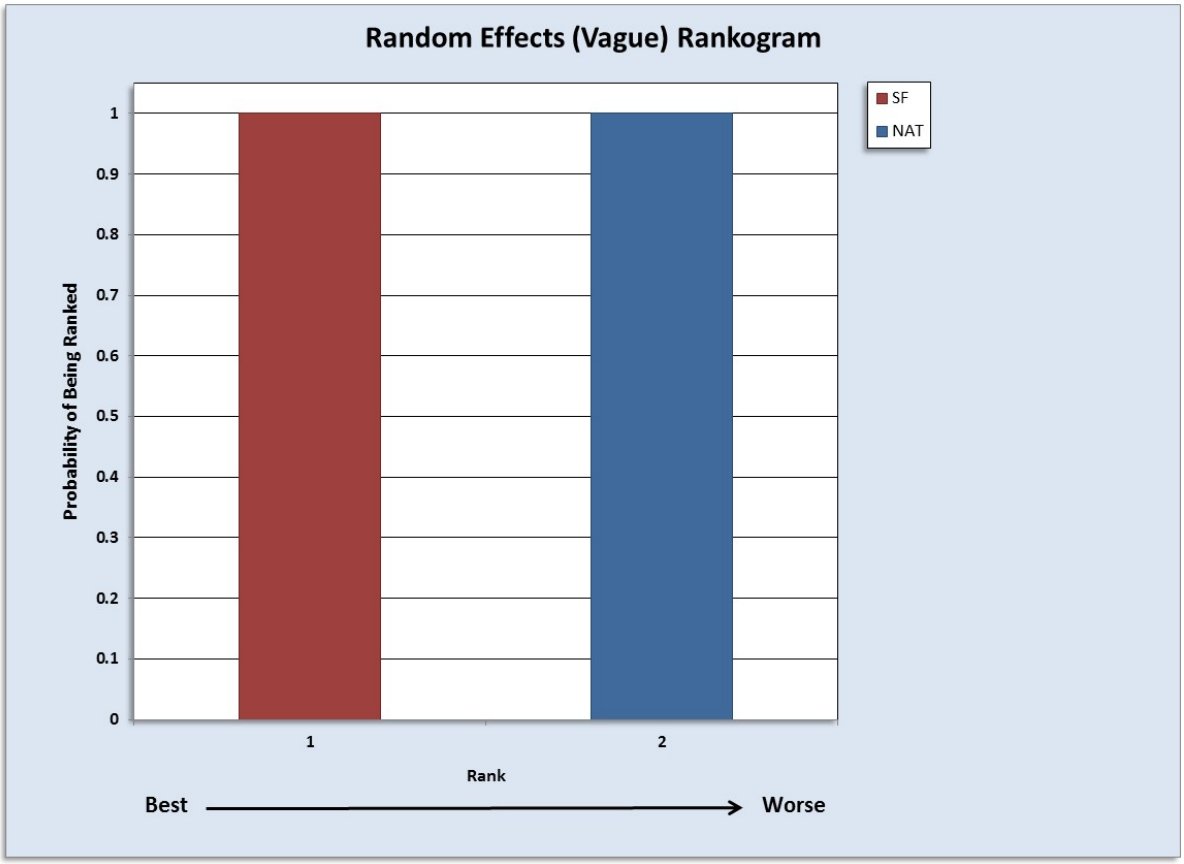


Figure Kii: League table based on results of fixed effects and random effects (vague prior) models. Where odds ratio (O.R.) is greater than 1 treatment at top left is superior

NAT	
0.07 (0.04 – 0.11)	**SF**

Figure Kiii: Rankogram summarising surface under the cumulative ranking (SUCRA).

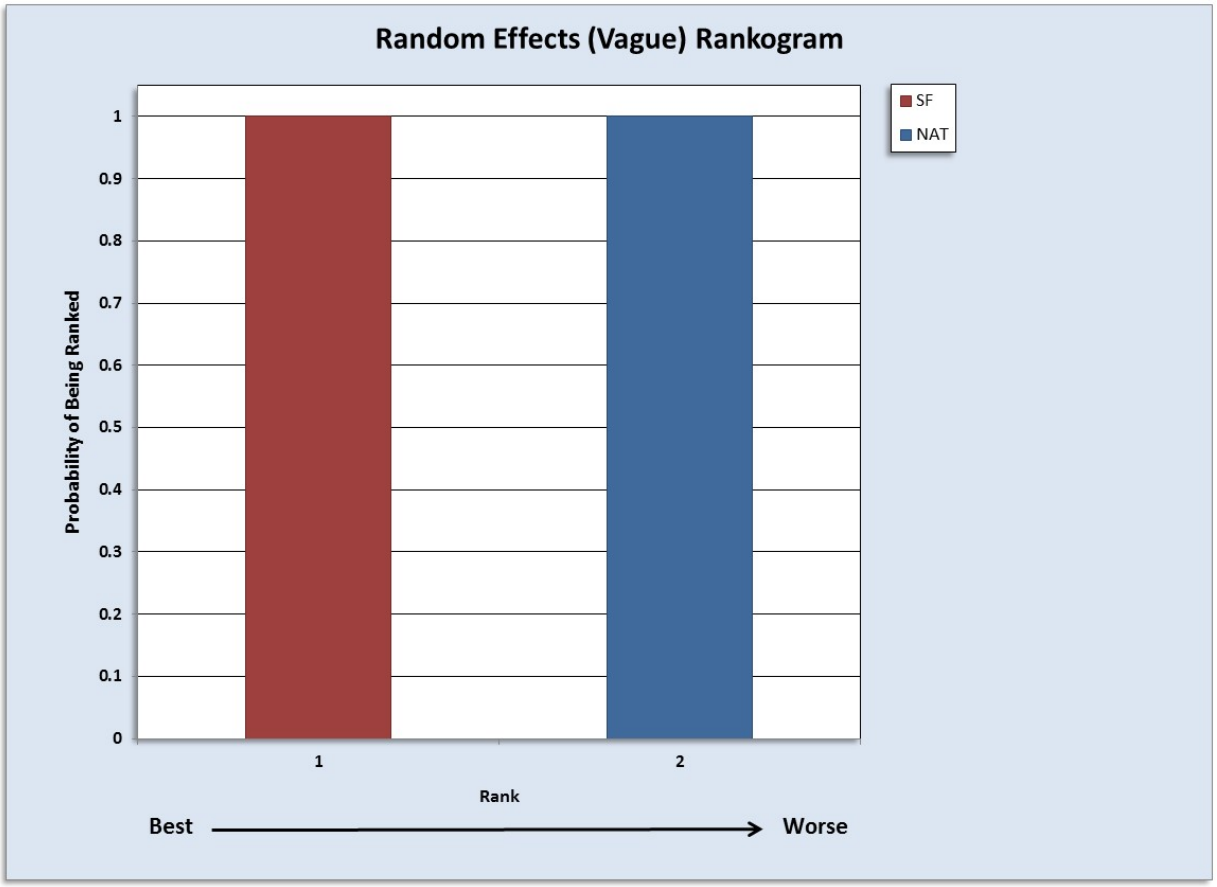


Resection Rates: Phase II/III plus Cohort studies

Figure Kiv: League table based on results of fixed effects and random effects (vague prior) models. Where odds ratio (O.R.) is greater than 1 treatment at top left is superior

NAT	
0.16 (0.13 – 0.20)	SF

Figure Kv: Rankogram summarising surface under the cumulative ranking (SUCRA).

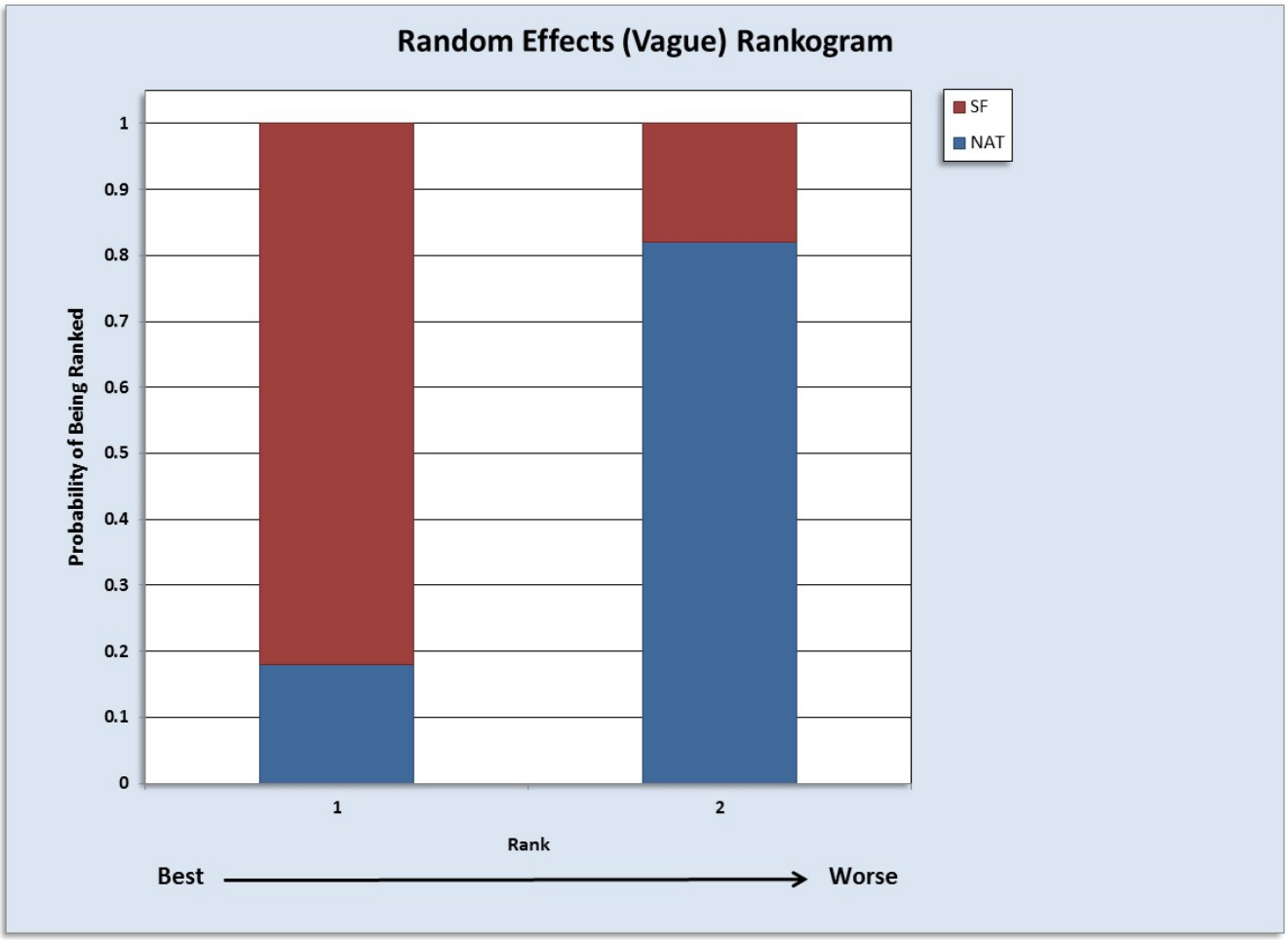


R0 Resection Rates: Phase II/III Studies

Figure Kvi: League table based on results of fixed effects and random effects (vague prior) models. Where odds ratio (O.R.) is greater than 1 treatment at top left is superior

NAT	
0.40 (0.28 – 0.57)	**SF**

Figure Kvii: Rankogram summarising surface under the cumulative ranking (SUCRA).



R0 Resection Rates: Phase II/III plus Cohort studies

Figure Kviii: Results of fixed effects and random effects (vague prior) models

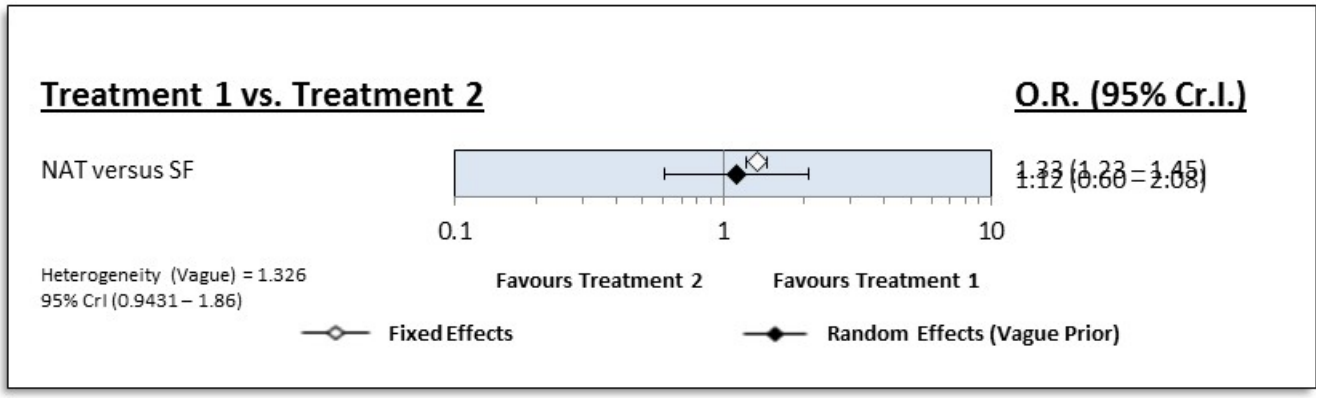
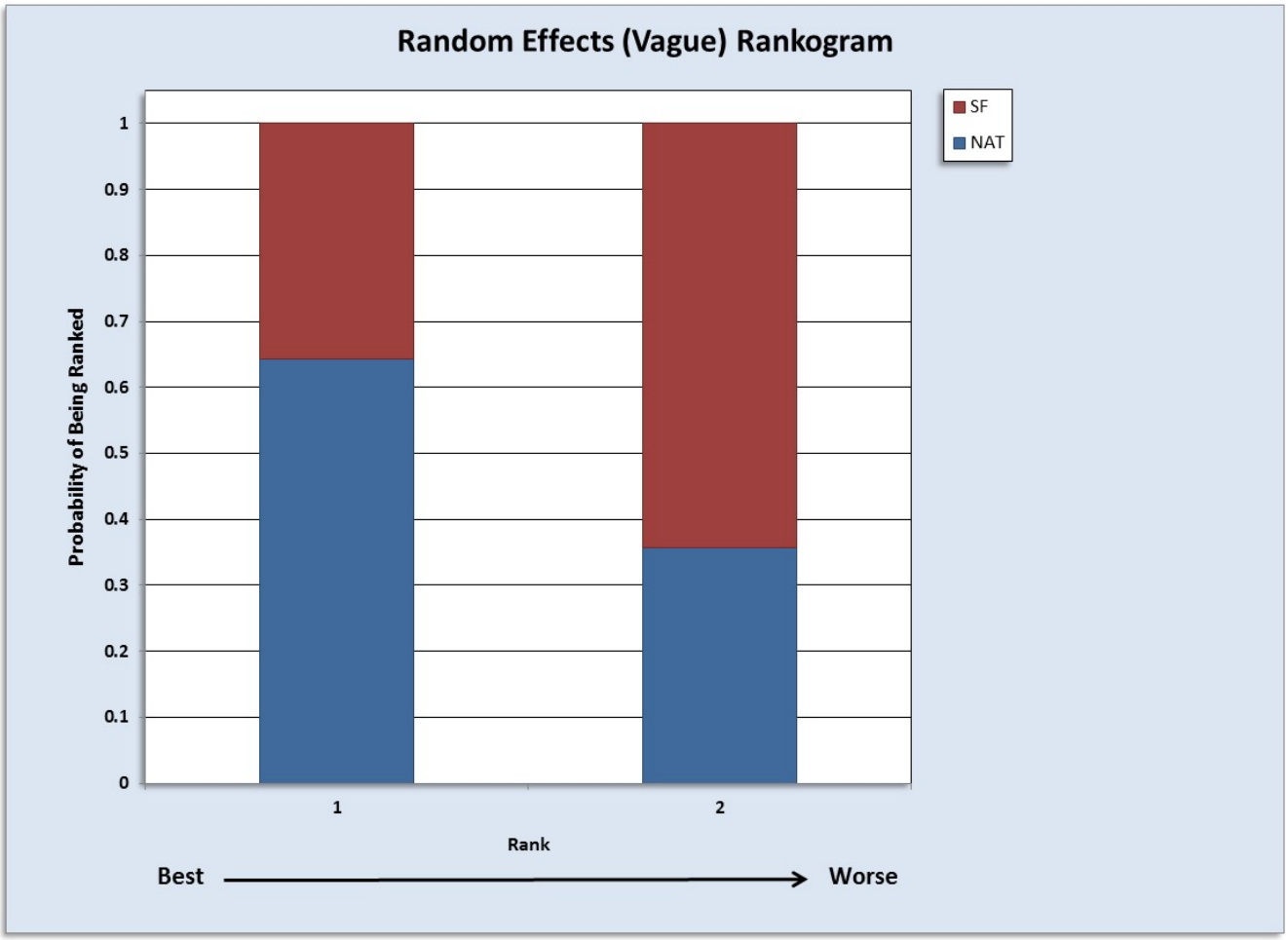


Figure Kix: League table based on results of fixed effects and random effects (vague prior) models. Where odds ratio (O.R.) is greater than 1 treatment at top left is superior

NAT	
1.33 (1.23 – 1.45)	**SF**

Figure Kx: Rankogram summarising surface under the cumulative ranking (SUCRA).



1-year Survival: Phase II/III Studies

Figure Kxi: Results of fixed effects and random effects (vague prior) models

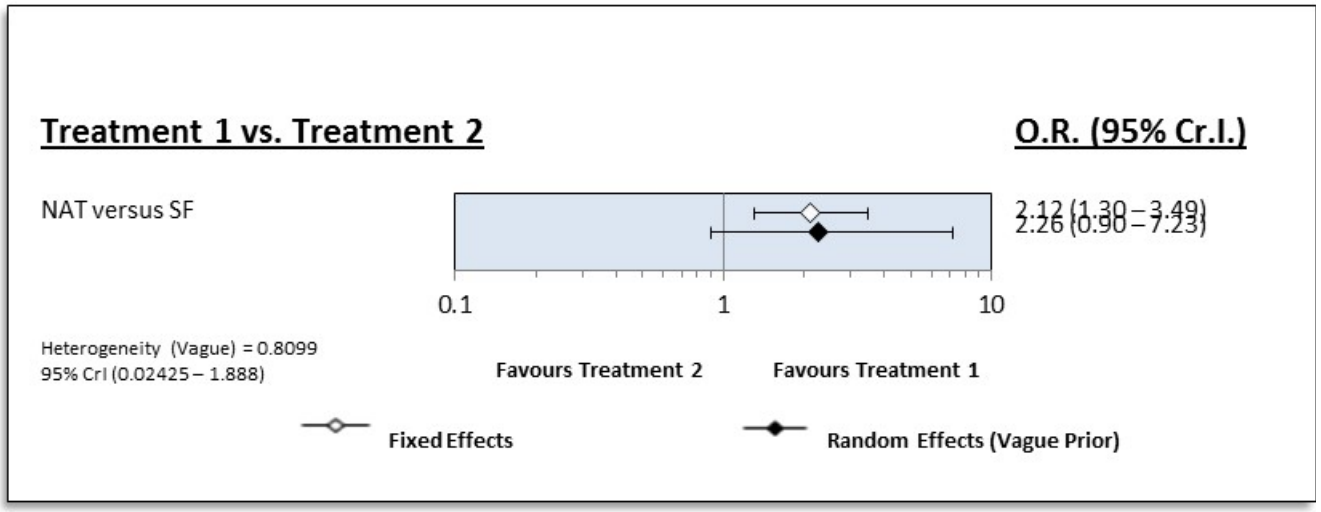
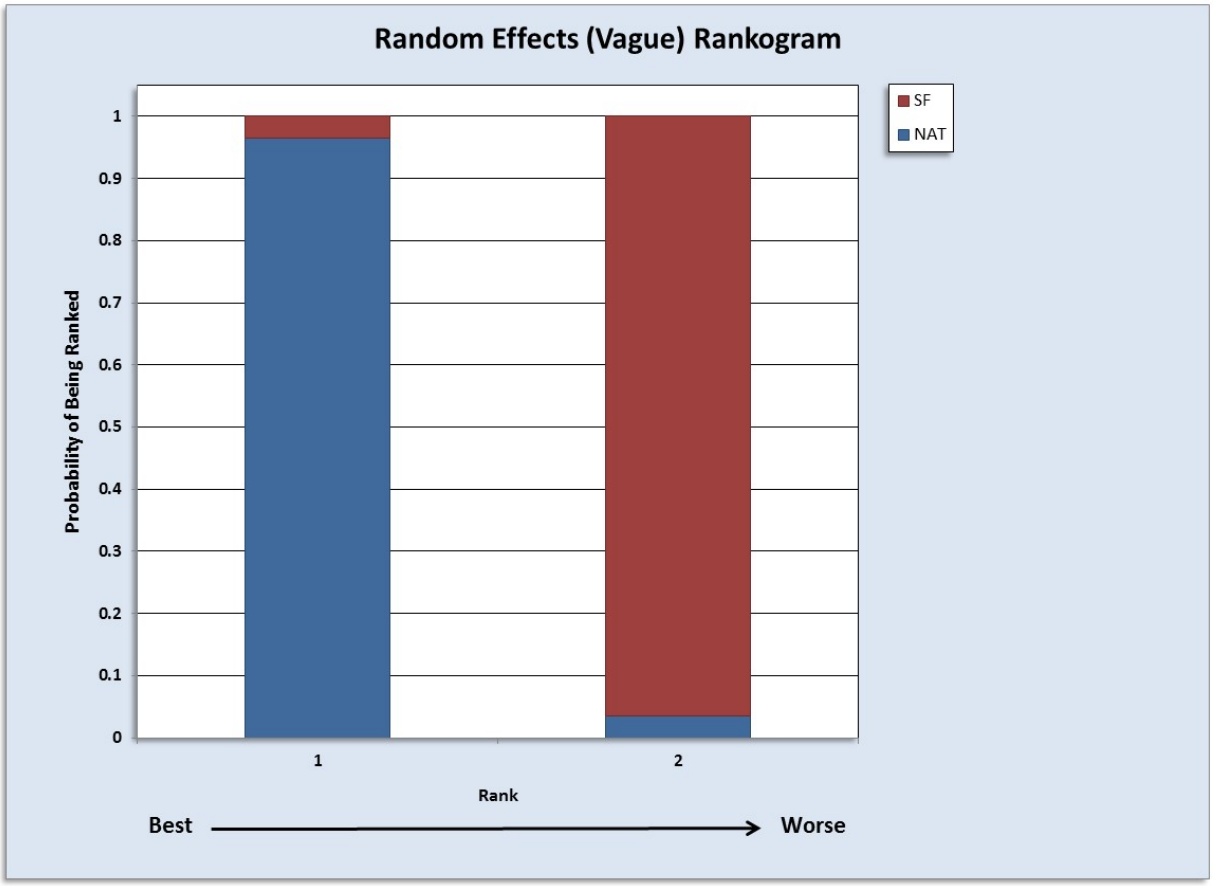


Figure Kxii: League table based on results of fixed effects and random effects (vague prior) models. Where odds ratio (O.R.) is greater than 1 treatment at top left is superior

NAT	
2.12 (1.30 – 3.49)	**SF**

Figure Kxiii: Rankogram summarising surface under the cumulative ranking (SUCRA).



1-year Survival: Phase II/III plus Cohort studies

Figure Kxiv: Results of fixed effects and random effects (vague prior) models

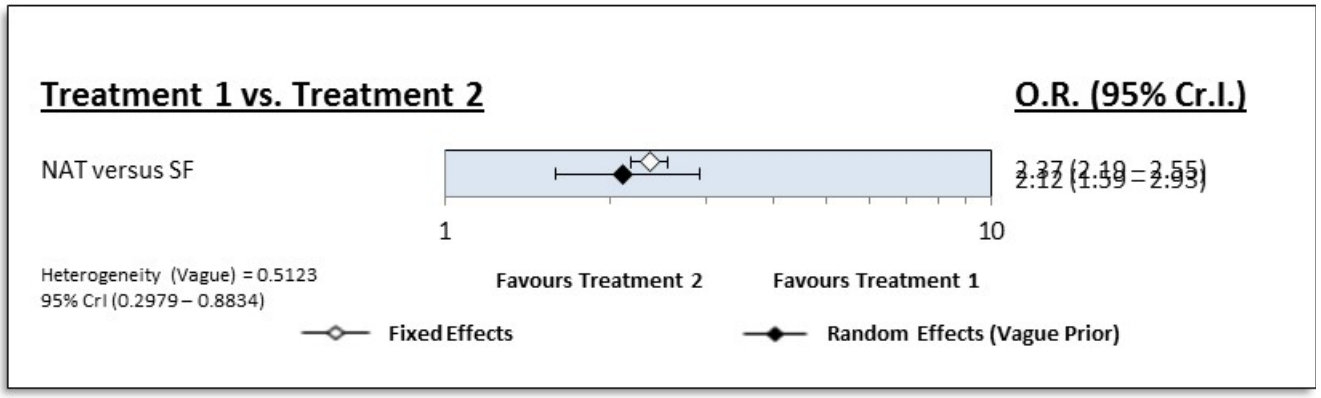
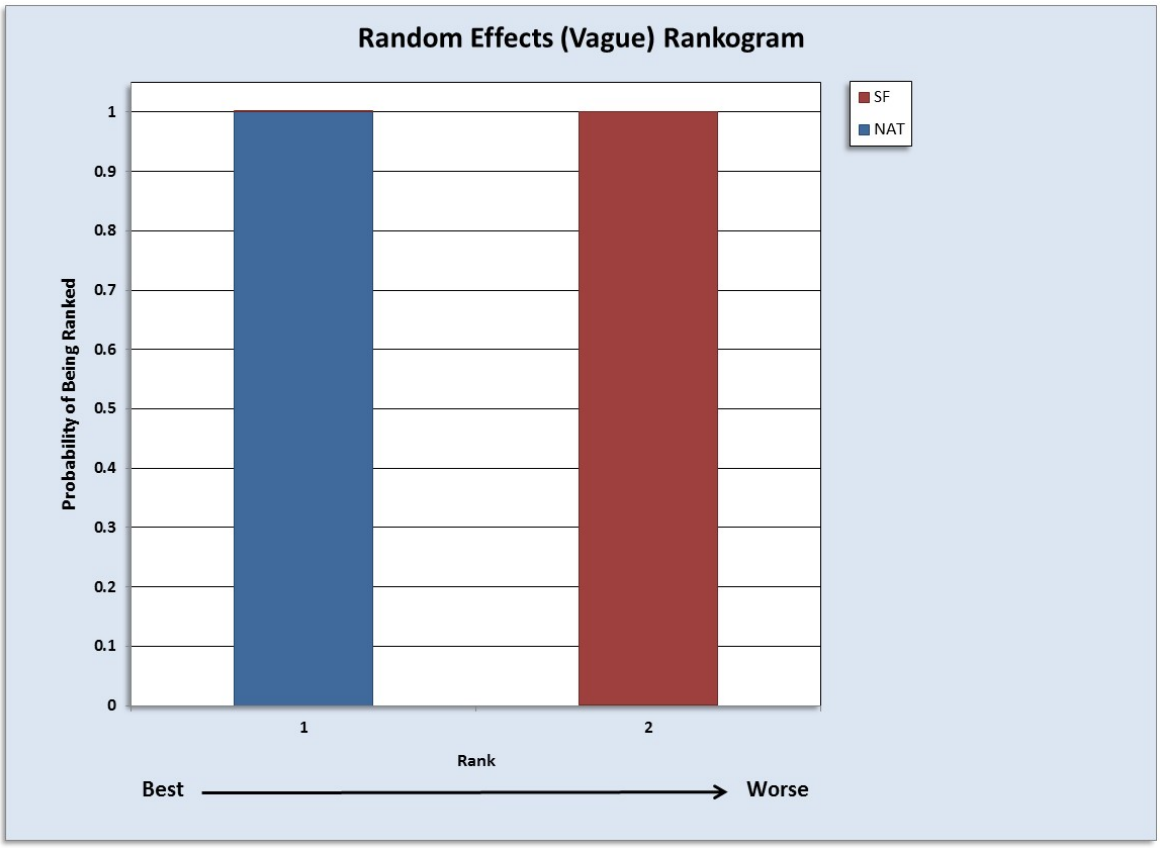


Figure Kxv: League table based on results of fixed effects and random effects (vague prior) models. Where odds ratio (O.R.) is greater than 1 treatment at top left is superior

NAT	
2.37 (2.19 – 2.55)	SF

Figure Kxvi: Rankogram summarising surface under the cumulative ranking (SUCRA).

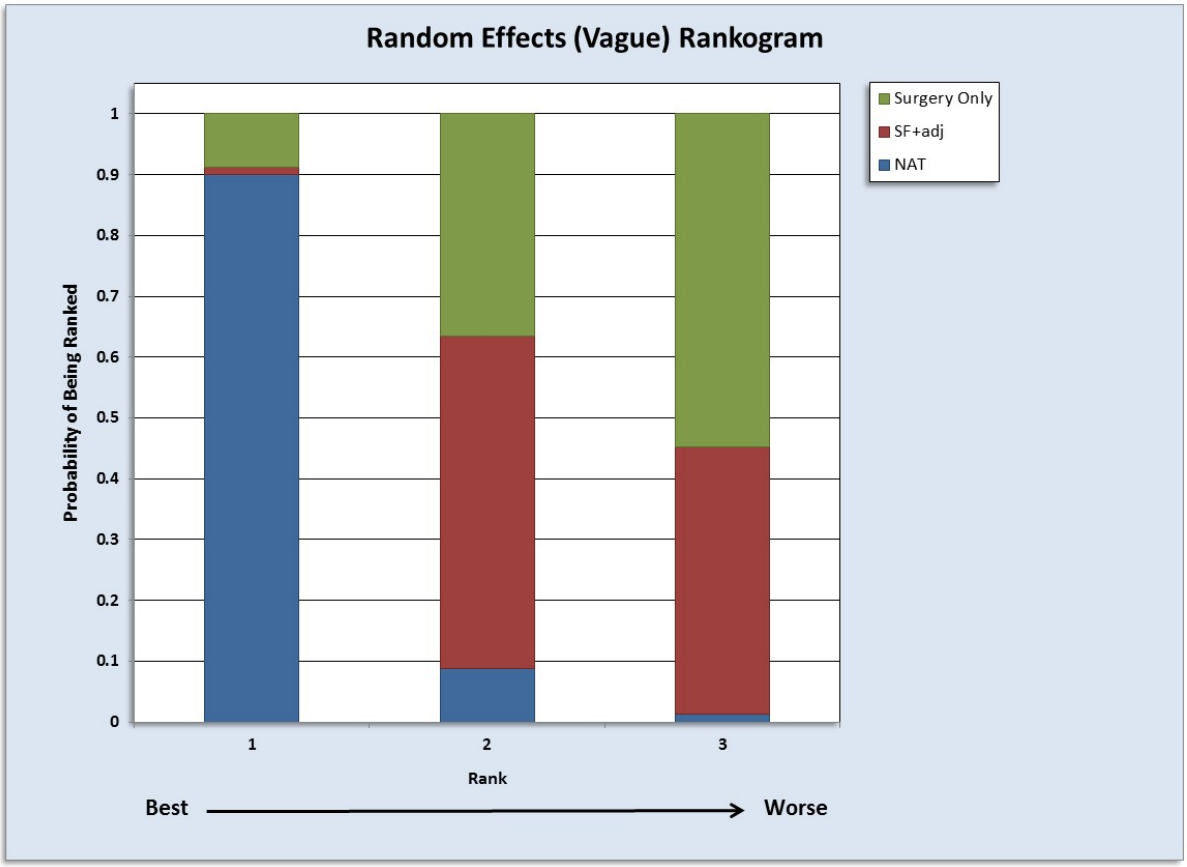


Sensitivity Network: 1-year Survival: Phase II/III plus RCTs

Figure Kxvii: League table based on results of fixed effects and random effects (vague prior) models. Where odds ratio (O.R.) is greater than 1 treatment at top left is superior

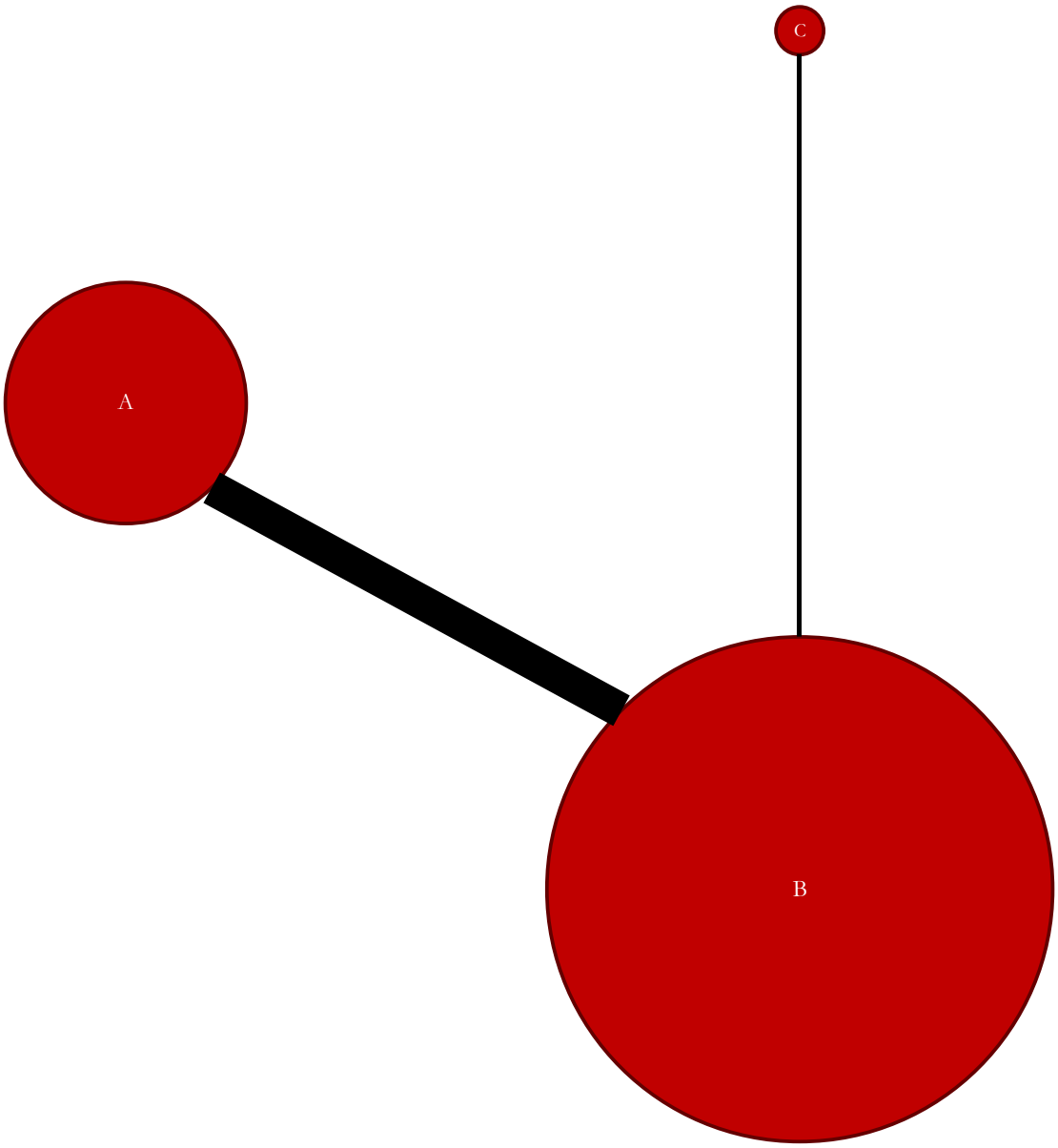
NAT		
2.23 (1.03 – 5.73)	**SF+adj**	
2.36 (0.55 – 12.40)	1.05 (0.29 – 3.90)	**Surgery Only**

Figure Kxviii: Rankogram summarising surface under the cumulative ranking (SUCRA).



Sensitivity Network: 1-year Survival: Phase II/III plus RCTs plus cohort studies

Figure Kxix: Bayesian Network Meta-analysis of Neoadjuvant therapy versus Upfront surgery plus adjuvant therapy versus surgery only



Drug	Abbreviation
NAT	A
SF+adj	B
Surgery Only	C

Figure Kxx: Results of fixed effects and random effects (vague prior) models

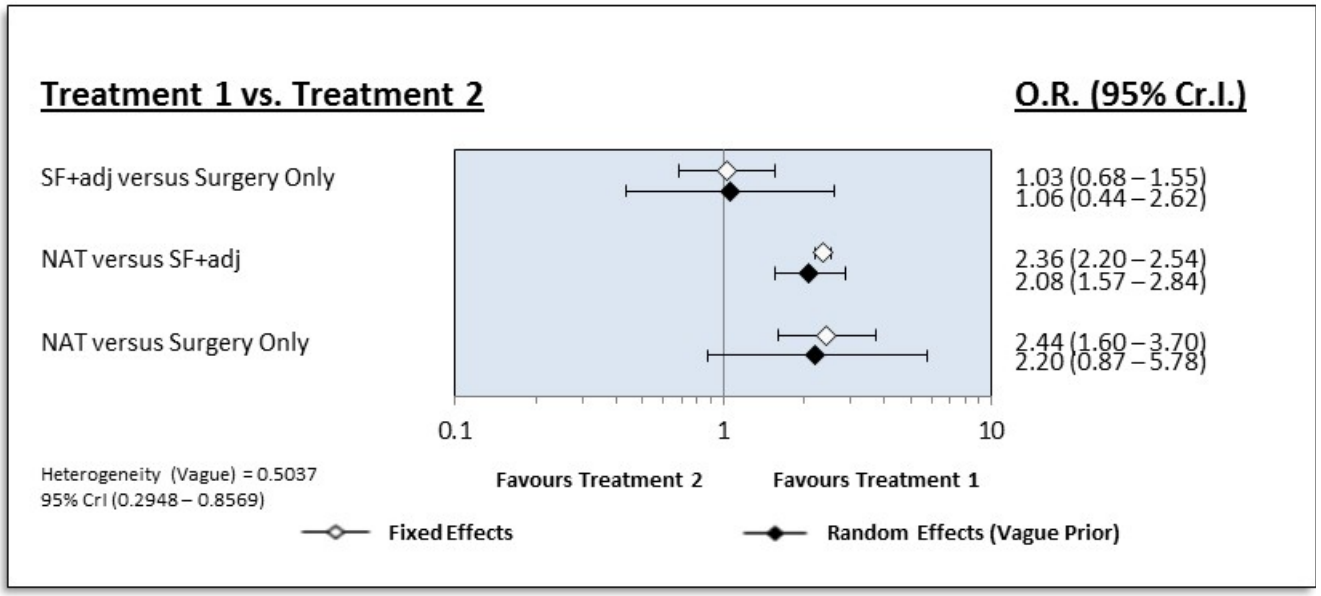
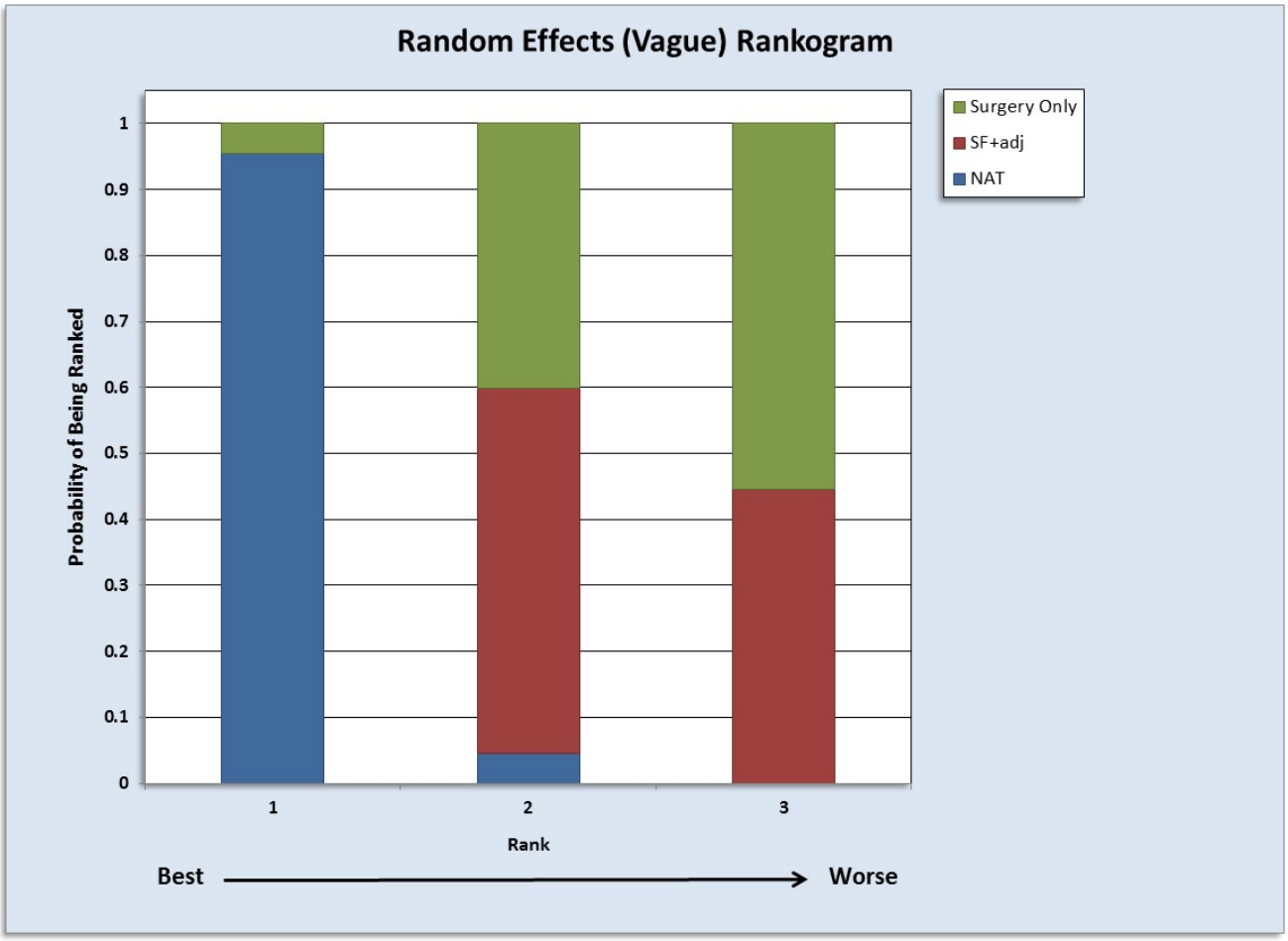


Figure Kxxi: League table based on results of fixed effects and random effects (vague prior) models. Where odds ratio (O.R.) is greater than 1 treatment at top left is superior

NAT		
2.08 (1.57 – 2.84)	**SF+adj**	
2.20 (0.87 – 5.78)	1.06 (0.44 – 2.62)	**Surgery Only**

Figure Kxxii: Rankogram summarising surface under the cumulative ranking (SUCRA).

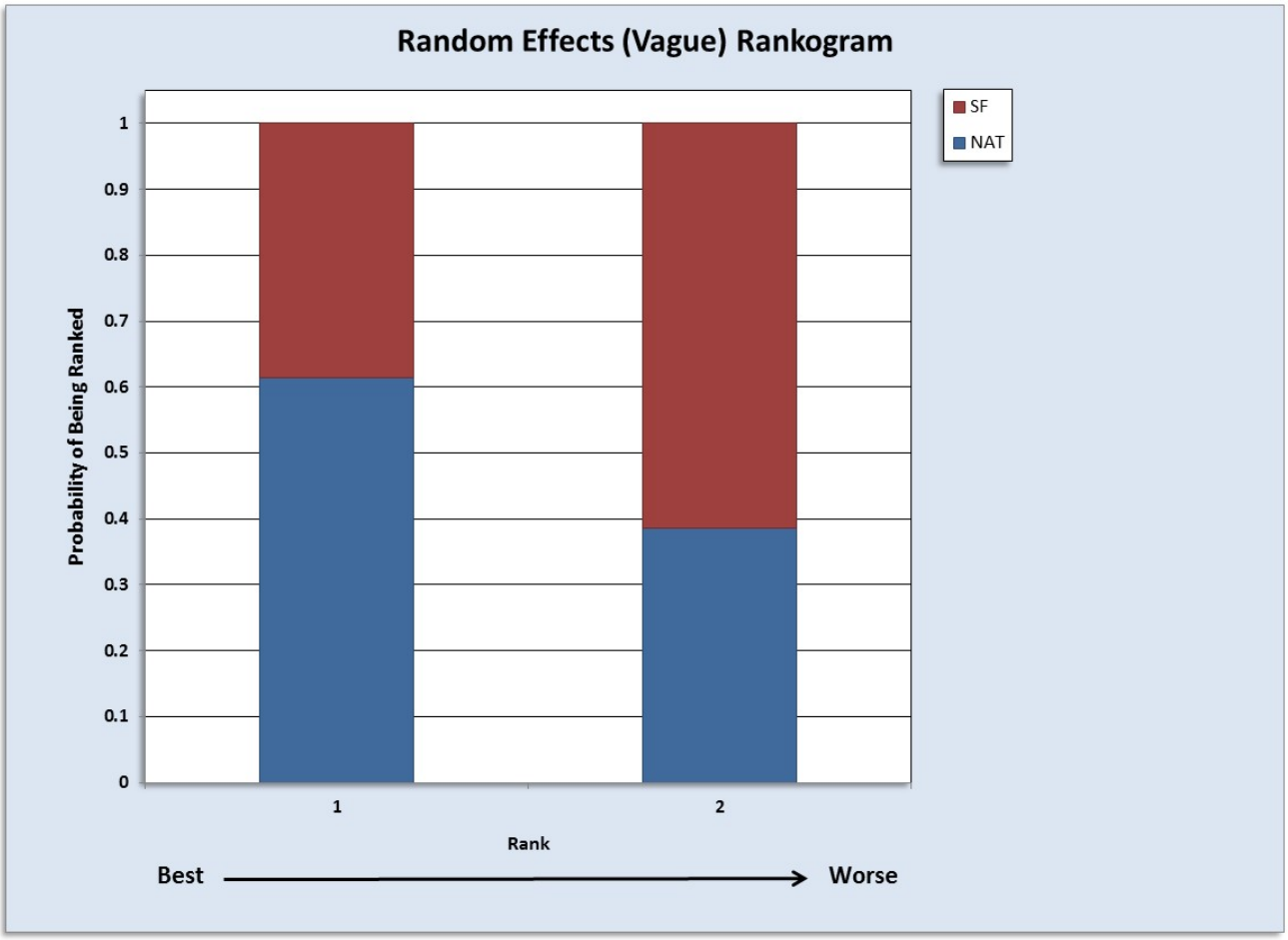


2-year Survival: Phase II/III Studies

Figure Kxxiii: League table based on results of fixed effects and random effects (vague prior) models. Where odds ratio (O.R.) is greater than 1 treatment at top left is superior

NAT	
1.32 (0.87 – 1.97)	SF

Figure Kxxiv: Rankogram summarising surface under the cumulative ranking (SUCRA).



2-year Survival: Phase II/III Studies plus Cohort Studies

Figure Kxxv: Results of fixed effects and random effects (vague prior) models

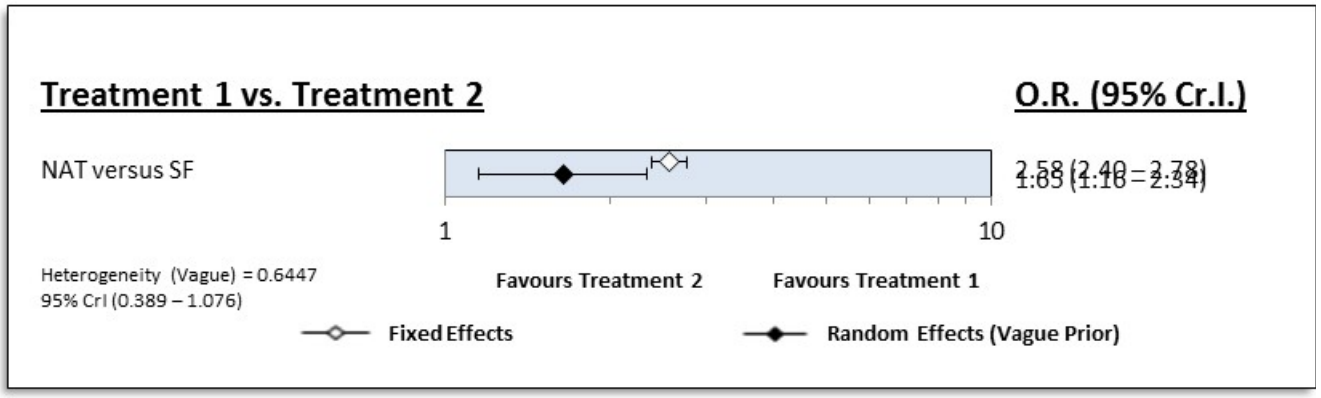
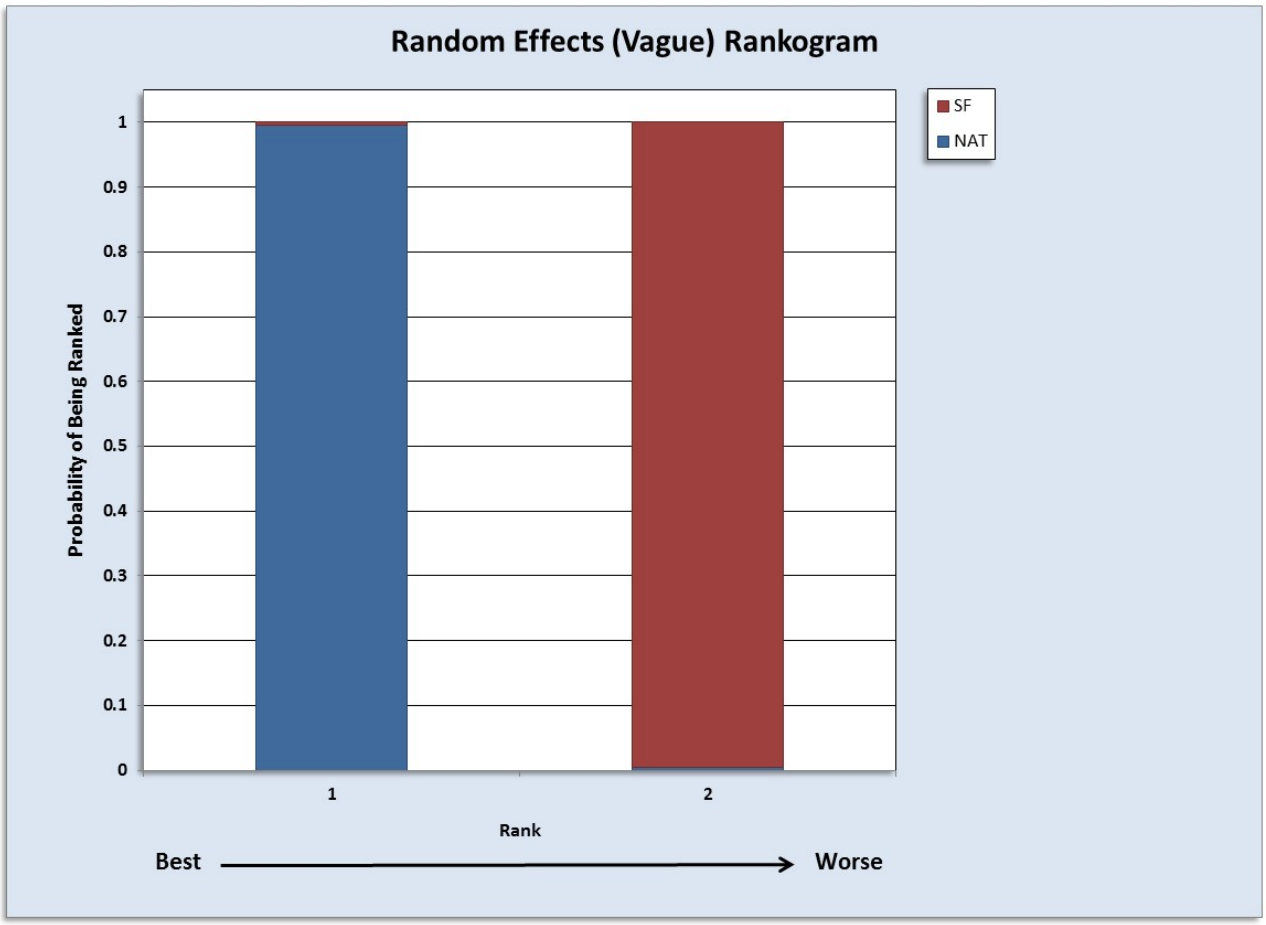


Figure Kxxvi: League table based on results of fixed effects and random effects (vague prior) models. Where odds ratio (O.R.) is greater than 1 treatment at top left is superior

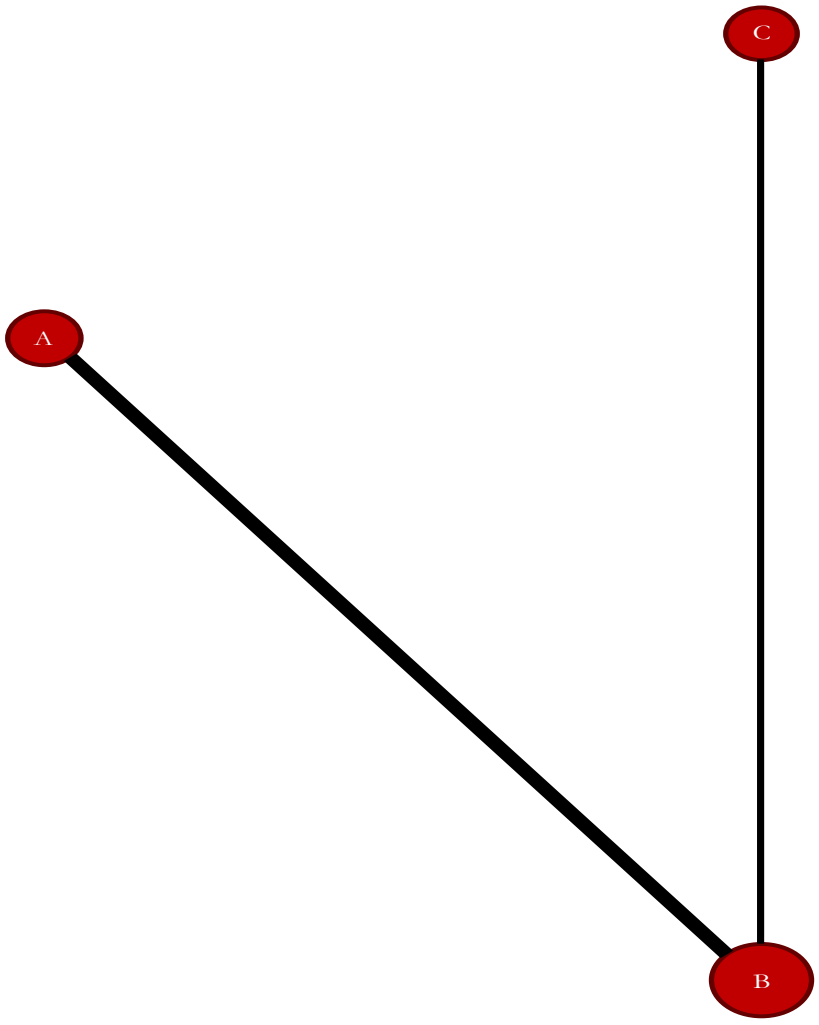
NAT	
2.58 (2.40 – 2.78)	**SF**

Figure Kxxvii: Rankogram summarising surface under the cumulative ranking (SUCRA).



Sensitivity Network: 2-year Survival: Phase II/III plus RCTs

Figure Kxxviii: Bayesian Network Meta-analysis of Neoadjuvant therapy versus Upfront surgery plus adjuvant therapy versus surgery only



Drug	Abbreviation
NAT	A
SF+adj	B
Surgery Only	C

Figure Kxxix: Results of fixed effects and random effects (vague prior) models

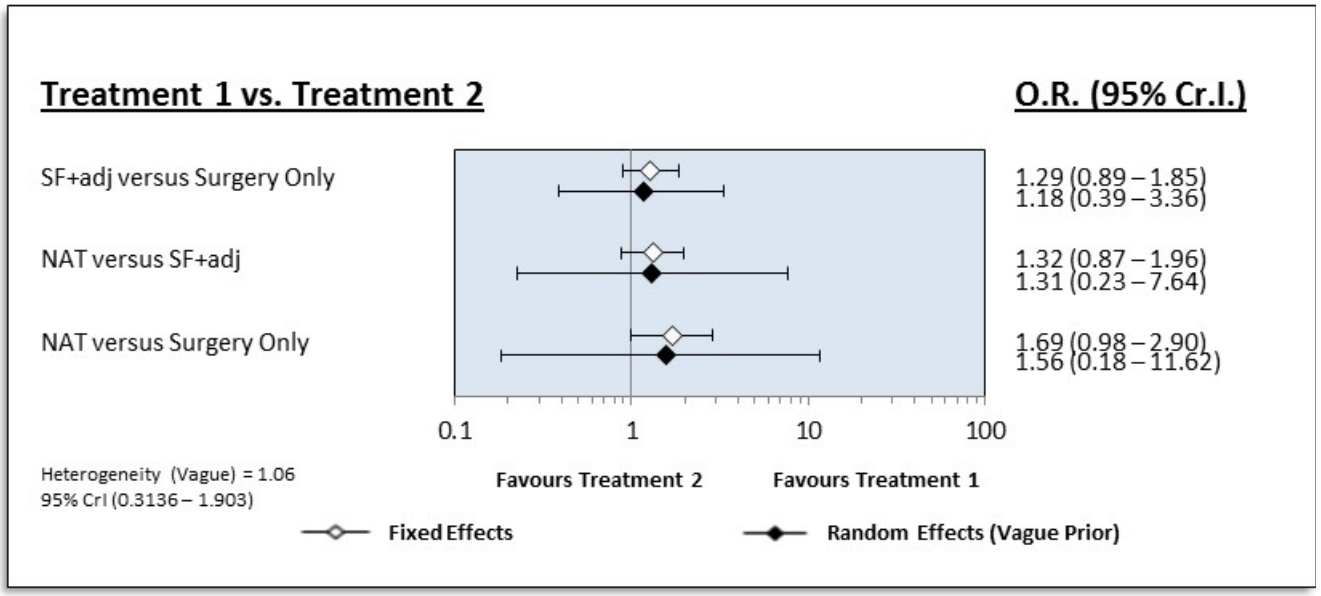
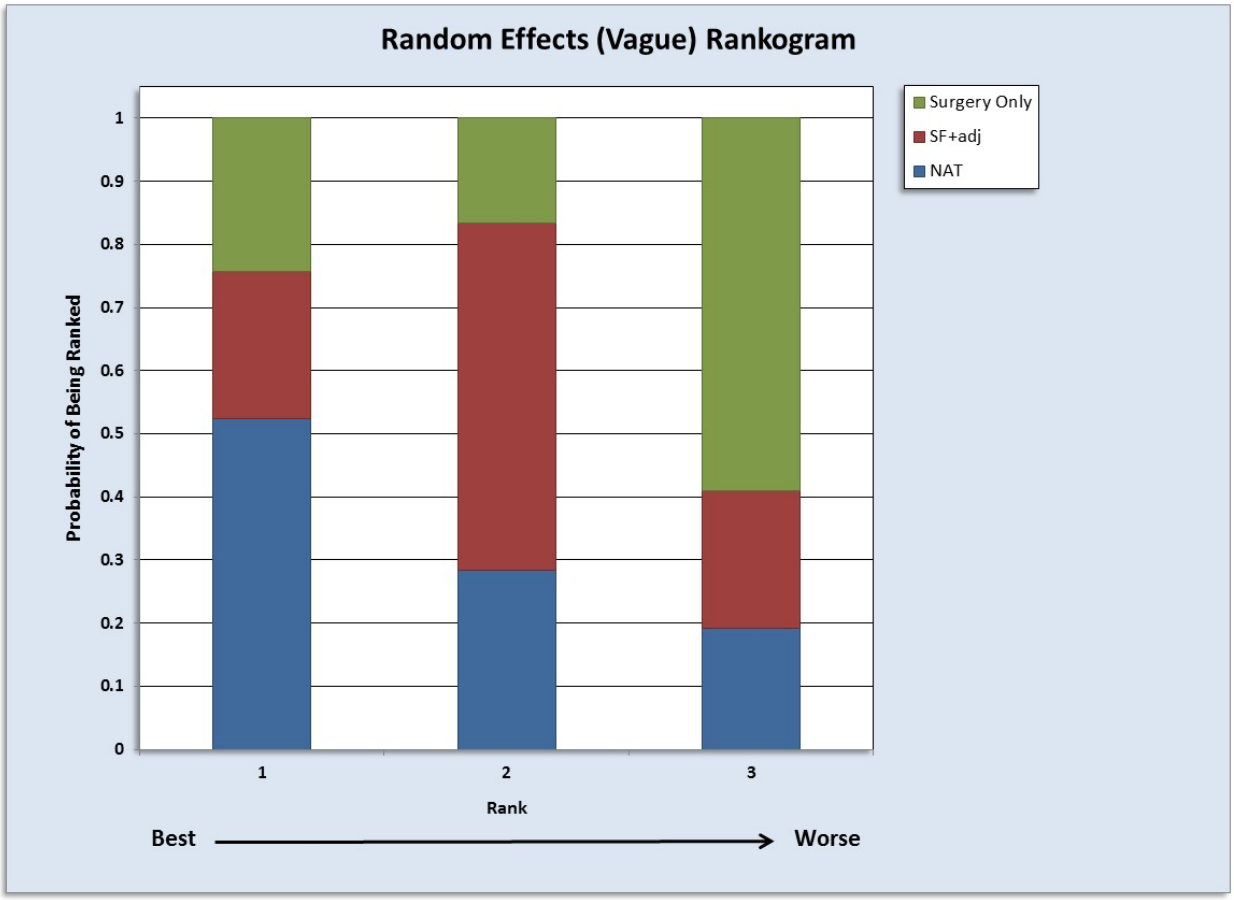


Figure Kxxx: League table based on results of fixed effects and random effects (vague prior) models. Where odds ratio (O.R.) is greater than 1 treatment at top left is superior

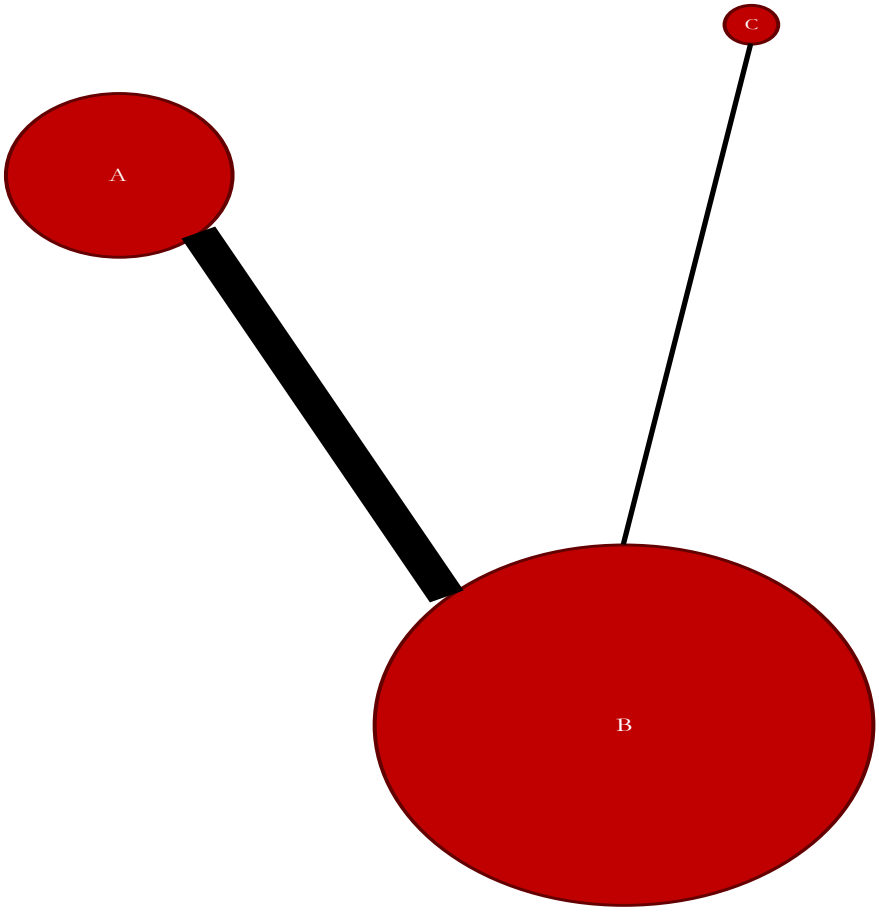
NAT		
1.32 (0.87 – 1.96)	**SF+adj**	
1.69 (0.98 – 2.90)	1.29 (0.89 – 1.85)	**Surgery Only**

Figure Kxxxi: Rankogram summarising surface under the cumulative ranking (SUCRA).



Sensitivity Network: 2-year Survival: Phase II/III plus RCTs plus cohort studies

Figure Kxxxii: Bayesian Network Meta-analysis of Neoadjuvant therapy versus Upfront surgery plus adjuvant therapy versus surgery only

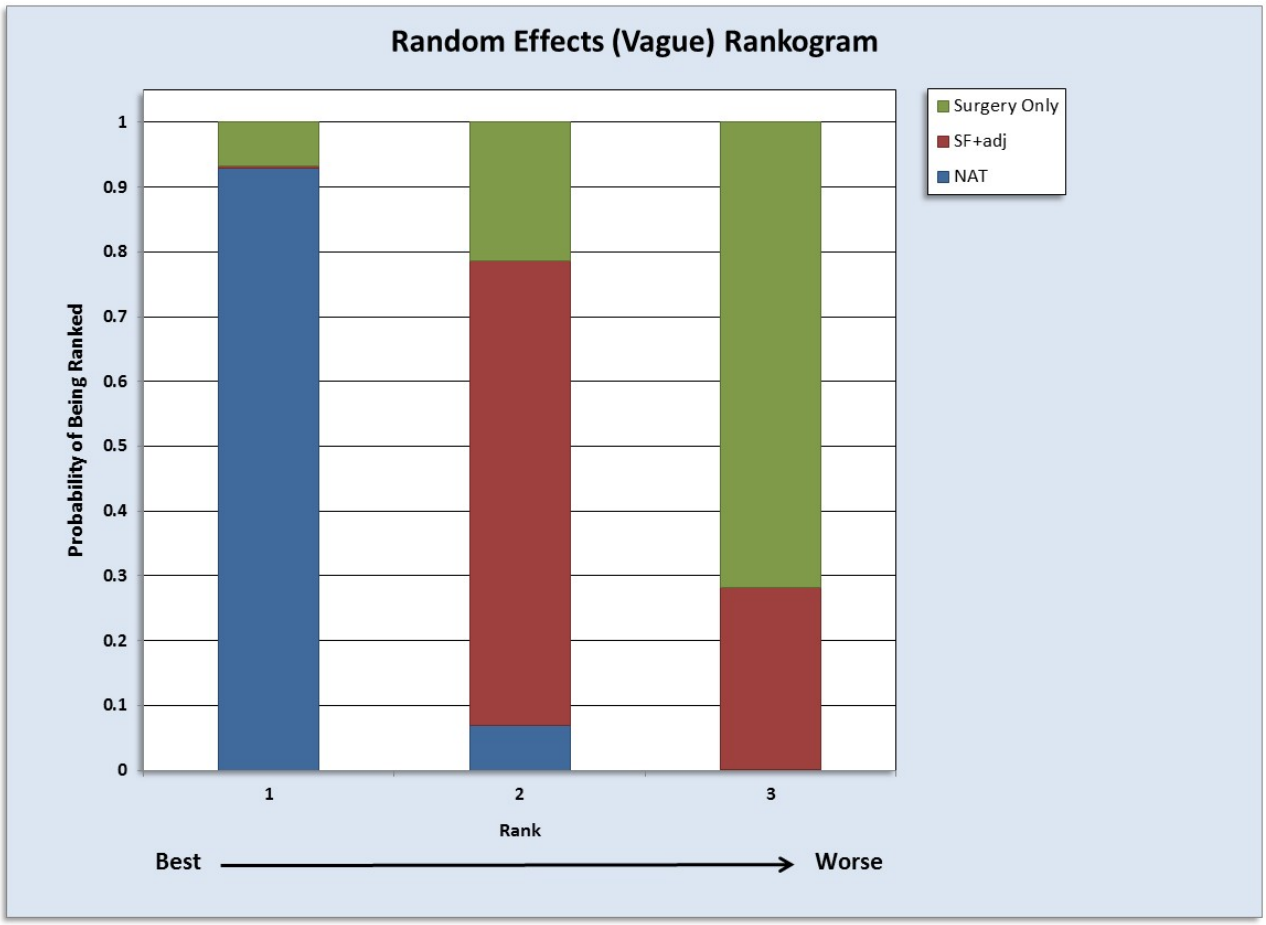


Drug	Abbreviation
NAT	A
SF+adj	B
Surgery Only	C

Figure Kxxxiii: League table based on results of fixed effects and random effects (vague prior) models. Where odds ratio (O.R.) is greater than 1 treatment at top left is superior

NAT		
2.58 (2.40 – 2.78)	**SF+adj**	
3.33 (2.30 – 4.83)	1.29 (0.90 – 1.85)	**Surgery Only**

Figure Kxxxivl: Rankogram summarising surface under the cumulative ranking (SUCRA).



3-year Survival: Phase II/III Studies

Figure Kxxxv: Results of fixed effects and random effects (vague prior) models

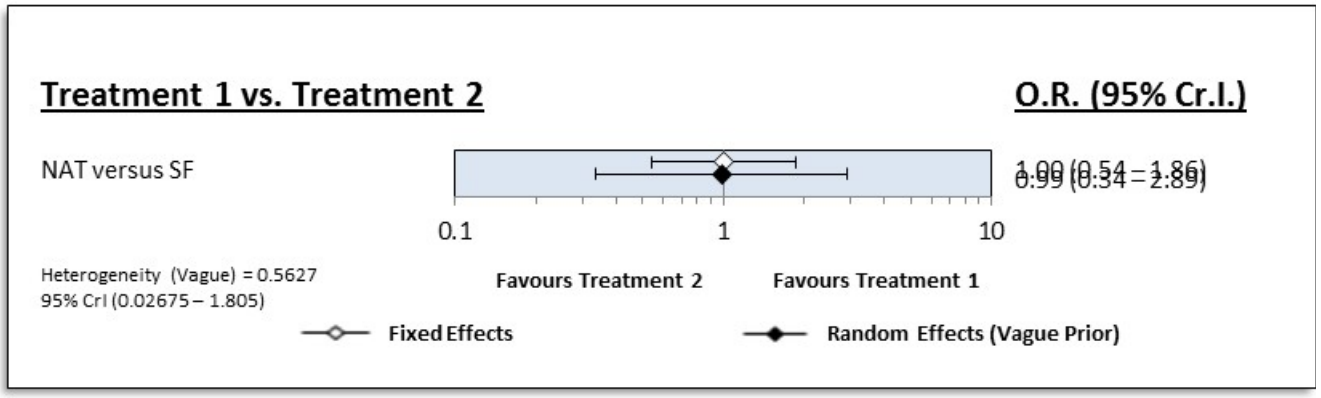
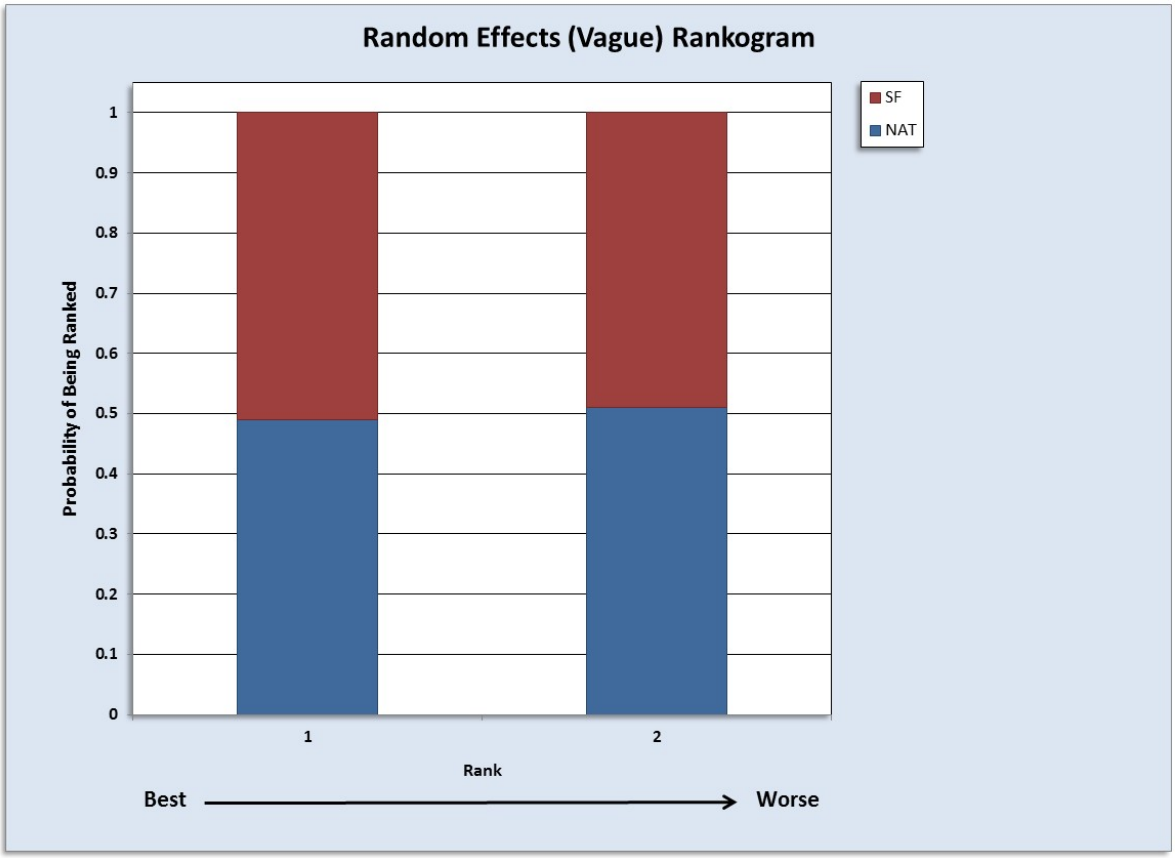


Figure Kxxxvi: League table based on results of fixed effects and random effects (vague prior) models. Where odds ratio (O.R.) is greater than 1 treatment at top left is superior

NAT	
1.00 (0.54 – 1.86)	SF

Figure Kxxxvii: Rankogram summarising surface under the cumulative ranking (SUCRA).



3-year Survival: Phase II/III Studies plus Cohort Studies

Figure Kxxxviii: Results of fixed effects and random effects (vague prior) models

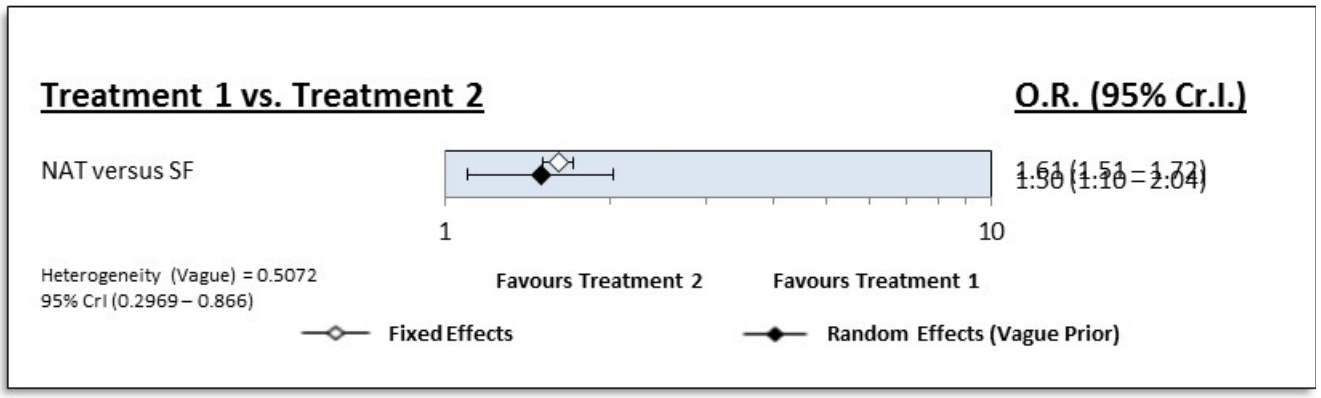
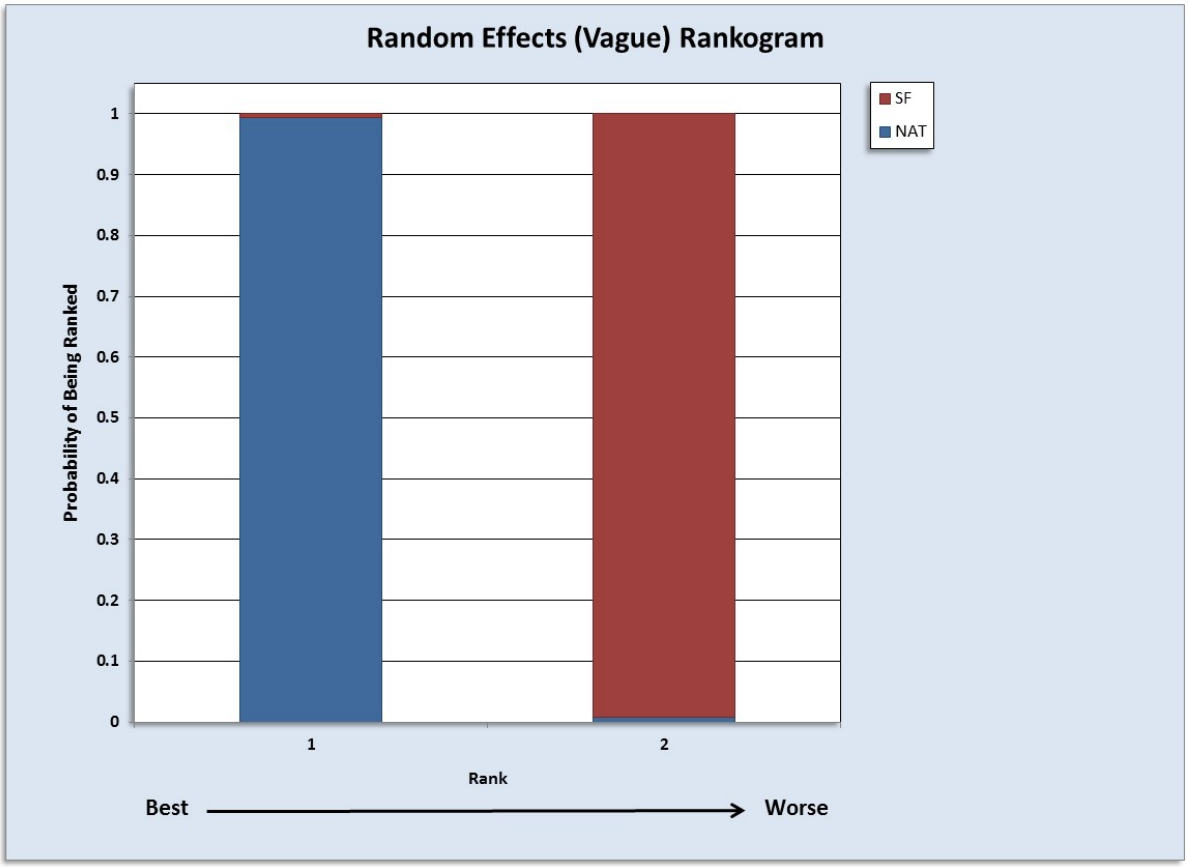


Figure Kxxxix: League table based on results of fixed effects and random effects (vague prior) models. Where odds ratio (O.R.) is greater than 1 treatment at top left is superior

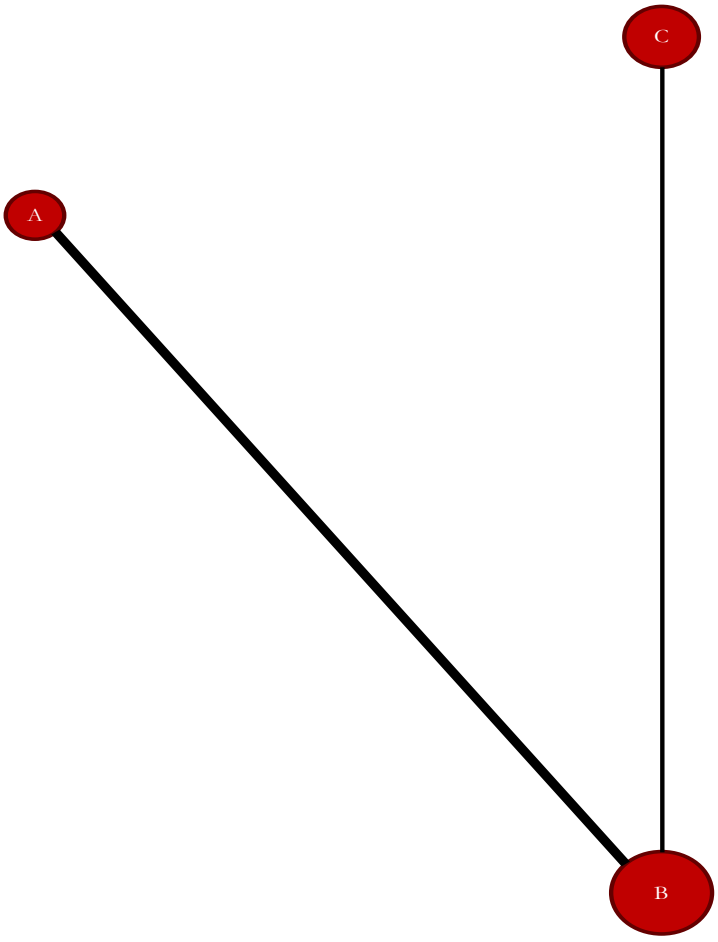
NAT	
1.61 (1.51 – 1.72)	**SF**

Figure Kxl: Rankogram summarising surface under the cumulative ranking (SUCRA).



Sensitivity Network: 3-year Survival: Phase II/III plus RCTs

Figure Kxli: Bayesian Network Meta-analysis of Neoadjuvant therapy versus Upfront surgery plus adjuvant therapy versus surgery only



Drug	Abbreviation
NAT	A
SF+adj	B
Surgery Only	C

Figure Kxlii: Results of fixed effects and random effects (vague prior) models

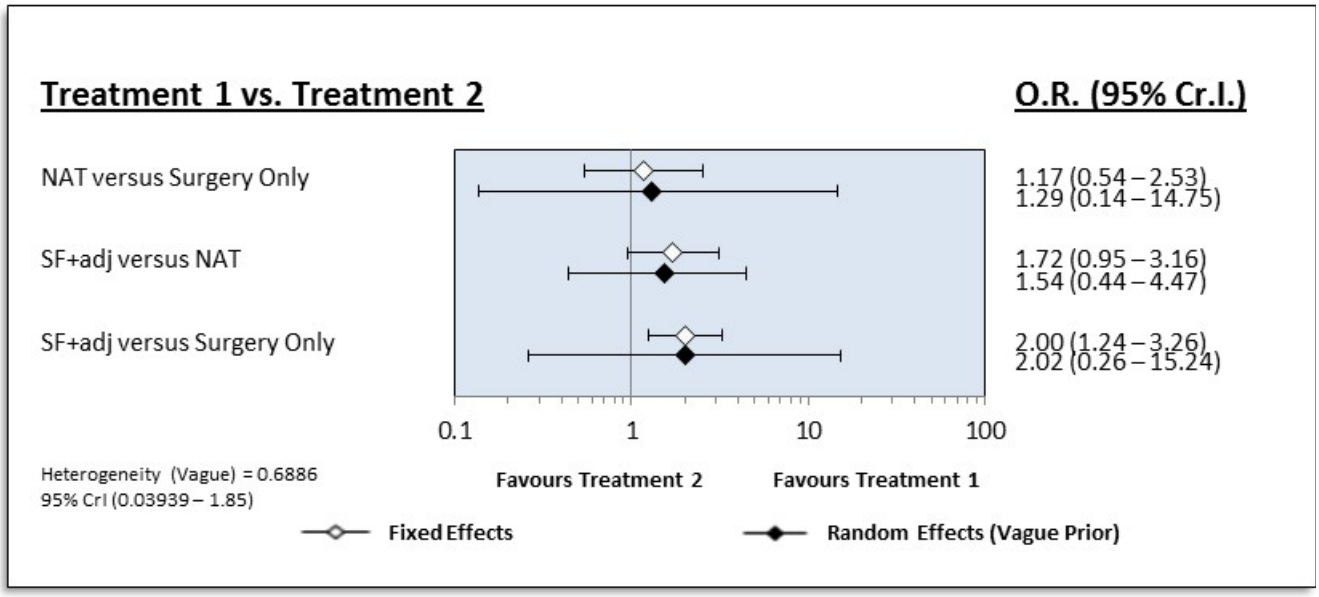
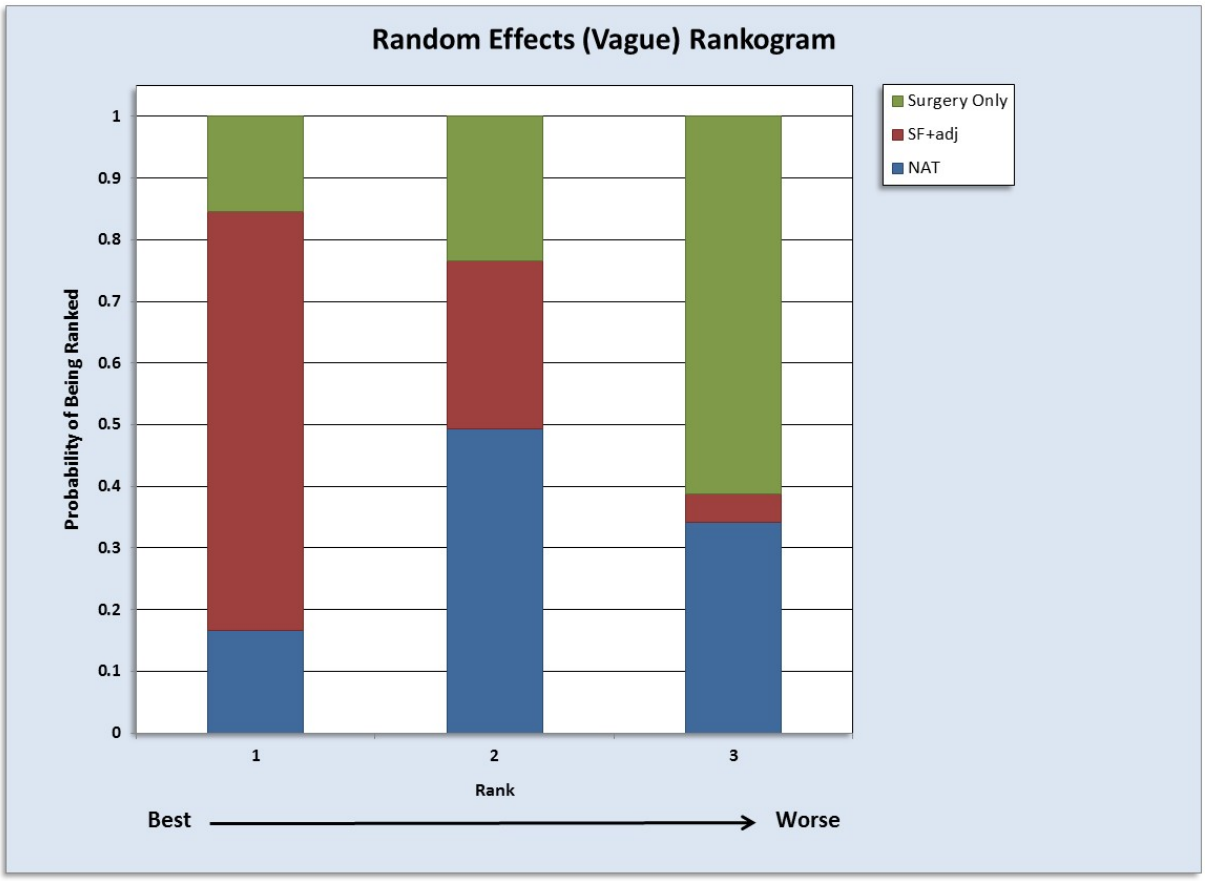


Figure Kxliii: League table based on results of fixed effects and random effects (vague prior) models. Where odds ratio (O.R.) is greater than 1 treatment at top left is superior

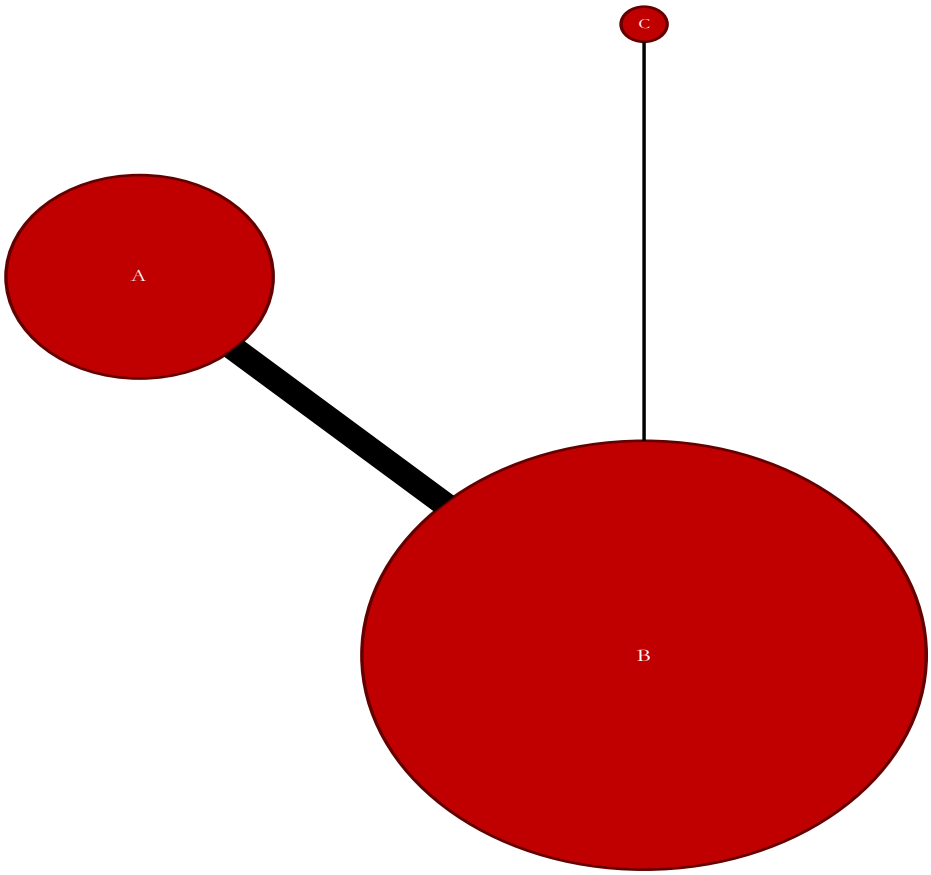
SF+adj		
1.72 (0.95 – 3.16)	**NAT**	
2.00 (1.24 – 3.26)	1.17 (0.54 – 2.53)	**Surgery Only**

Figure Kxliv: Rankogram summarising surface under the cumulative ranking (SUCRA).



Sensitivity Network: 3-year Survival: Phase II/III plus RCTs plus cohort studies

Figure Kxlv: Bayesian Network Meta-analysis of Neoadjuvant therapy versus Upfront surgery plus adjuvant therapy versus surgery only

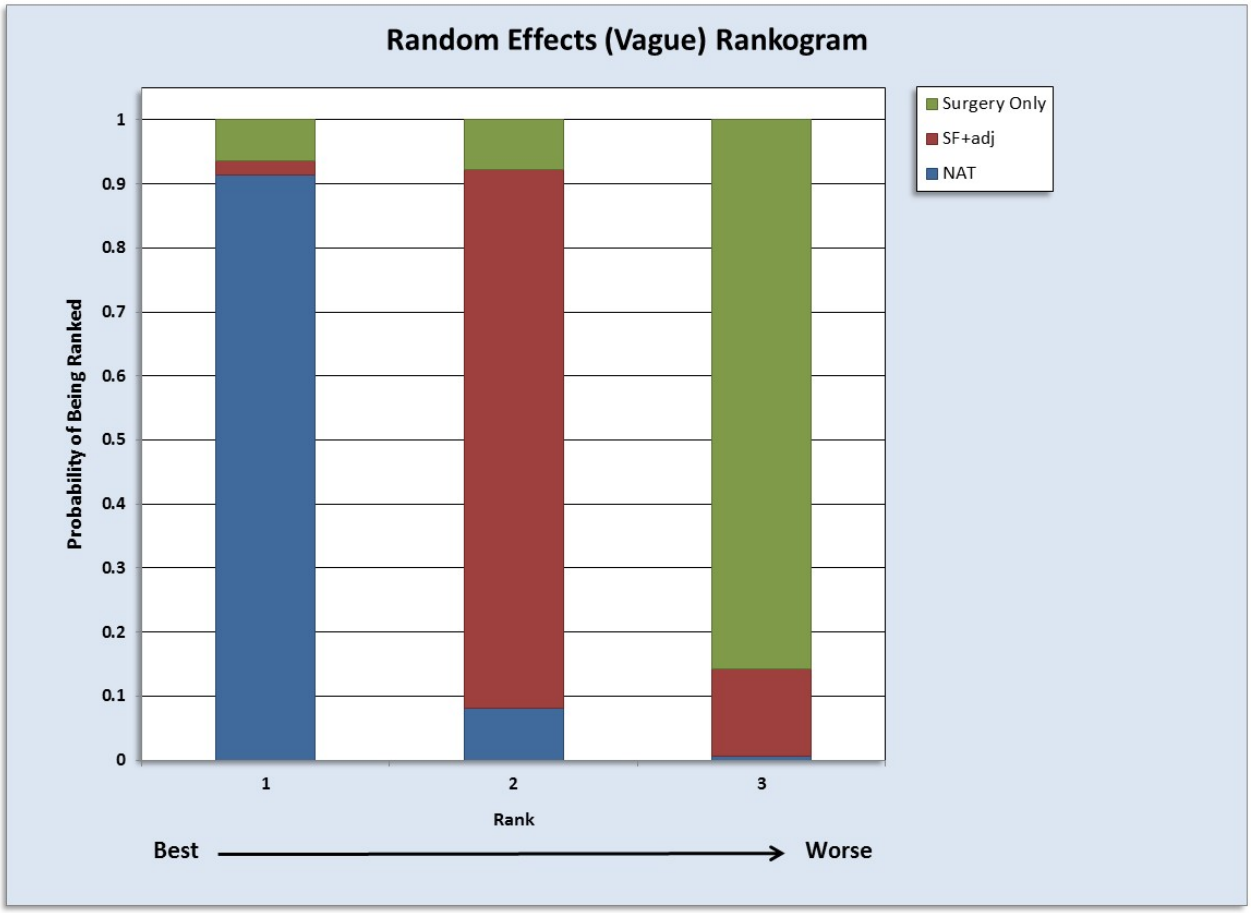


Drug	Abbreviation
NAT	A
SF+adj	B
Surgery Only	C

Figure Kxlvi: League table based on results of fixed effects and random effects (vague prior) models. Where odds ratio (O.R.) is greater than 1 treatment at top left is superior

NAT		
1.60 (1.49 – 1.71)	**SF+adj**	
3.21 (1.98 – 5.27)	2.01 (1.25 – 3.28)	**Surgery Only**

Figure Kxlvii: Rankogram summarising surface under the cumulative ranking (SUCRA).



4-year Survival: Phase II/III studies

Figure Kxlviii: Results of fixed effects and random effects (vague prior) models

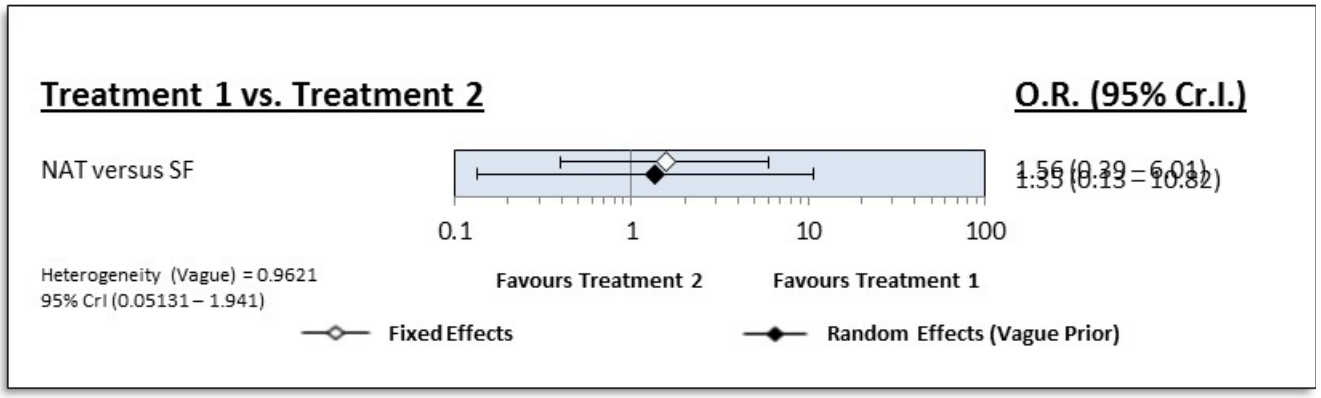
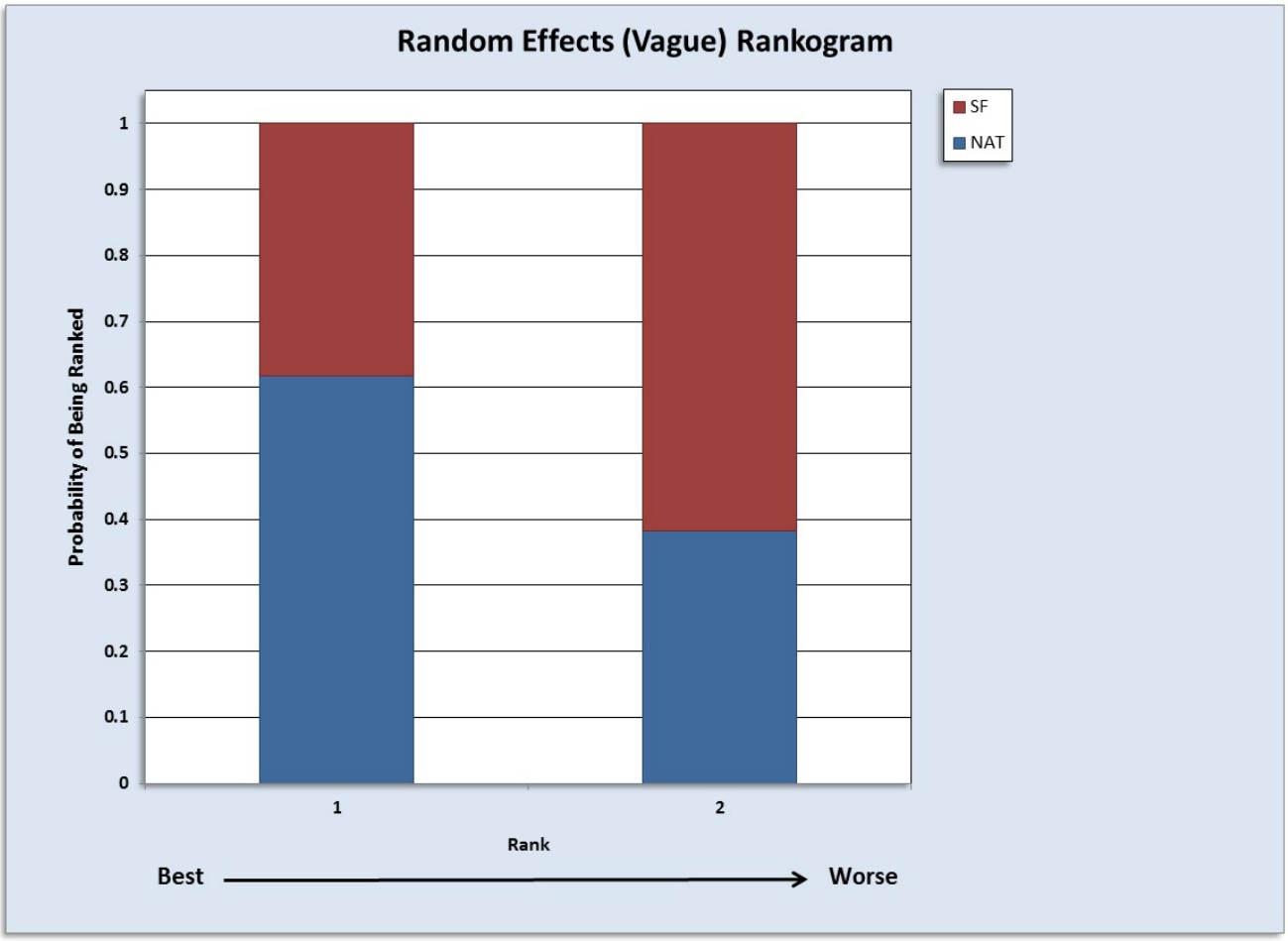


Figure Kxlix: League table based on results of fixed effects and random effects (vague prior) models. Where odds ratio (O.R.) is greater than 1 treatment at top left is superior

NAT	
1.56 (0.39 – 6.01)	**SF**

Figure Kl: Rankogram summarising surface under the cumulative ranking (SUCRA).



4-year Survival: Phase II/III plus cohort studies

Figure Kli: Results of fixed effects and random effects (vague prior) models

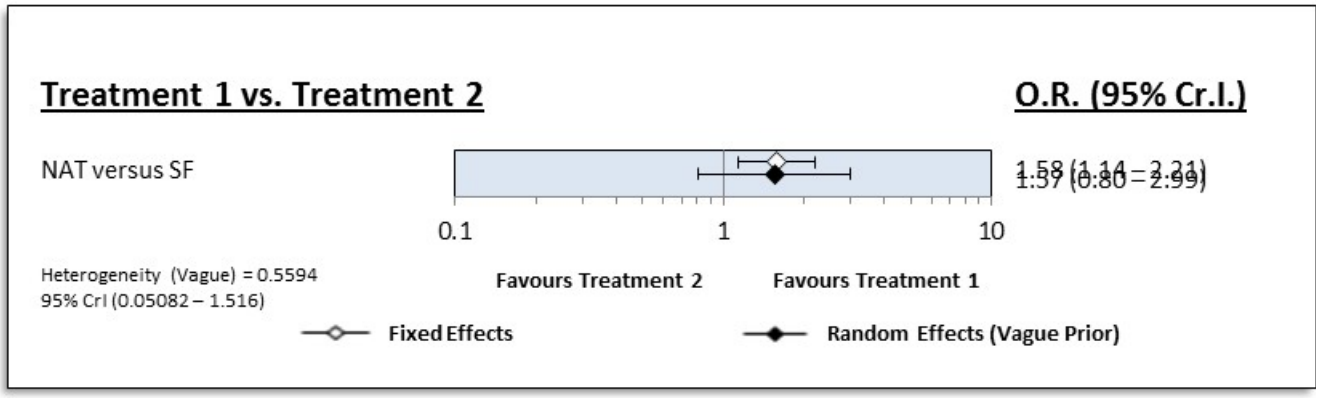
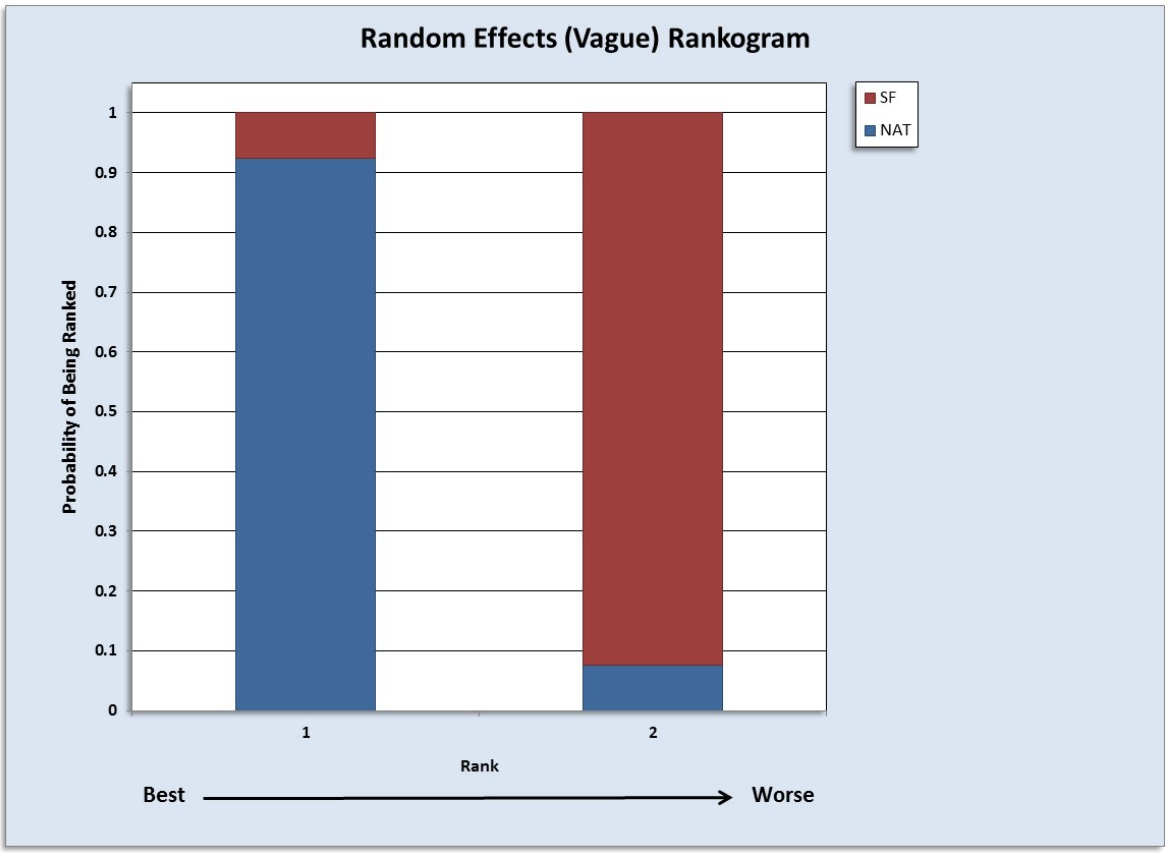


Figure Klii: League table based on results of fixed effects and random effects (vague prior) models. Where odds ratio (O.R.) is greater than 1 treatment at top left is superior

NAT	
1.58 (1.14 – 2.21)	SF

Figure Kliii: Rankogram summarising surface under the cumulative ranking (SUCRA).



5-year Survival: Phase II/III studies

Figure Kliv: Results of fixed effects and random effects (vague prior) models

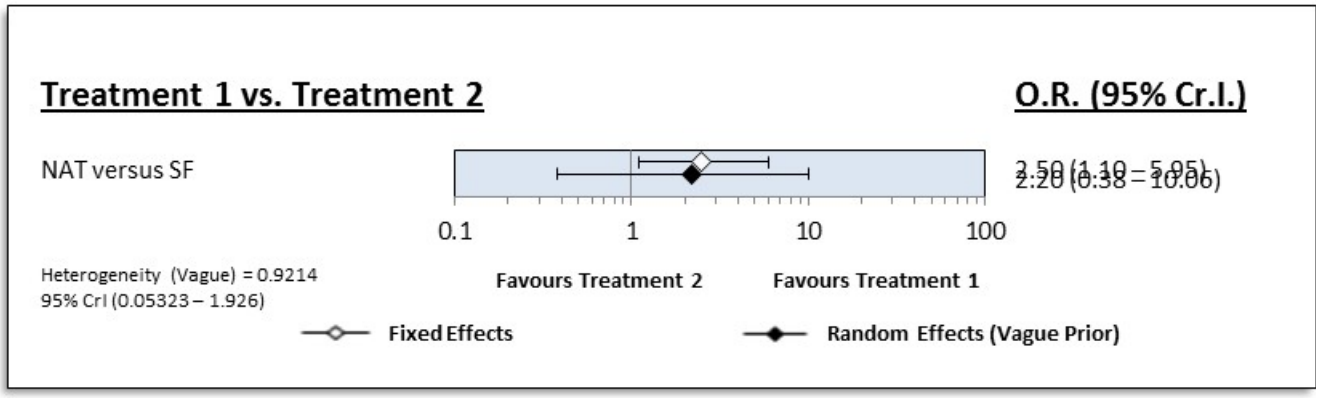
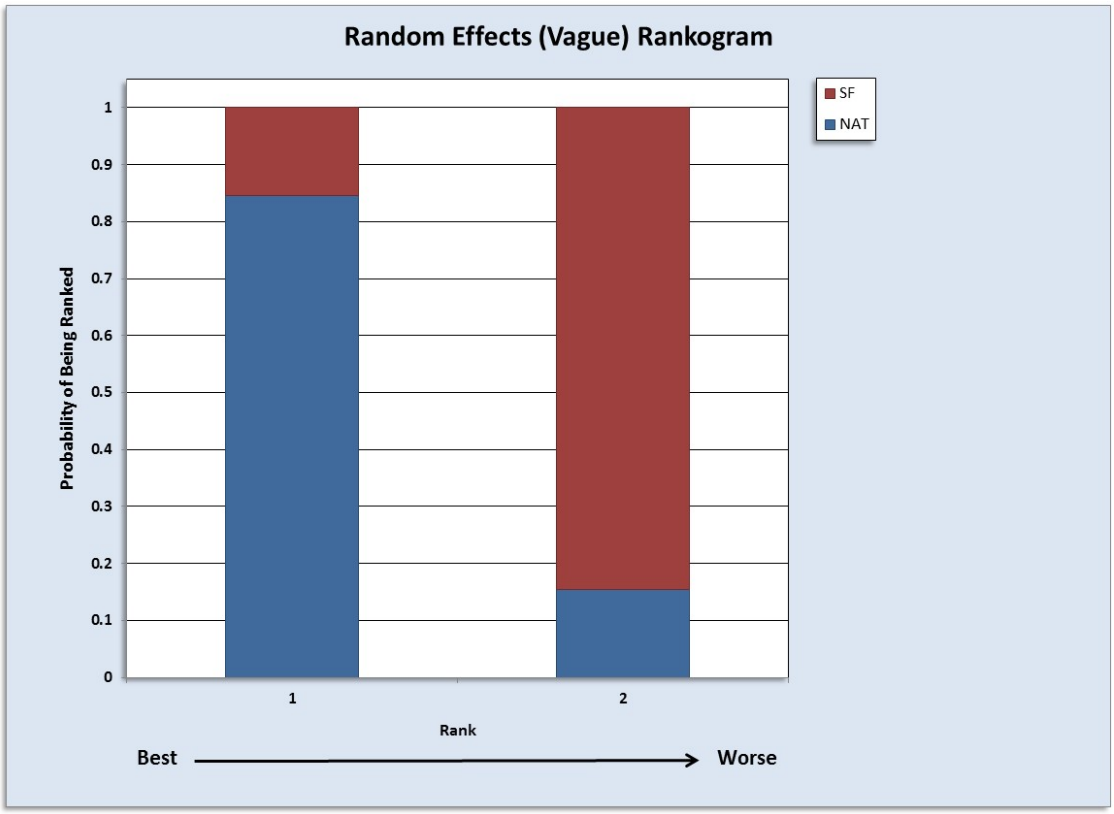


Figure Klv: League table based on results of fixed effects and random effects (vague prior) models. Where odds ratio (O.R.) is greater than 1 treatment at top left is superior

NAT	
2.50 (1.10 – 5.95)	**SF**

Figure Klvi: Rankogram summarising surface under the cumulative ranking (SUCRA).



5-year Survival: Phase II/III plus cohort studies

Figure Klvii: Results of fixed effects and random effects (vague prior) models

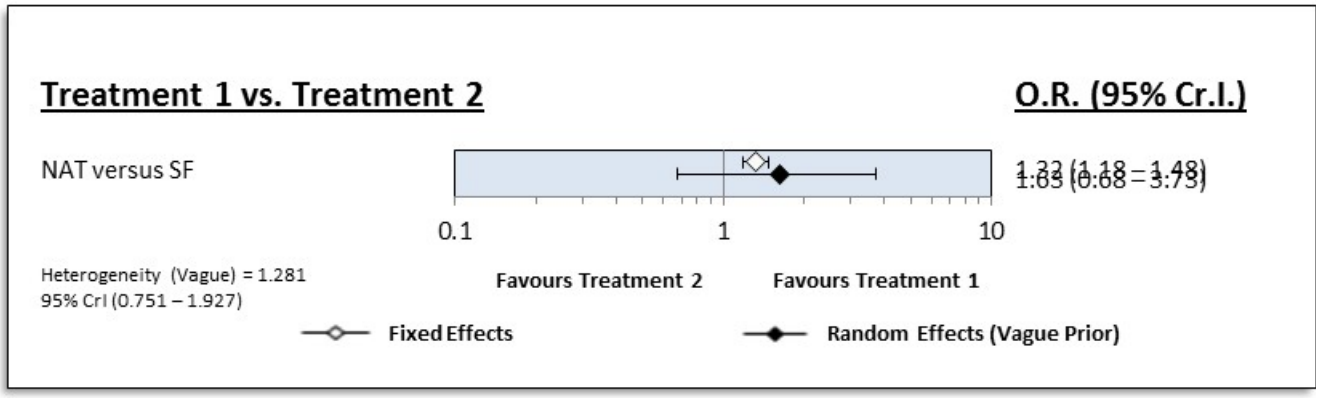
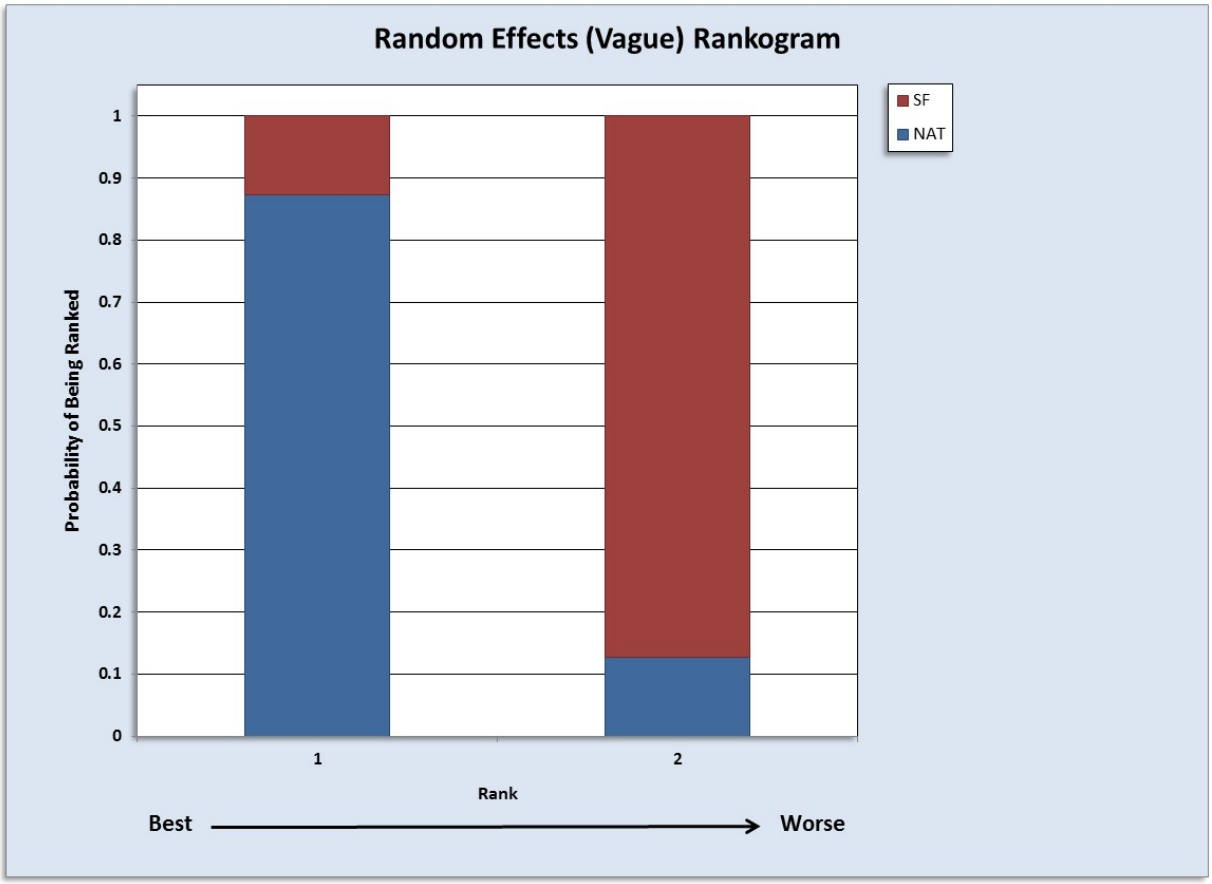


Figure Klviii: League table based on results of fixed effects and random effects (vague prior) models. Where odds ratio (O.R.) is greater than 1 treatment at top left is superior

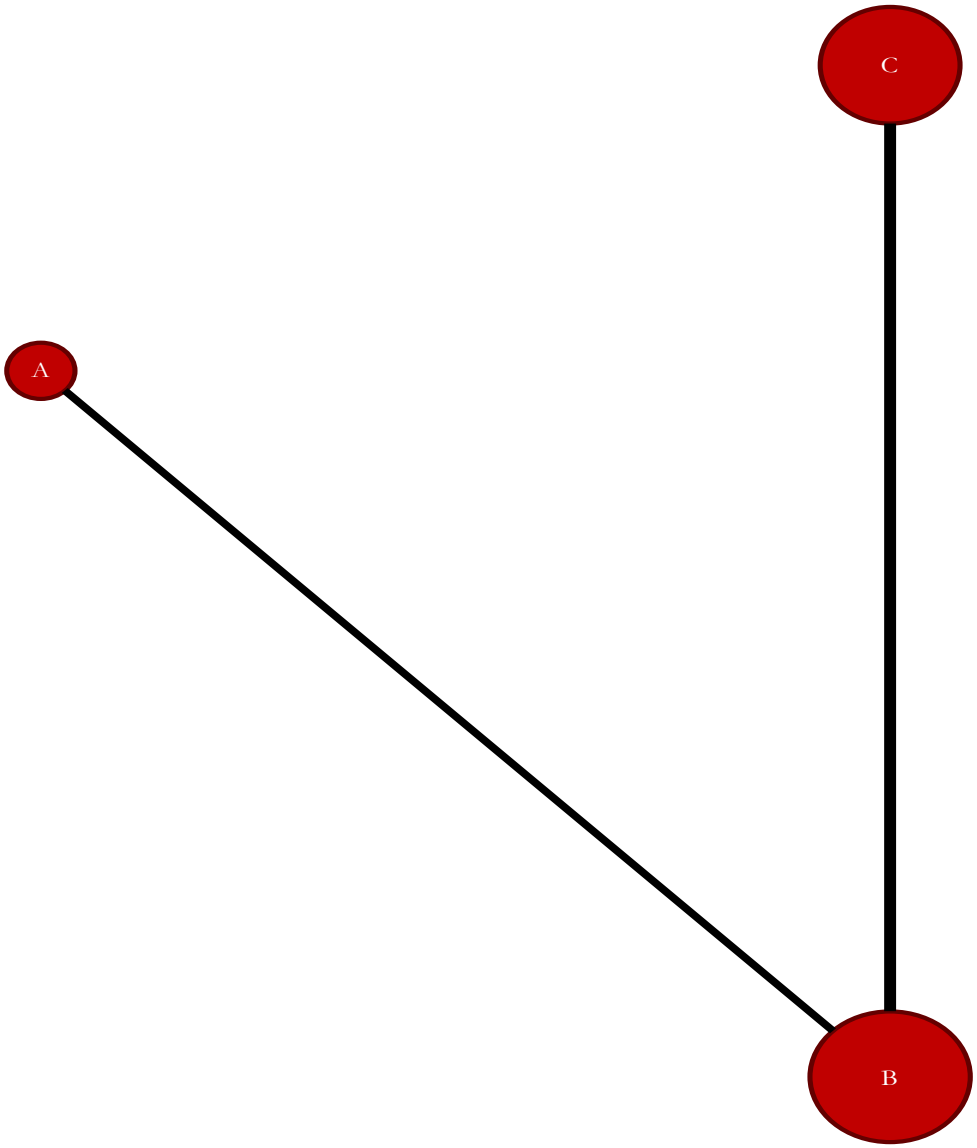
NAT	
1.32 (1.18 – 1.48)	SF

Figure Klix: Rankogram summarising surface under the cumulative ranking (SUCRA).



Sensitivity Network: 5-year Survival: Phase II/III plus RCTs

Figure Klx: Bayesian Network Meta-analysis of Neoadjuvant therapy versus Upfront surgery plus adjuvant therapy versus surgery only

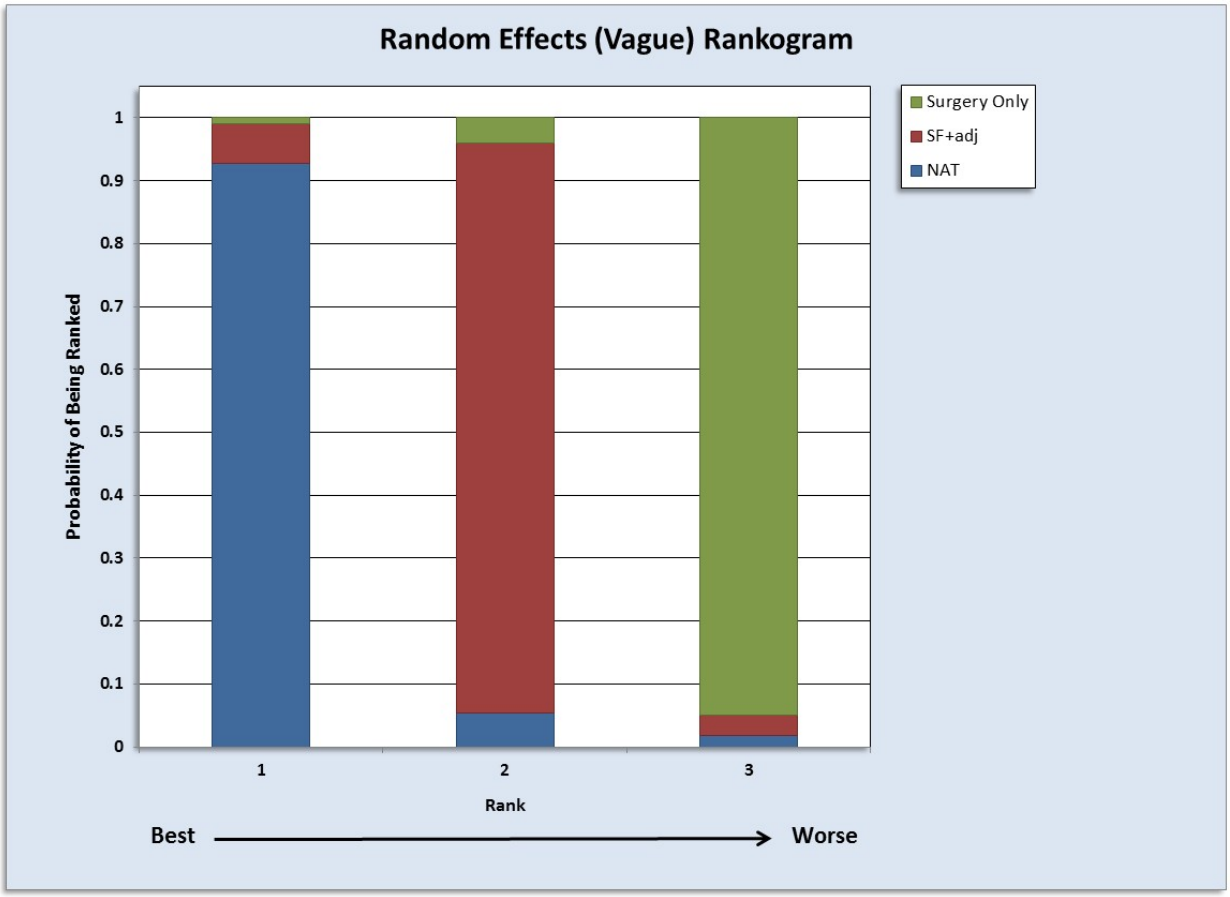


Drug	Abbreviation
NAT	A
SF+adj	B
Surgery Only	C

Figure Klxi: League table based on results of fixed effects and random effects (vague prior) models. Where odds ratio (O.R.) is greater than 1 treatment at top left is superior

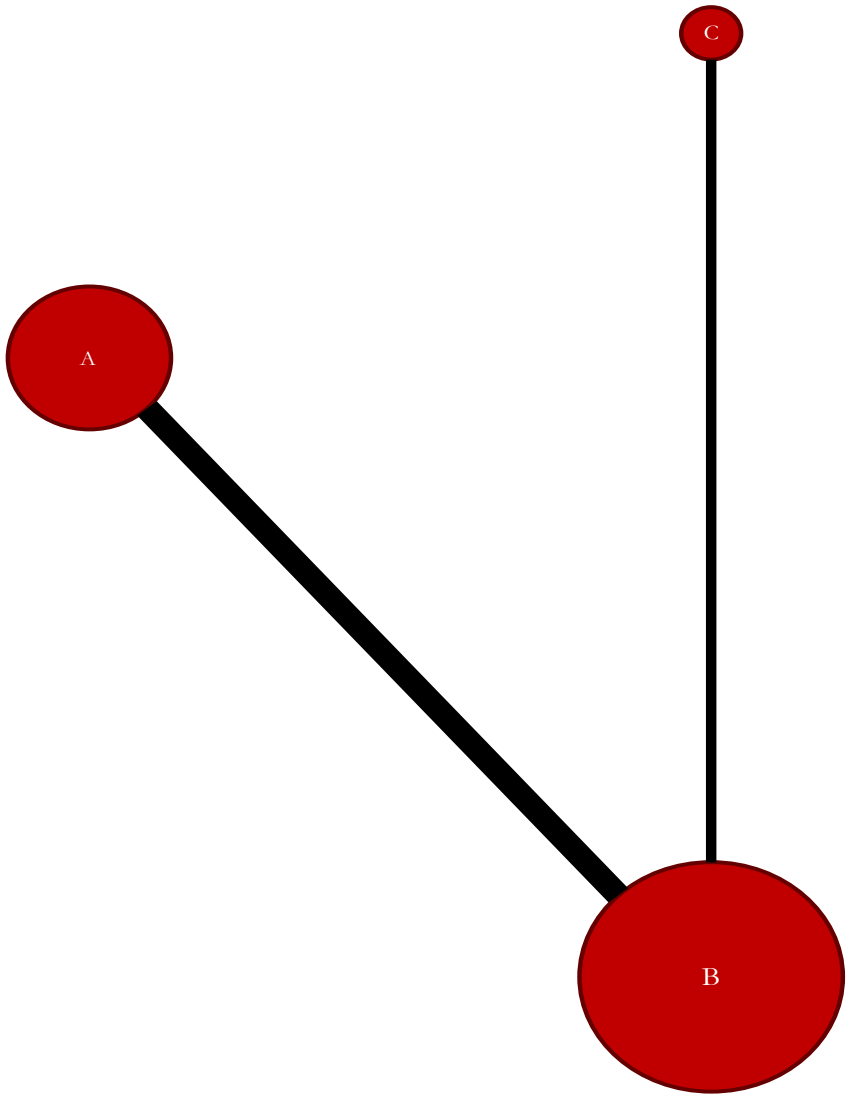
NAT		
2.49 (1.10 – 5.85)	**SF+adj**	
4.56 (1.86 – 11.54)	1.83 (1.26 – 2.67)	**Surgery Only**

Figure Klxii: Rankogram summarising surface under the cumulative ranking (SUCRA).



Sensitivity Network: 5-year Survival: Phase II/III plus RCTs plus cohort studies

Figure Klxiii: Bayesian Network Meta-analysis of Neoadjuvant therapy versus Upfront surgery plus adjuvant therapy versus surgery only



Drug	Abbreviation
NAT	A
SF+adj	B
Surgery Only	C

Figure Klxiv: Results of fixed effects and random effects (vague prior) models

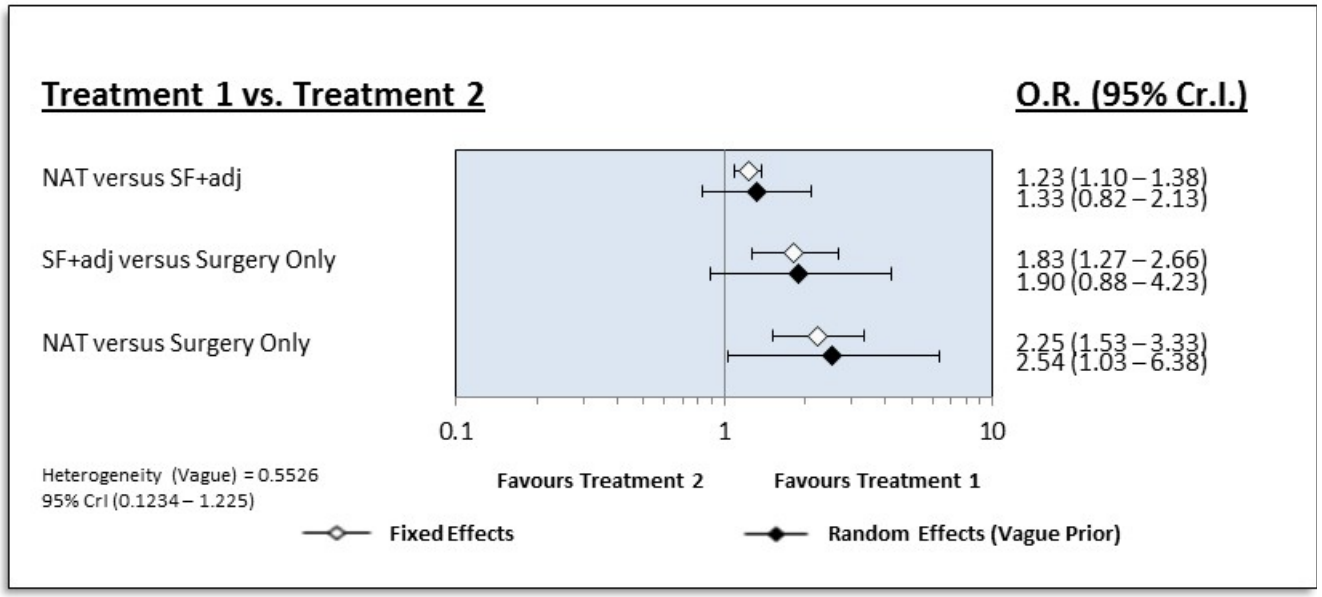
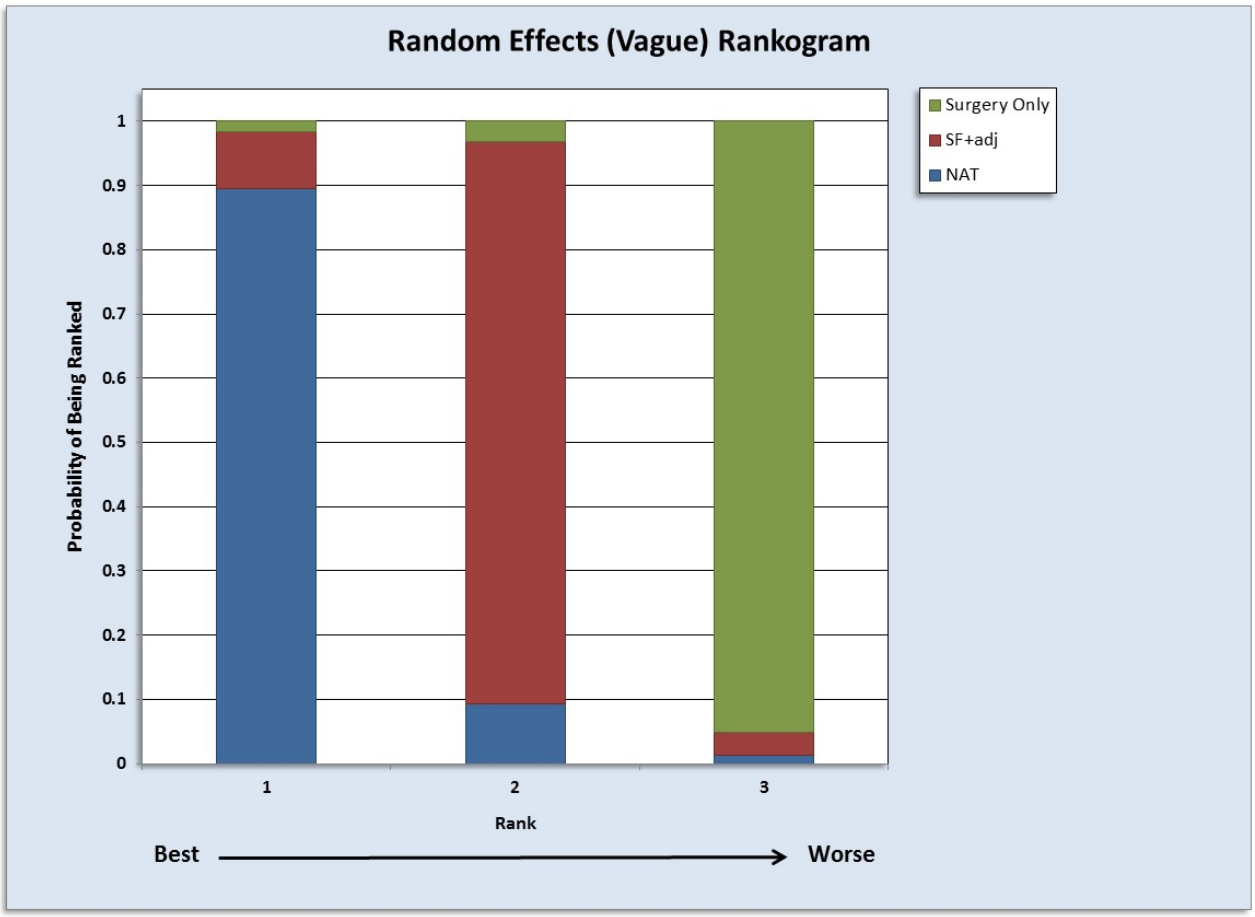


Figure Klxv: League table based on results of fixed effects and random effects (vague prior) models. Where odds ratio (O.R.) is greater than 1 treatment at top left is superior

NAT		
1.23 (1.10 – 1.38)	**SF+adj**	
2.25 (1.53 – 3.33)	1.83 (1.27 – 2.66)	**Surgery Only**

Figure Klxvi: Rankogram summarising surface under the cumulative ranking (SUCRA).



Risk-of-Bias Assessment

Figure Klxvii: Assessment of Risk of Bias for each included trials of neoadjuvant versus upfront surgery plus adjuvant therapy

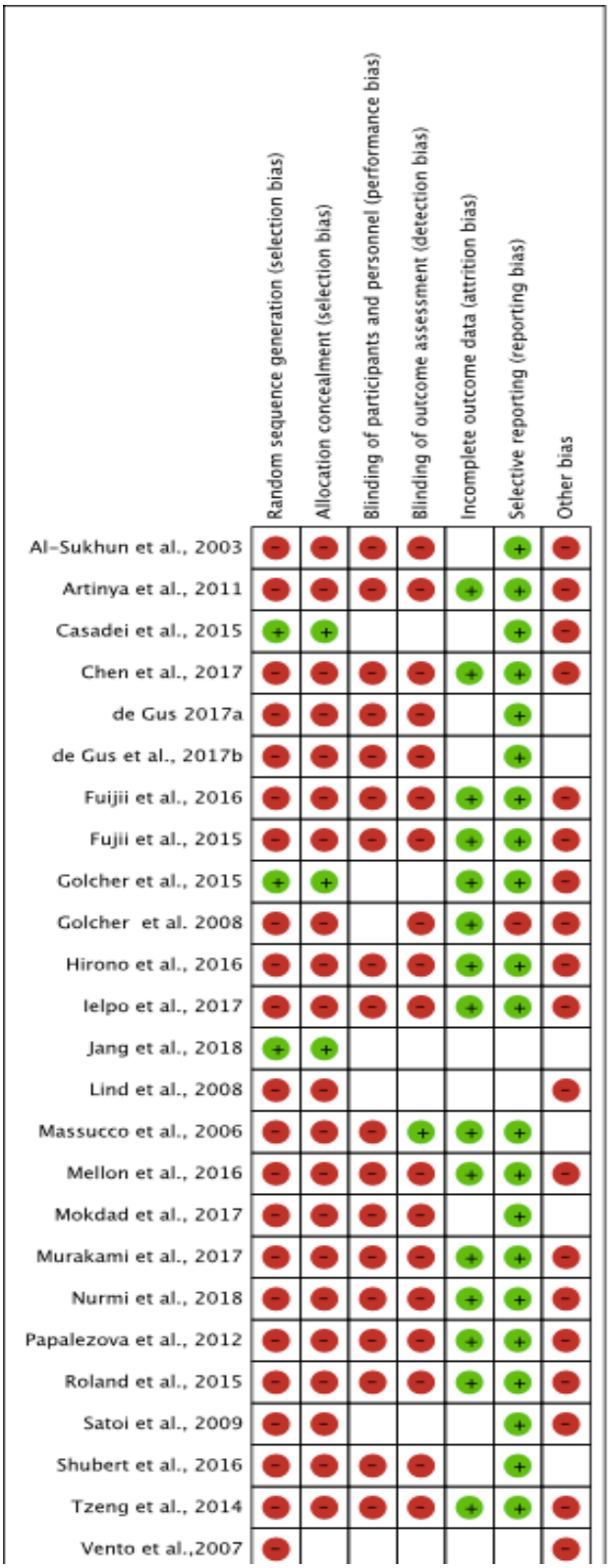
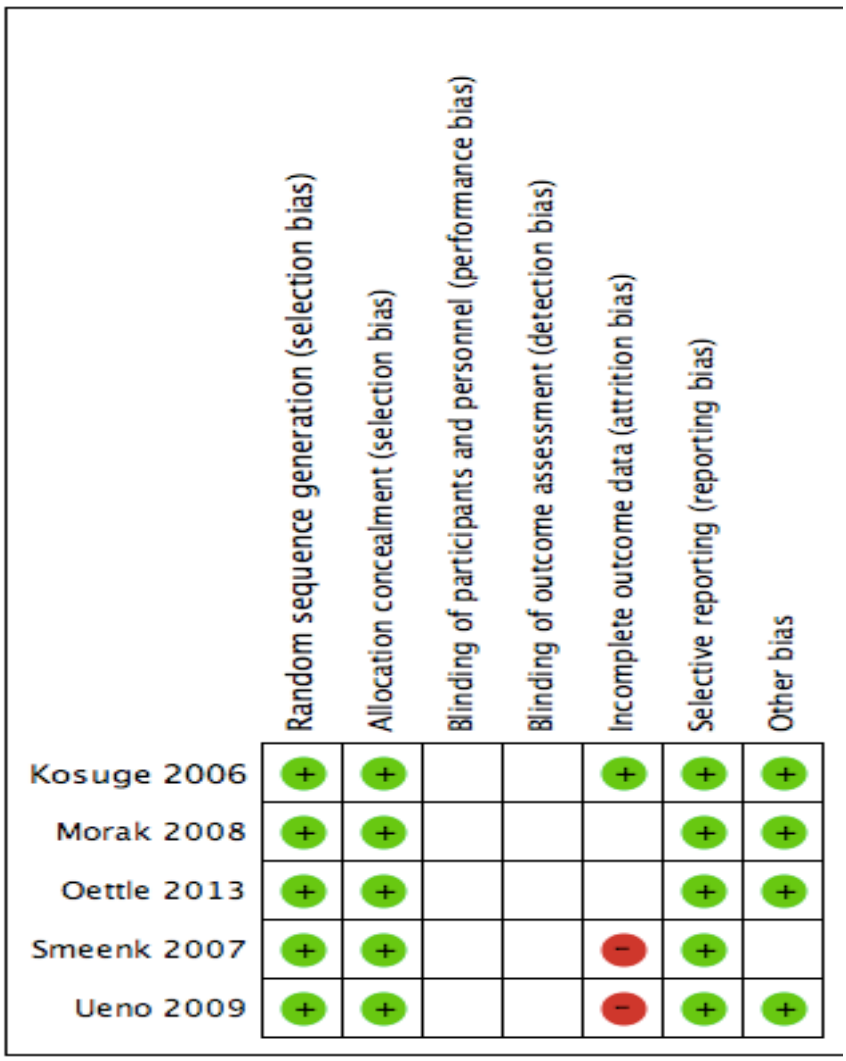
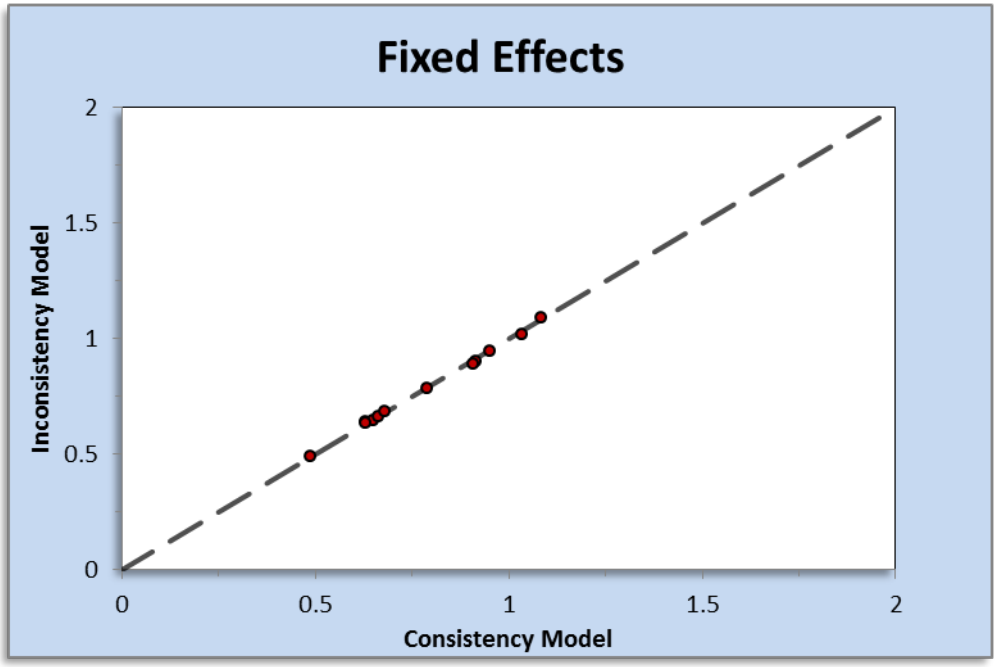


Figure Klxxvi: Assessment of Risk of Bias of Randomised Controlled Trials comparing Upfront Surgery plus Adjuvant Therapy versus Surgery Only.

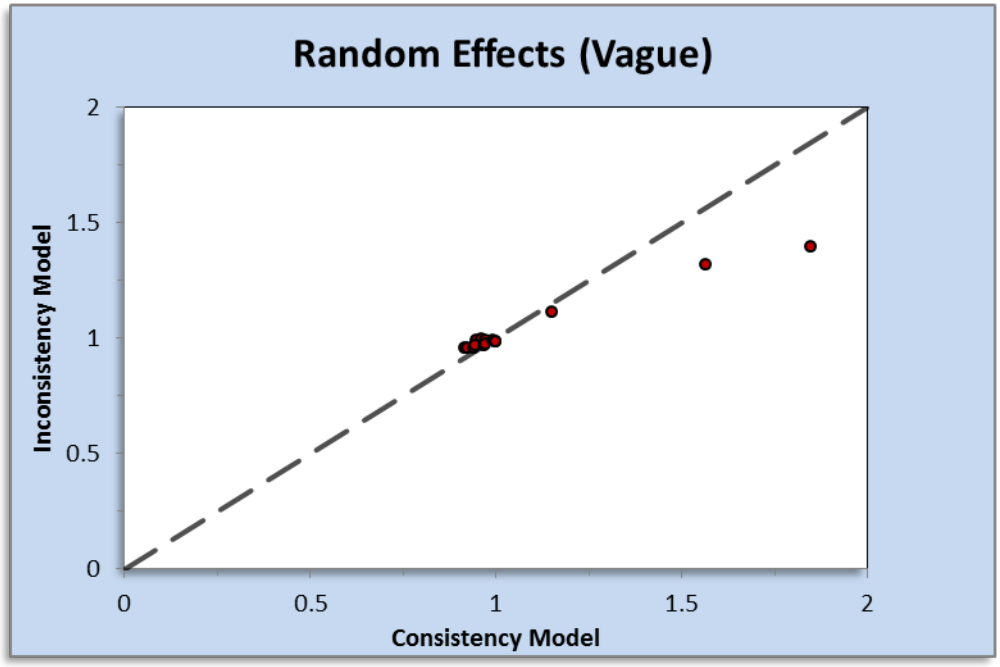


Assessment of Convergence and Inconsistency

Figure Klxviii: 1-year Survival Model Including Phase II/III Trials and Randomised Controlled Trials

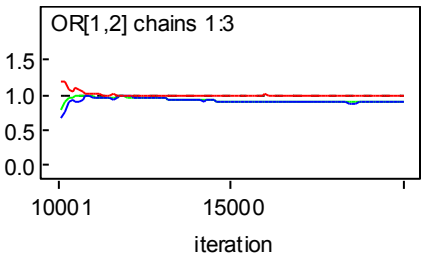


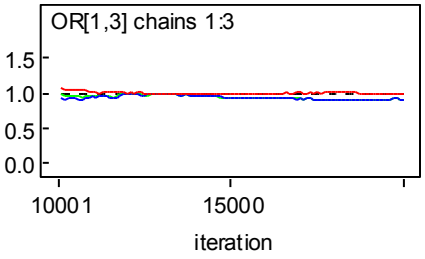
	Fixed Effects Odds Ratios		
	Treatment Comparisons	Consistency Model	InConsistency Model
OR[1,2]	SF+adj versus NAT	0.5679	0.5689
OR[1,3]	Surgery Only versus NAT	0.5493	N/A
OR[2,3]	Surgery Only versus SF+adj	0.9674	0.9695



	Random Effects Odds Ratios		
	Treatment Comparisons	Consistency Model	InConsistency Model
OR[1,2]	SF+adj versus NAT	0.5568	0.5332
OR[1,3]	Surgery Only versus NAT	0.524	N/A
OR[2,3]	Surgery Only versus SF+adj	0.9421	0.9374

Gelman Rubin statistic





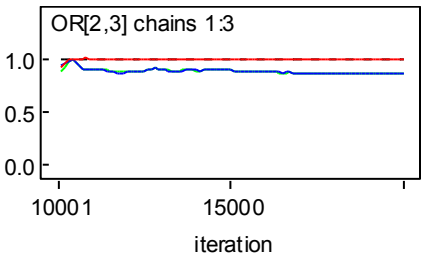
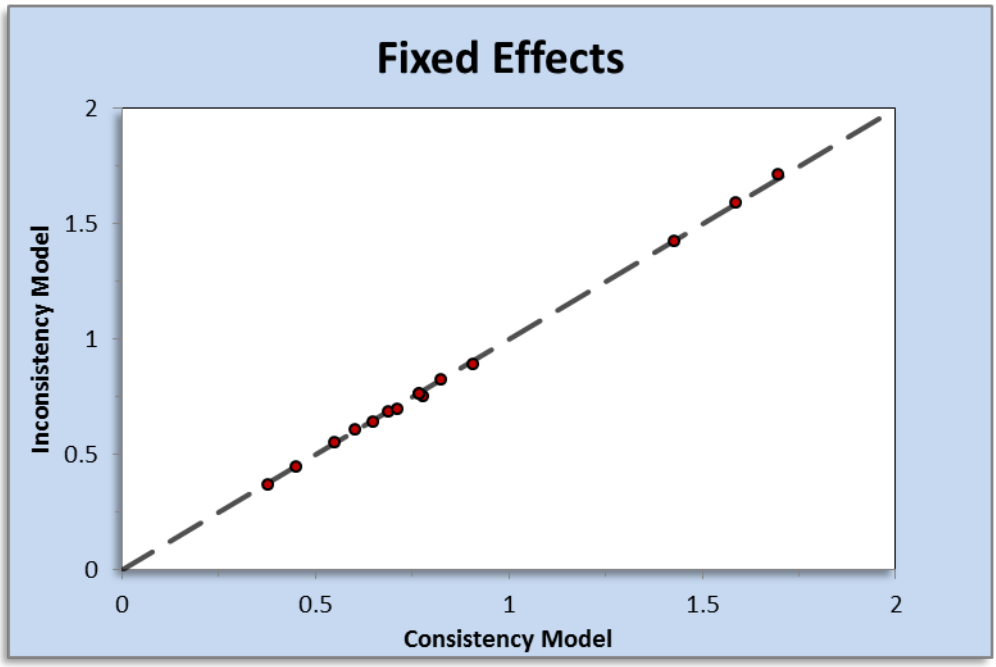
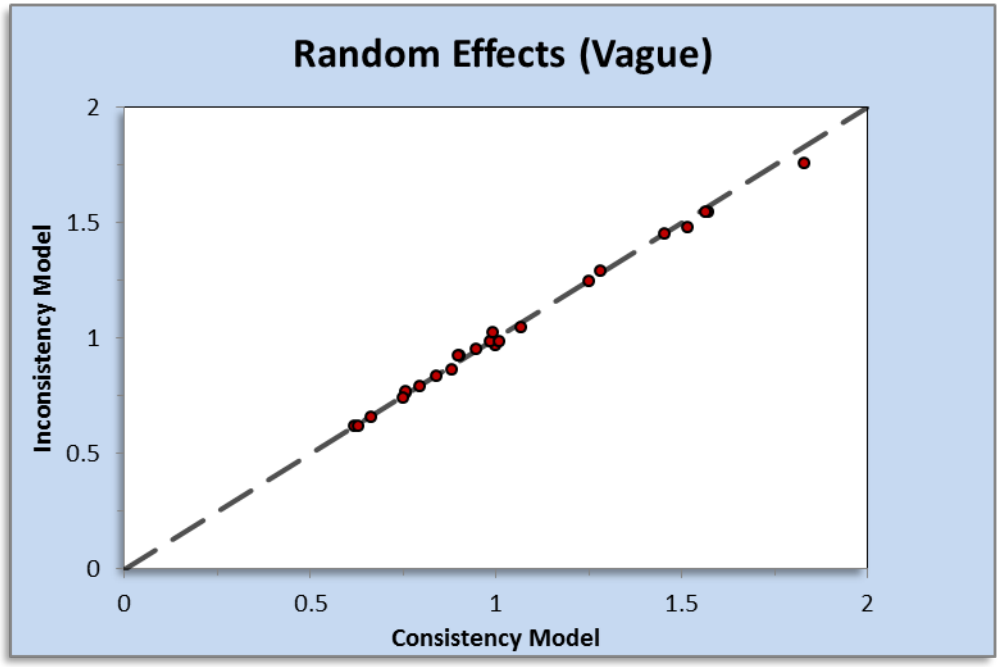


Figure Klxix: 1-year Survival Model Including Phase II/III Trials, Randomised Controlled Trials and Cohort Studies

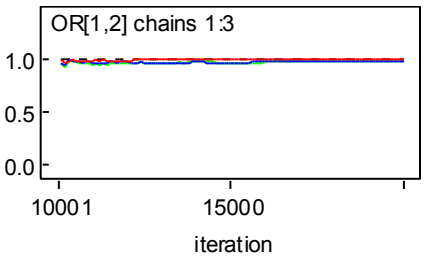


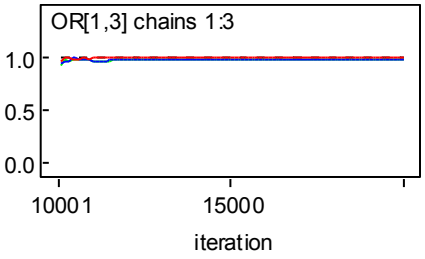
	Fixed Effects Odds Ratios		
	Treatment Comparisons	Consistency Model	InConsistency Model
OR[1,2]	SF+adj versus NAT	0.4263	0.4269
OR[1,3]	Surgery Only versus NAT	0.4125	N/A
OR[2,3]	Surgery Only versus SF+adj	0.9674	0.9644



	Random Effects Odds Ratios		
	Treatment Comparisons	Consistency Model	InConsistency Model
OR[1,2]	SF+adj versus NAT	0.509	0.5091
OR[1,3]	Surgery Only versus NAT	0.4815	N/A
OR[2,3]	Surgery Only versus SF+adj	0.9457	0.9426

Gelman Rubin statistic





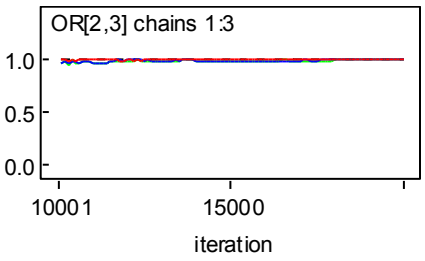
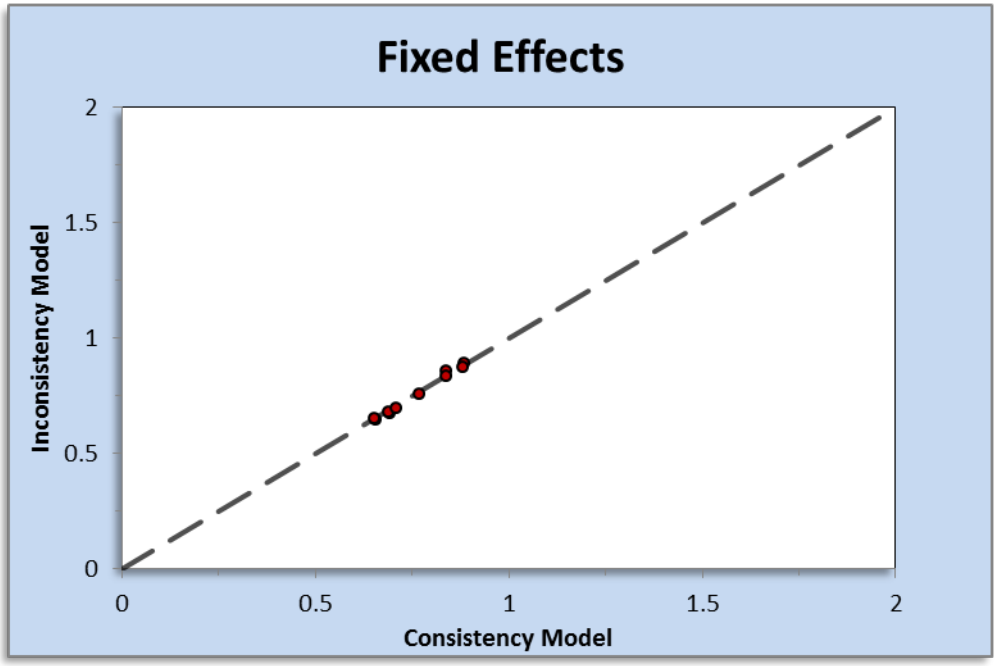
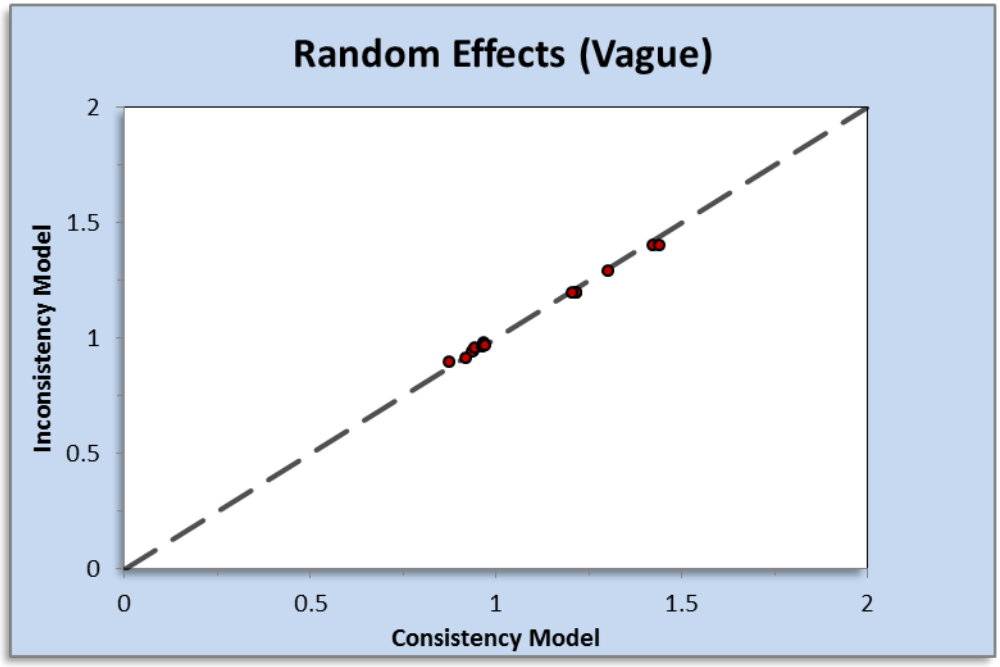


Figure Klxx: 2-year Survival Model Including Phase II/III Trials and Randomised Controlled Trials

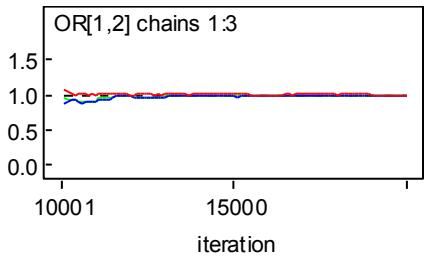


	Fixed Effects Odds Ratios		
	Treatment Comparisons	Consistency Model	InConsistency Model
OR[1,2]	SF+adj versus NAT	0.7607	0.7635
OR[1,3]	Surgery Only versus NAT	0.5911	N/A
OR[2,3]	Surgery Only versus SF+adj	0.7771	0.7766



	Random Effects Odds Ratios		
	Treatment Comparisons	Consistency Model	InConsistency Model
OR[1,2]	SF+adj versus NAT	0.8498	0.8450
OR[1,3]	Surgery Only versus NAT	0.6415	N/A
OR[2,3]	Surgery Only versus SF+adj	0.758	0.7498

Gelman Rubin statistic



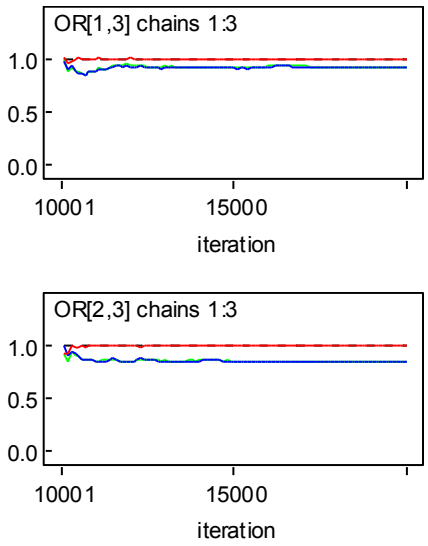
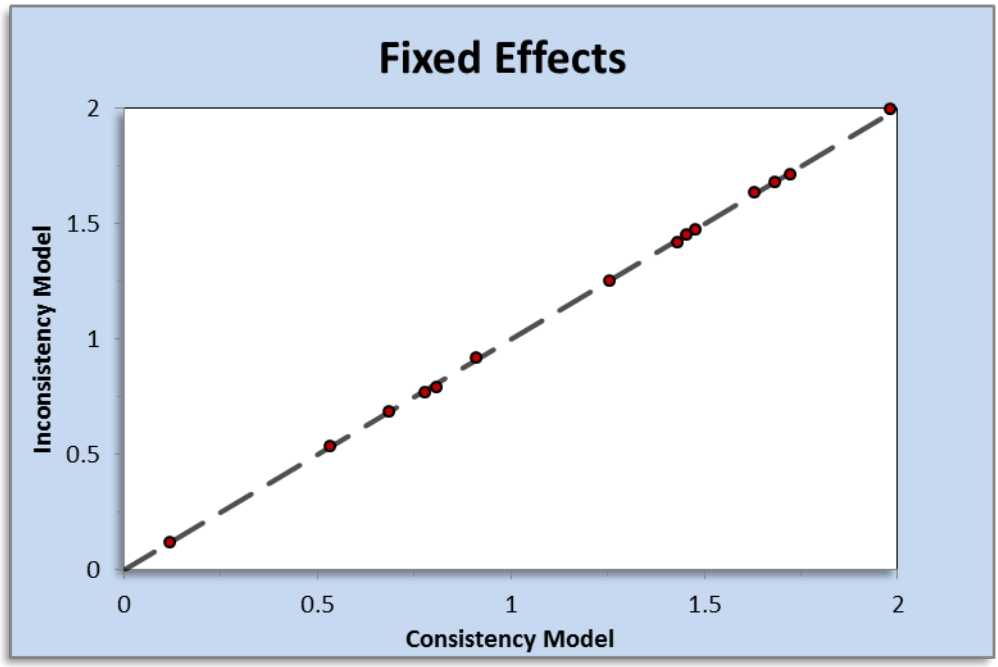
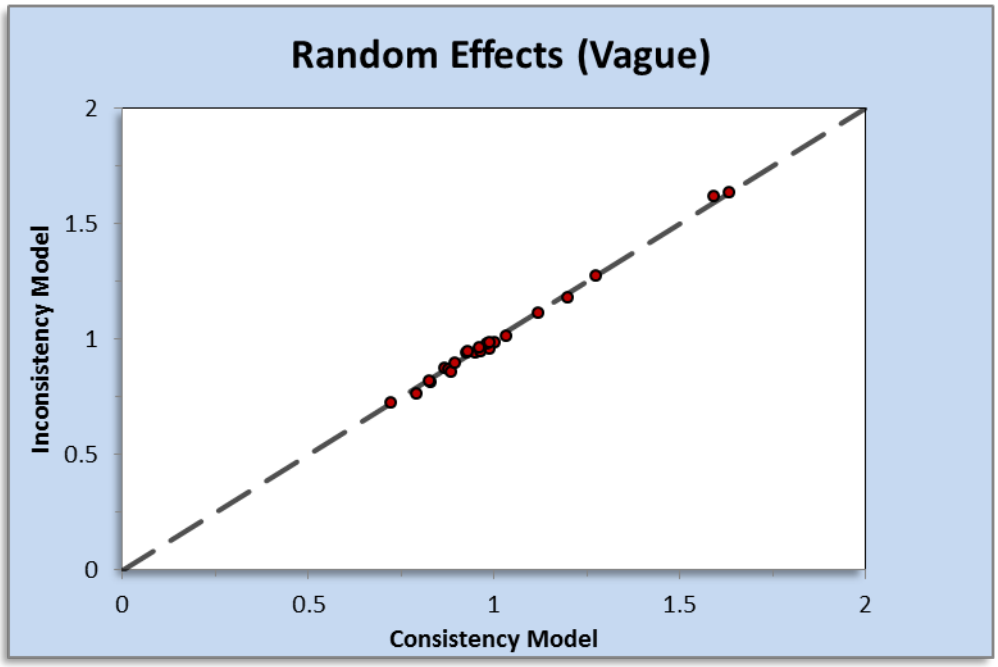


Figure Klxxi: 2-year Survival Model Including Phase II/III Trials, Randomised Controlled Trials and Cohort Studies



	Fixed Effects Odds Ratios		
	Treatment Comparisons	Consistency Model	InConsistency Model
OR[1,2]	SF+adj versus NAT	0.3871	0.3877
OR[1,3]	Surgery Only versus NAT	0.2997	N/A
OR[2,3]	Surgery Only versus SF+adj	0.775	0.7769



	Random Effects Odds Ratios		
	Treatment Comparisons	Consistency Model	InConsistency Model
OR[1,2]	SF+adj versus NAT	0.6073	0.6065
OR[1,3]	Surgery Only versus NAT	0.4595	N/A
OR[2,3]	Surgery Only versus SF+adj	0.7591	0.7604

Gelman Rubin statistic

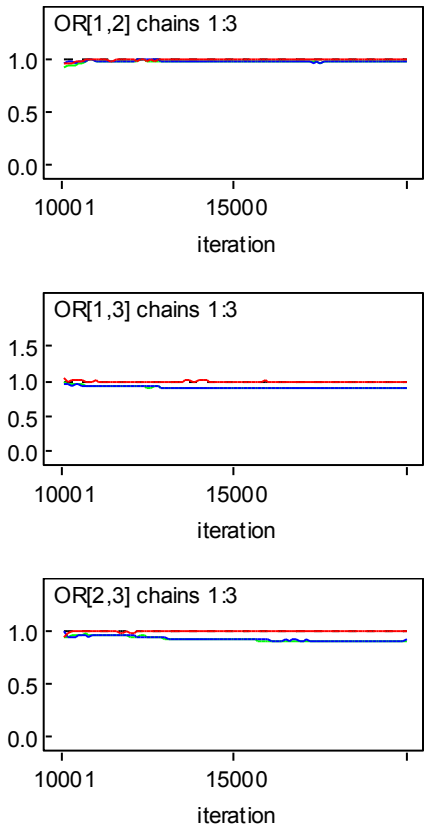
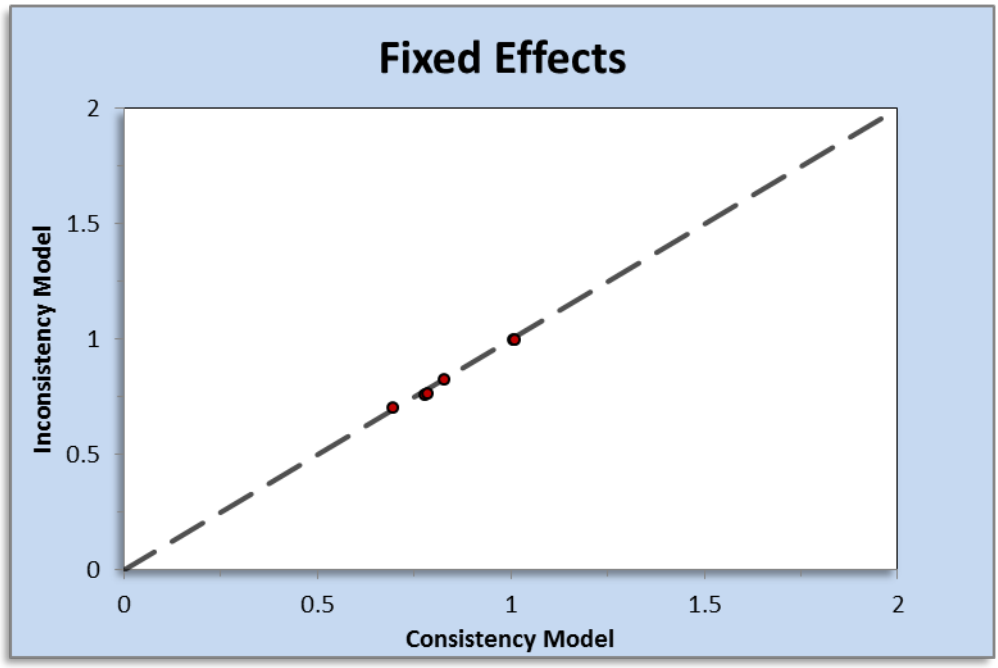
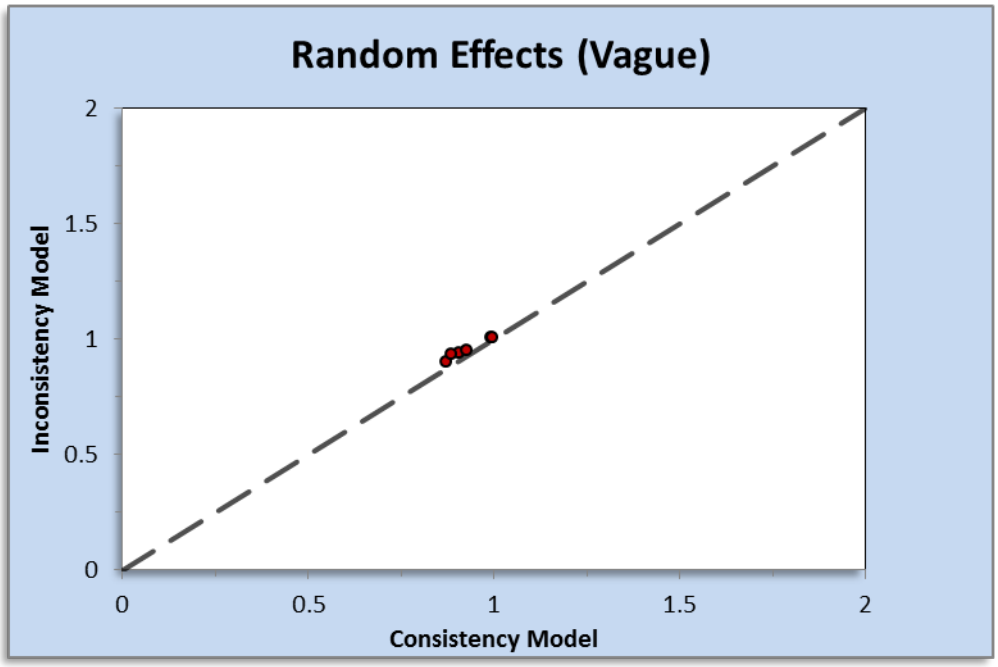


Figure Klxxii: 3-year Survival Model Including Phase II/III Trials and Randomised Controlled Trials

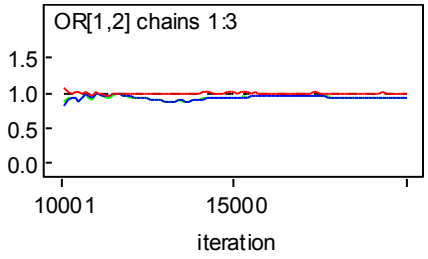


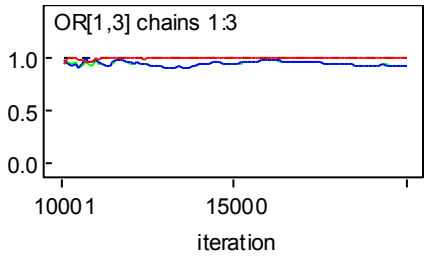
	Fixed Effects Odds Ratios		
	Treatment Comparisons	Consistency Model	InConsistency Model
OR[1,2]	Sfadj versus NAT	1.761	1.7372
OR[1,3]	Surgery Only versus NAT	0.875	N/A
OR[2,3]	Surgery Only versus Sfadj	0.4976	0.4977



	Random Effects Odds Ratios		
	Treatment Comparisons	Consistency Model	InConsistency Model
OR[1,2]	Sfadj versus NAT	1.737	1.7565
OR[1,3]	Surgery Only versus NAT	0.8592	N/A
OR[2,3]	Surgery Only versus Sfadj	0.4984	0.5002

Gelman Rubin statistic


OR[1,2] chains 1:3
1.5
1.0
0.5
0.0
10001
15000
iteration


OR[1,3] chains 1:3
1.0
0.5
0.0
10001
15000
iteration

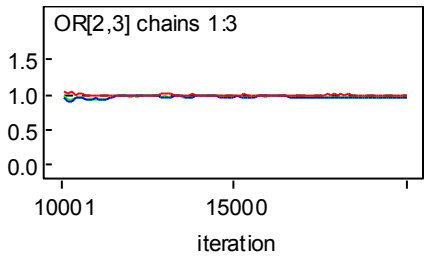
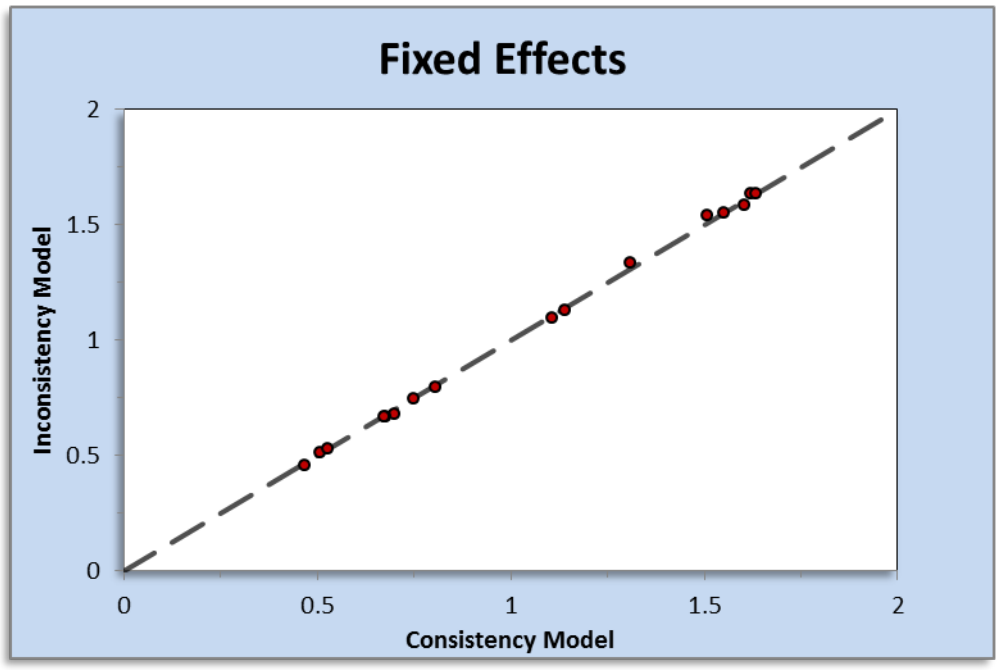
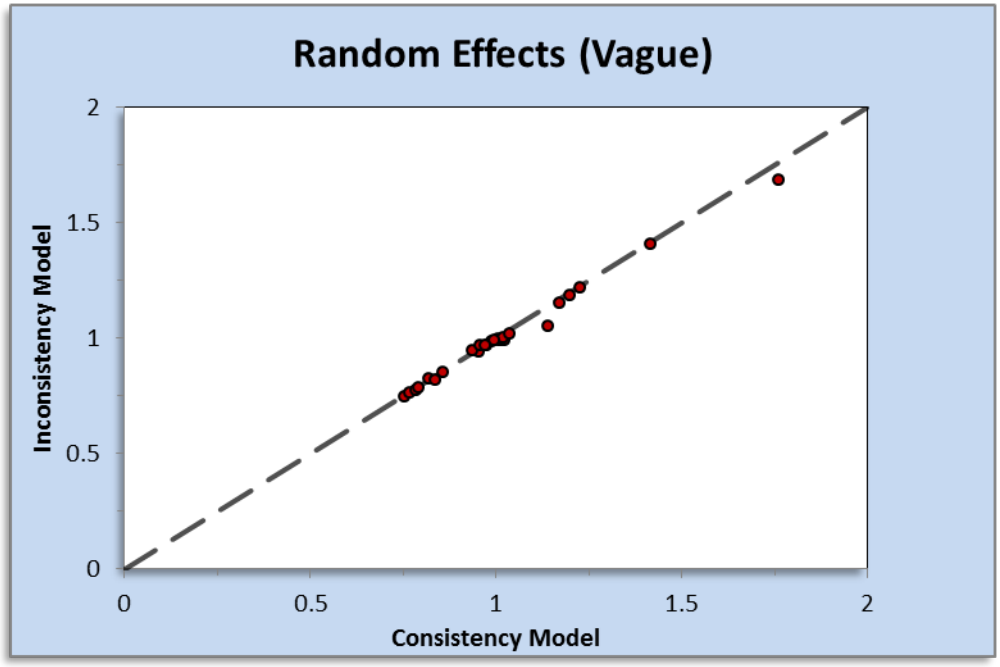

OR[2,3] chains 1:3
1.5
1.0
0.5
0.0
10001
15000
iteration

Figure Klxxiii: 3-year Survival Model Including Phase II/III Trials, Randomised Controlled Trials and Cohort Studies

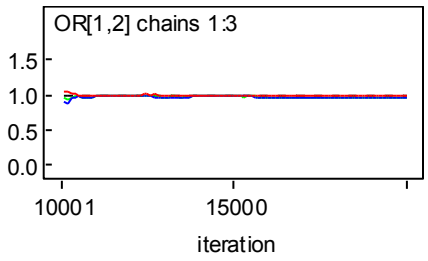


	Fixed Effects Odds Ratios		
	Treatment Comparisons	Consistency Model	InConsistency Model
OR[1,2]	Sfadj versus NAT	0.6204	0.6211
OR[1,3]	Surgery Only versus NAT	0.3085	N/A
OR[2,3]	Surgery Only versus Sfadj	0.4968	0.4980



	Random Effects Odds Ratios		
	Treatment Comparisons	Consistency Model	InConsistency Model
OR[1,2]	Sfadj versus NAT	0.6669	0.6673
OR[1,3]	Surgery Only versus NAT	0.3326	N/A
OR[2,3]	Surgery Only versus Sfadj	0.4973	0.4989

Gelman Rubin statistic



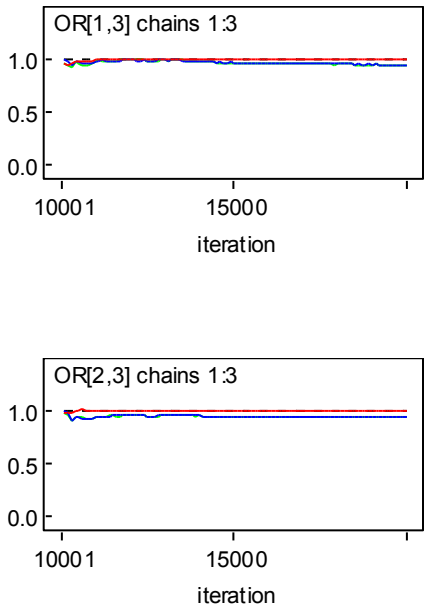
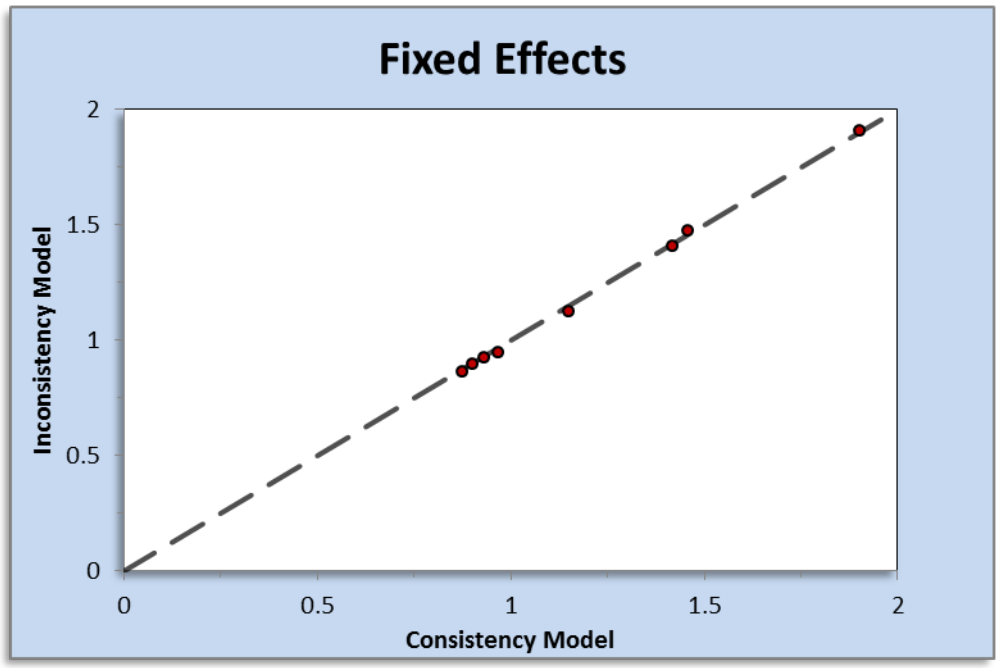
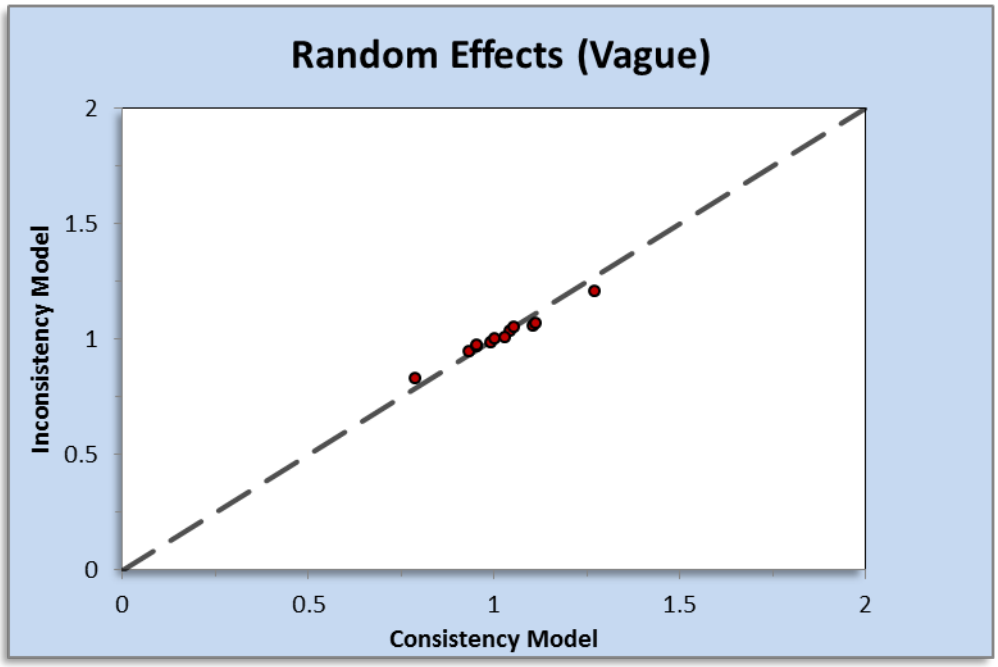


Figure Klxxiva: 5-year Survival Model Including Phase II/III Trials and Randomised Controlled Trials

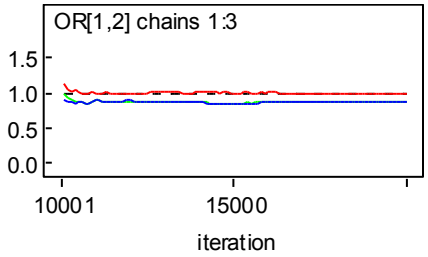


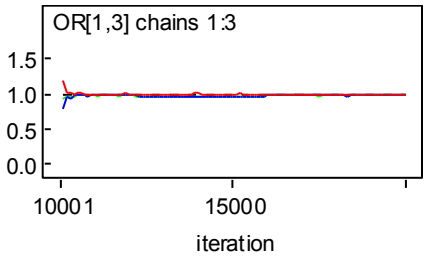
	Fixed Effects Odds Ratios		
	Treatment Comparisons	Consistency Model	InConsistency Model
OR[1,2]	SF+adj versus NAT	0.3943	0.4011
OR[1,3]	Surgery Only versus NAT	0.1876	N/A
OR[2,3]	Surgery Only versus SF+adj	0.4747	0.4742



	Random Effects Odds Ratios		
	Treatment Comparisons	Consistency Model	InConsistency Model
OR[1,2]	SF+adj versus NAT	0.5083	0.5230
OR[1,3]	Surgery Only versus NAT	0.1796	N/A
OR[2,3]	Surgery Only versus SF+adj	0.3545	0.3542

Gelman Rubin statistic




OR[1,3] chains 1:3
1.5
1.0
0.5
0.0
10001
15000
iteration

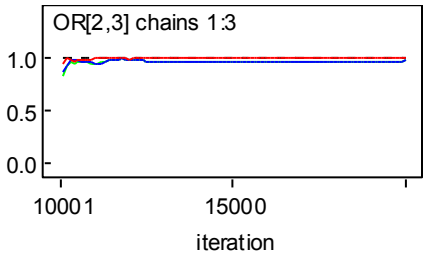
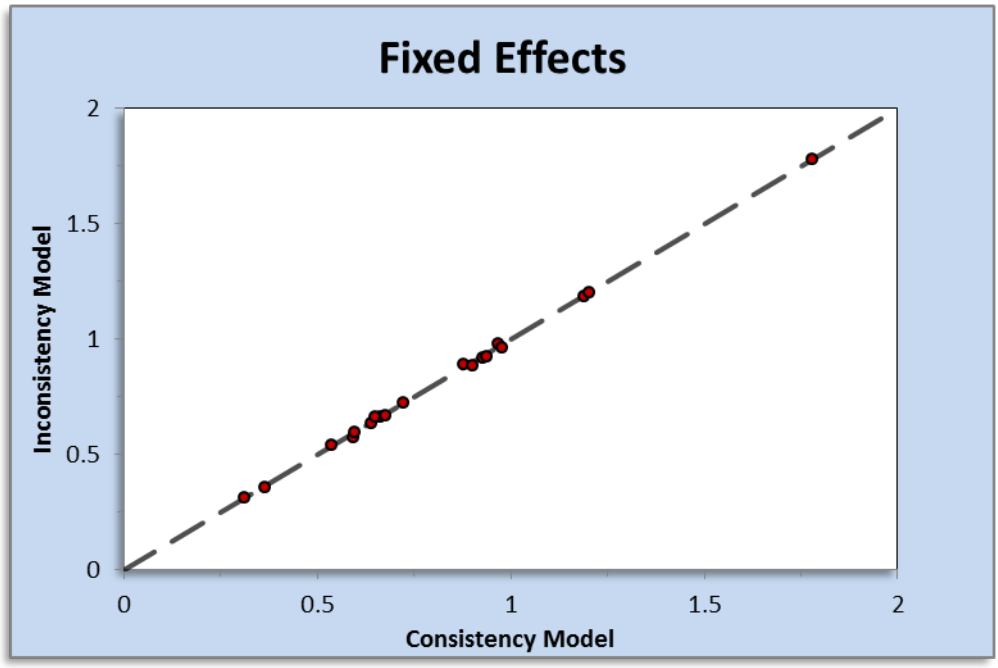
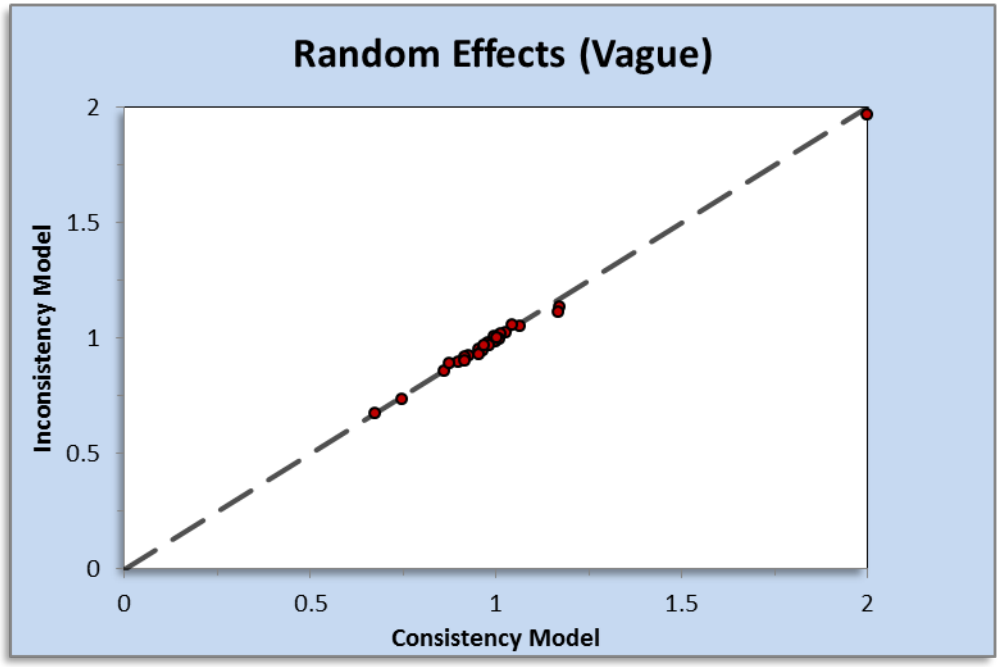

OR[2,3] chains 1:3
1.0
0.5
0.0
10001
15000
iteration

Figure Klxxivb: 5-year Survival Model Including Phase II/III Trials, Randomised Controlled Trials and Cohort Studies

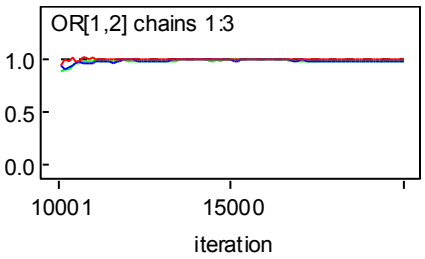


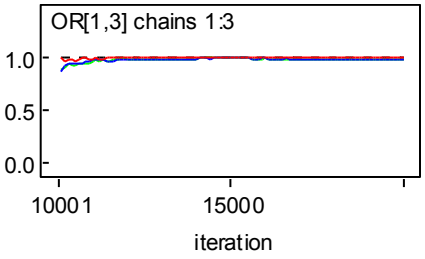
	Fixed Effects Odds Ratios		
	Treatment Comparisons	Consistency Model	InConsistency Model
OR[1,2]	SF+adj versus NAT	0.8124	0.8129
OR[1,3]	Surgery Only versus NAT	0.3874	N/A
OR[2,3]	Surgery Only versus SF+adj	0.4763	0.4758

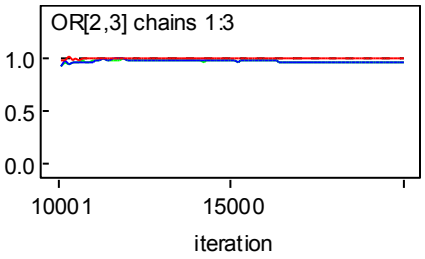


	Random Effects Odds Ratios		
	Treatment Comparisons	Consistency Model	InConsistency Model
OR[1,2]	SF+adj versus NAT	0.7595	0.7639
OR[1,3]	Surgery Only versus NAT	0.2743	N/A
OR[2,3]	Surgery Only versus SF+adj	0.3643	0.3628

Gelman Rubin statistic






OR[2,3] chains 1:3
1.0
0.5
0.0
10001
15000
iteration

Appendix L

Bayesian Network Meta-analysis: Resectable Pancreatic Cancer

** SF= surgery first pathway; NAT = neoadjuvant therapy; SF+adj = surgery first plus adjuvant therapy; surgery only= surgical resection no adjuvant therapy*

1-year Survival

Figure Li: Results of fixed effects and random effects (vague prior) models

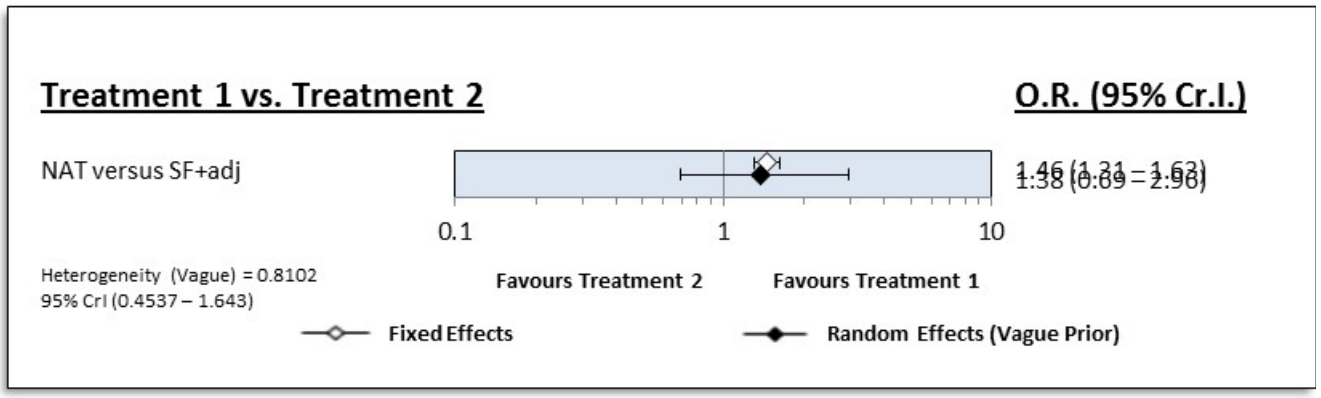
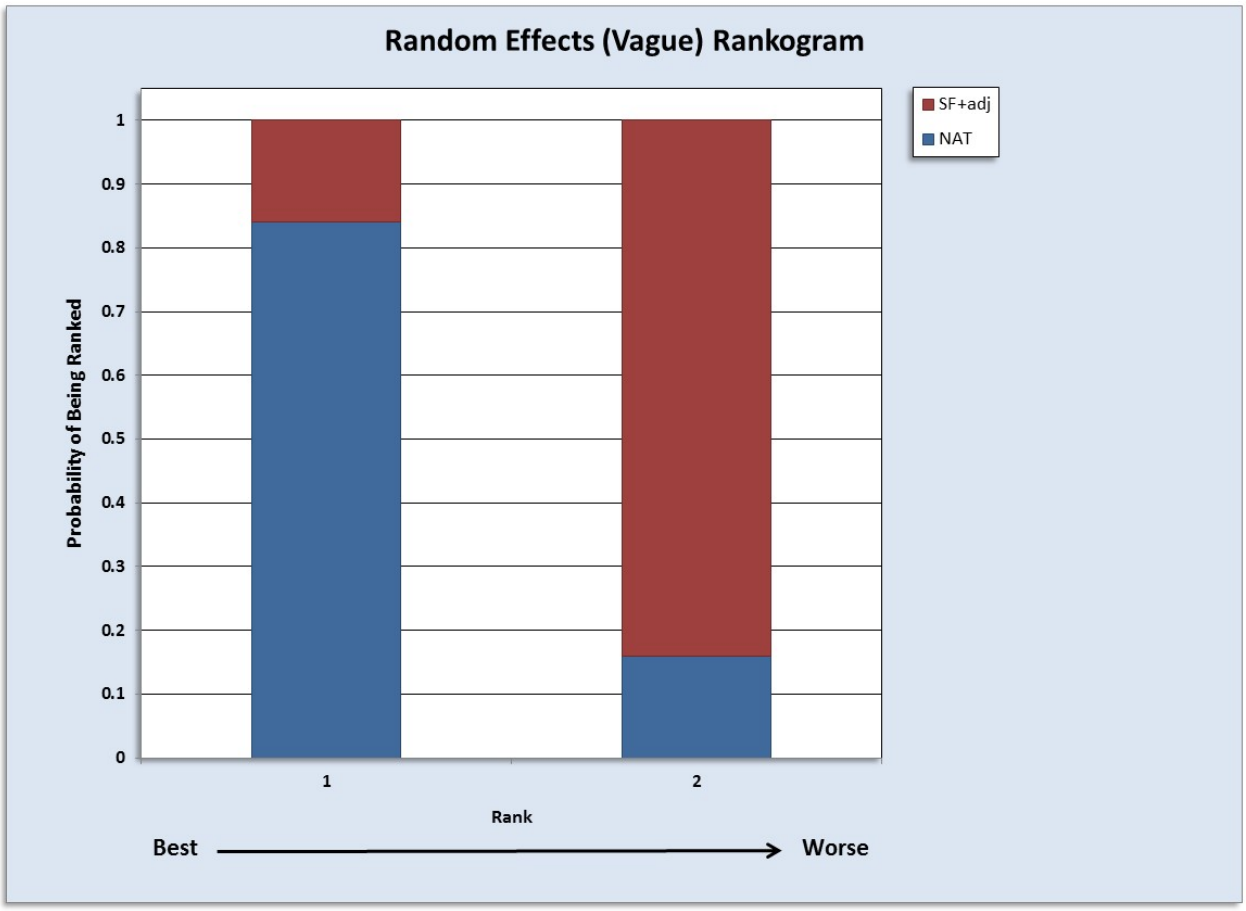


Figure Lii: League table based on results of fixed effects and random effects (vague prior) models. Where odds ratio (O.R.) is greater than 1 treatment at top left is superior

NAT	
1.46 (1.31 – 1.63)	SF+adj

Figure Liii: Rankogram summarising surface under the cumulative ranking (SUCRA).

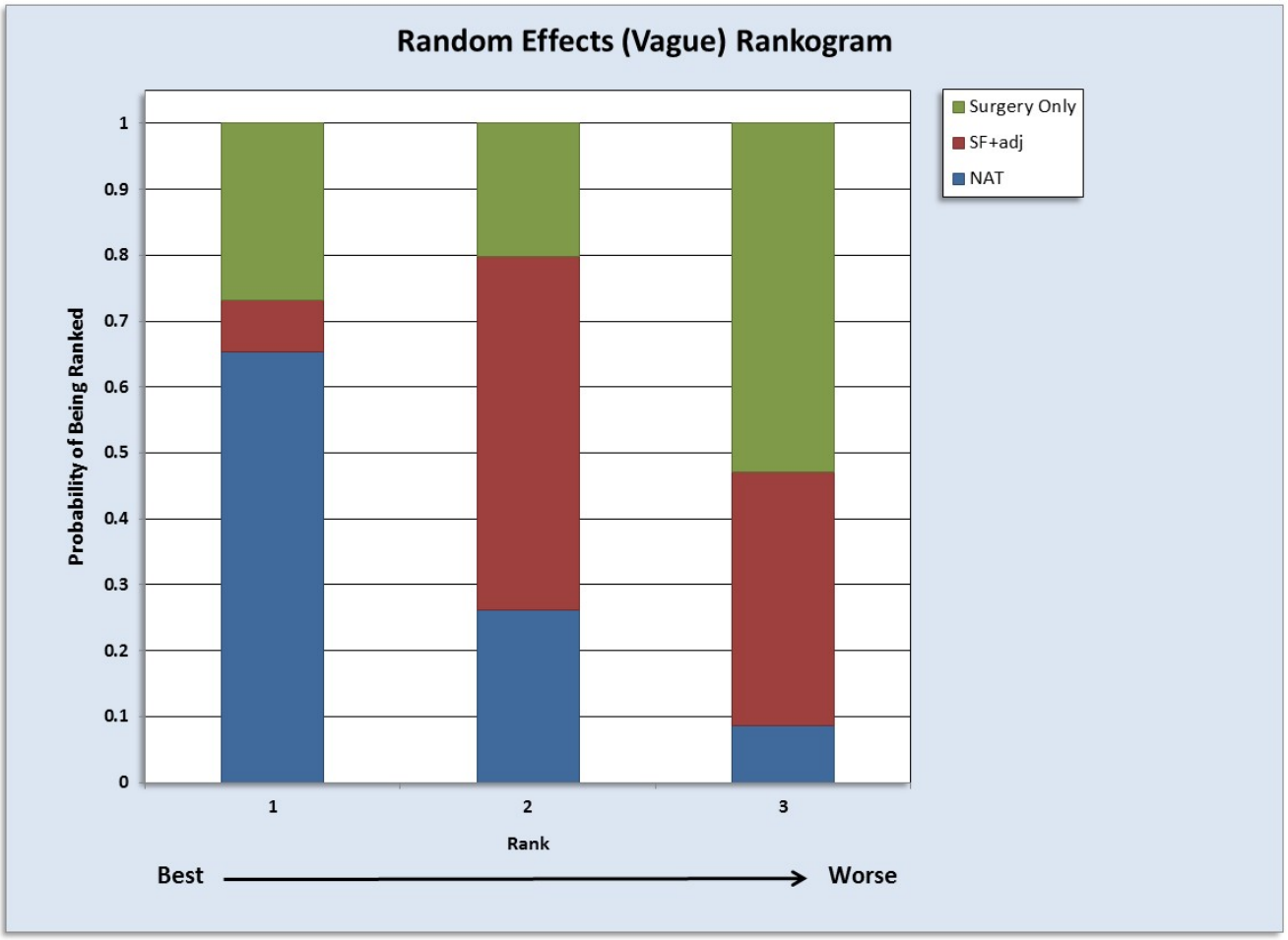


1-year survival: inclusion of RCTs

Figure Liv: League table based on results of fixed effects and random effects (vague prior) models. Where odds ratio (O.R.) is greater than 1 treatment at top left is superior

NAT		
1.46 (1.32 – 1.63)	**SF+adj**	
1.52 (0.99 – 2.31)	1.04 (0.69 – 1.56)	**Surgery Only**

Figure Lv: Rankogram summarising surface under the cumulative ranking (SUCRA).



2-year Survival Rates

Figure Lvi: Results of fixed effects and random effects (vague prior) models

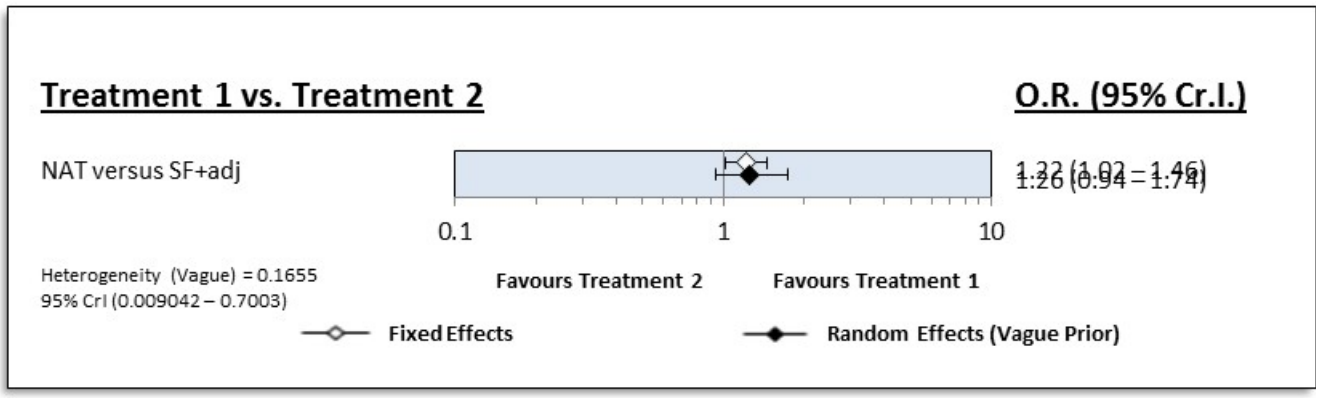
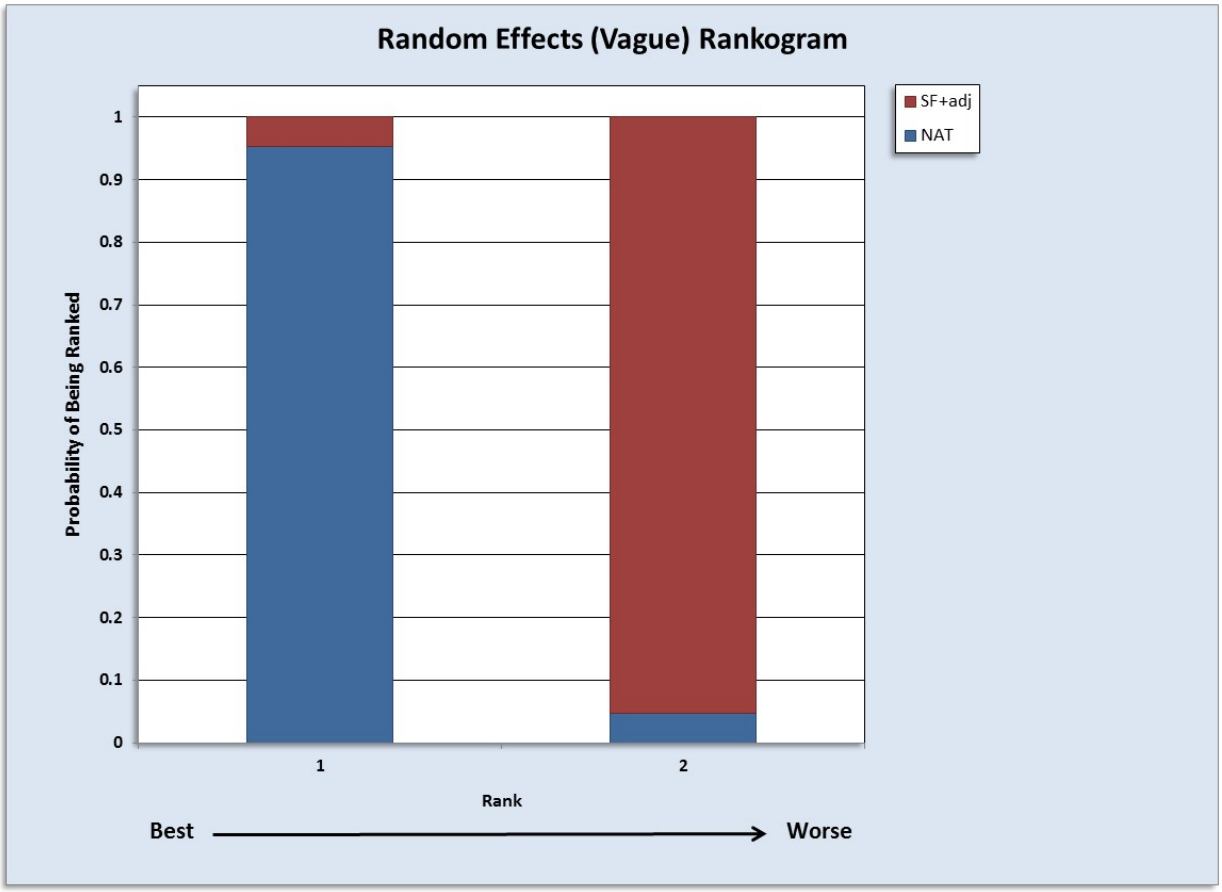


Figure Lvii: League table based on results of fixed effects and random effects (vague prior) models. Where odds ratio (O.R.) is greater than 1 treatment at top left is superior

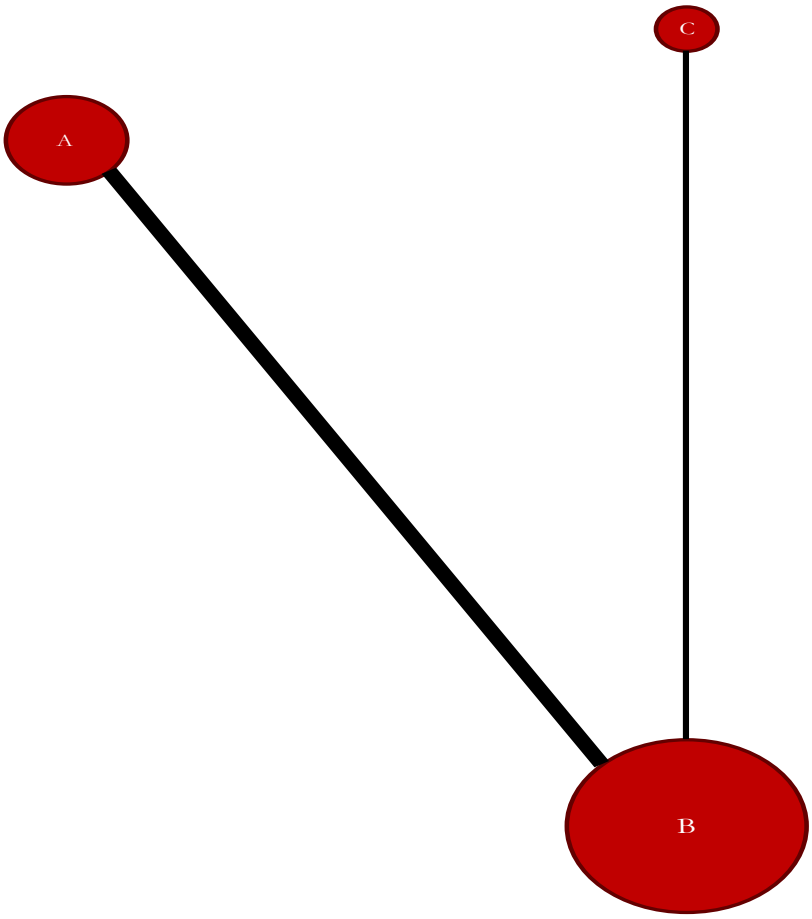
NAT	
1.22 (1.02 – 1.46)	**SF+adj**

Figure Lviii: Rankogram summarising surface under the cumulative ranking (SUCRA).



2-year survival: inclusion of RCTs

Figure Lix: Bayesian Network Meta-analysis of Neoadjuvant therapy versus Upfront surgery plus adjuvant therapy versus surgery only

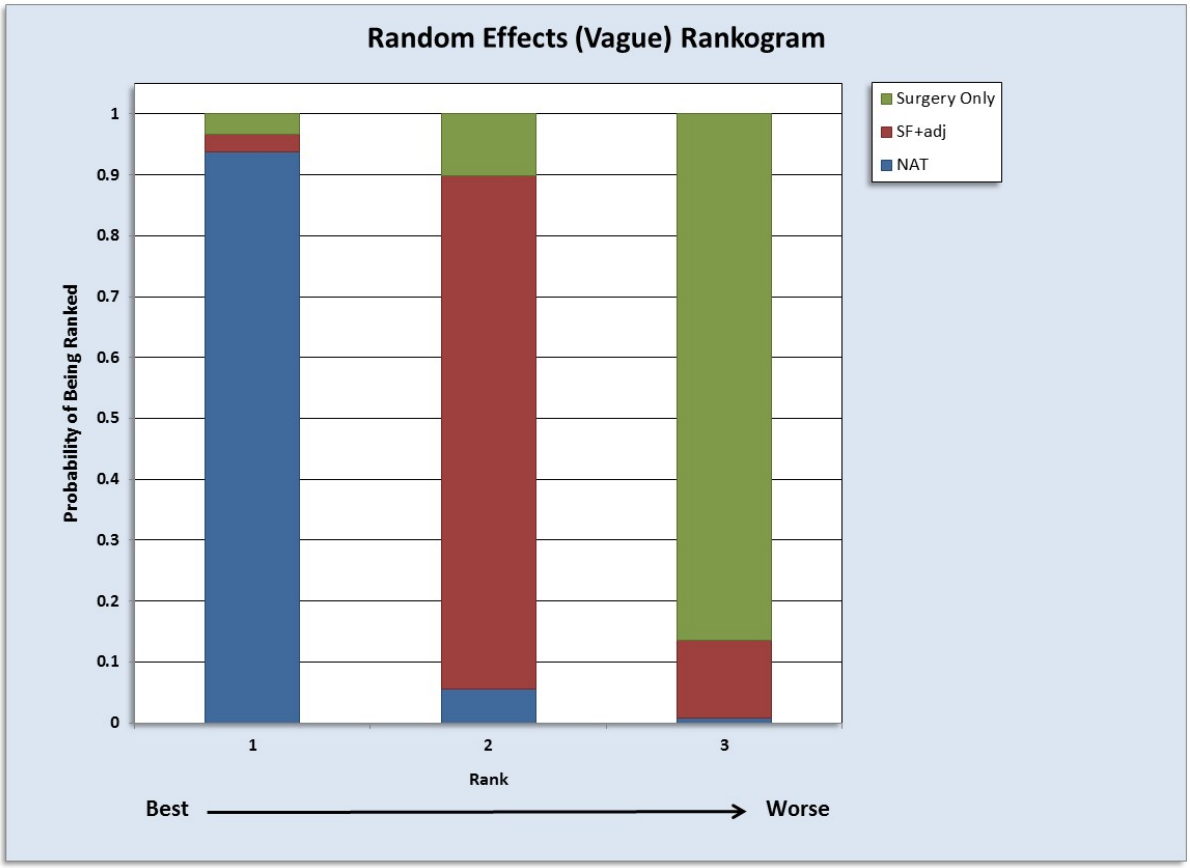


Drug	Abbreviation
NAT	A
SF+adj	B
Surgery Only	C

Figure Lx: League table based on results of fixed effects and random effects (vague prior) models. Where odds ratio (O.R.) is greater than 1 treatment at top left is superior

NAT		
1.23 (1.03 – 1.46)	**SF+adj**	
1.58 (1.06 – 2.37)	1.29 (0.89 – 1.85)	**Surgery Only**

Figure Lxi: Rankogram summarising surface under the cumulative ranking (SUCRA).



3-year Survival Rates

Figure Lxii: Results of fixed effects and random effects (vague prior) models

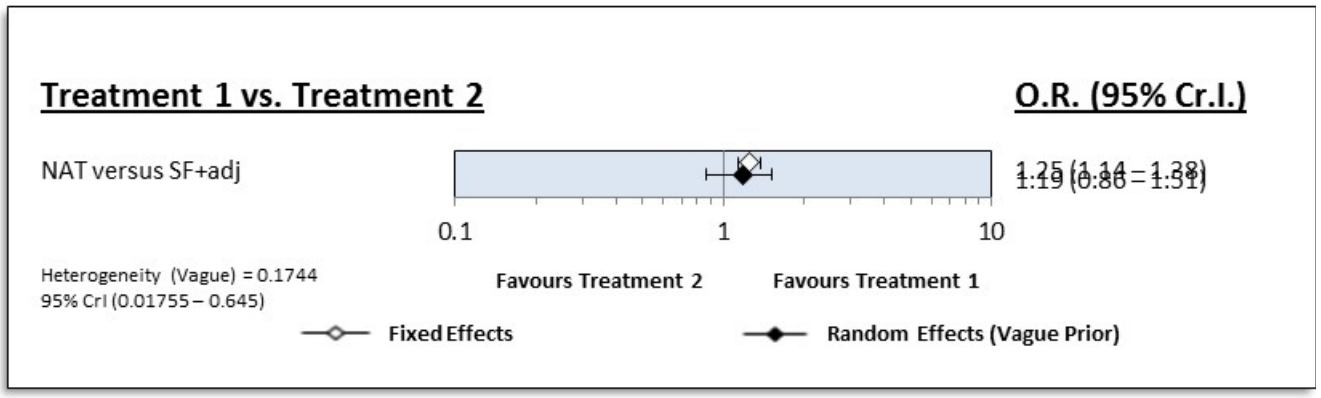
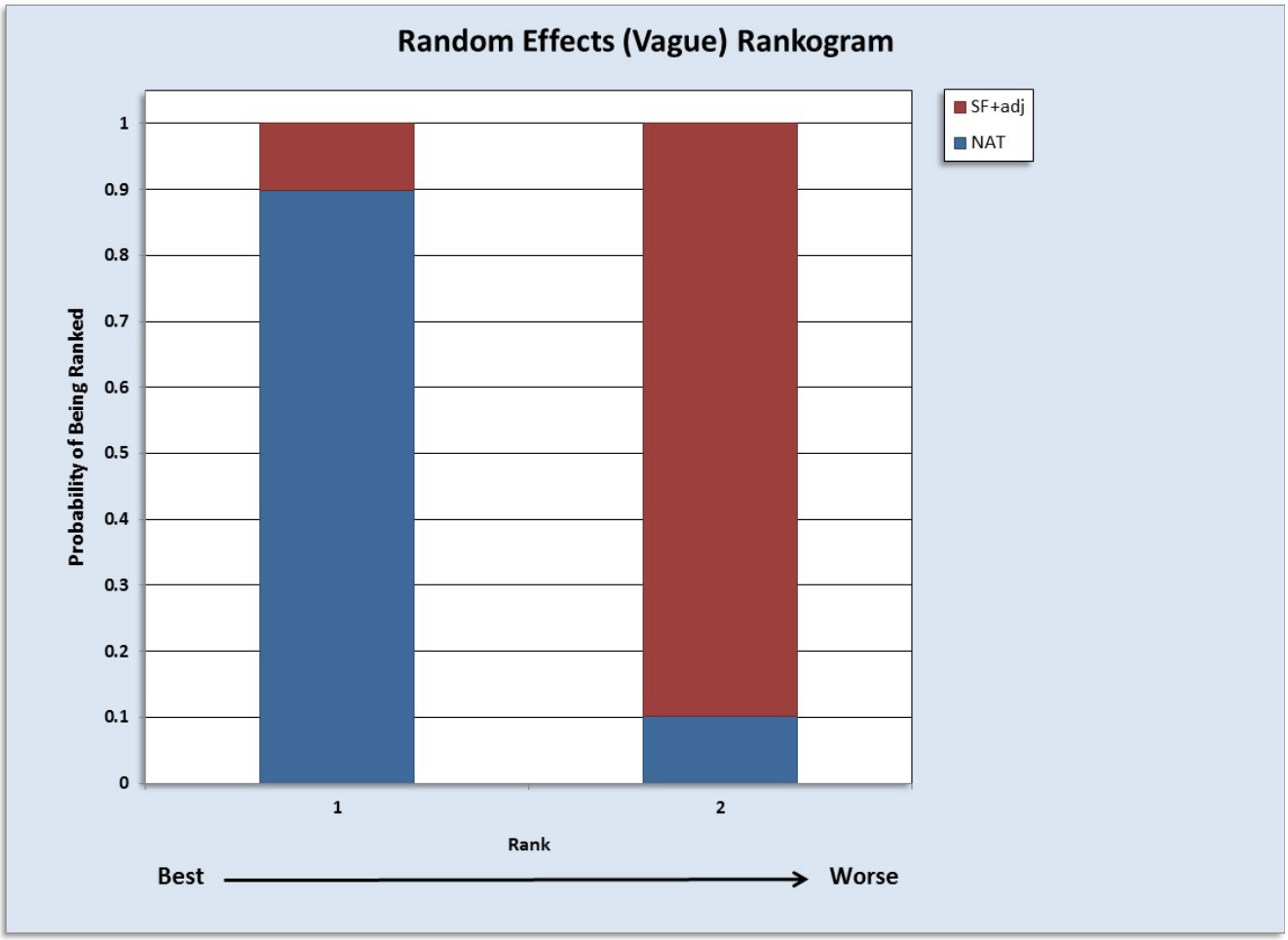


Figure Lxiii: League table based on results of fixed effects and random effects (vague prior) models. Where odds ratio (O.R.) is greater than 1 treatment at top left is superior

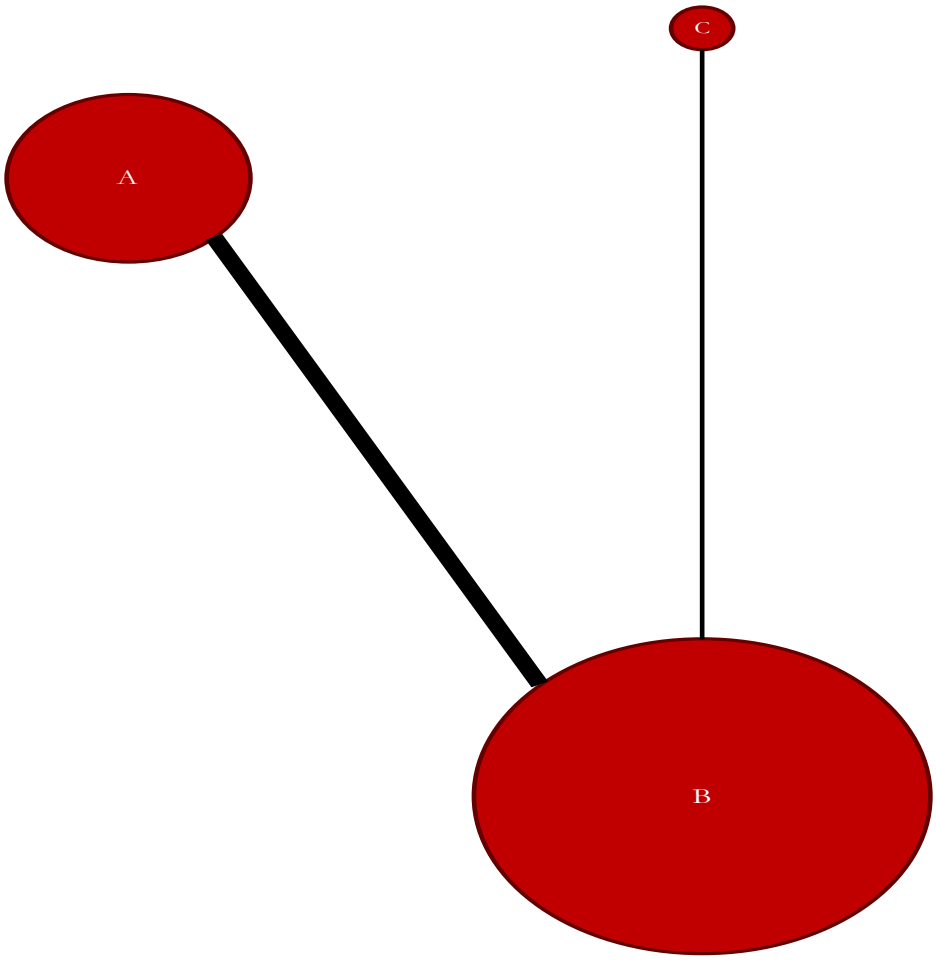
NAT	
1.25 (1.14 – 1.38)	**SF+adj**

Figure Lxviv: Rankogram summarising surface under the cumulative ranking (SUCRA).



3-year survival: inclusion of RCTs

Figure Lxv: Bayesian Network Meta-analysis of Neoadjuvant therapy versus Upfront surgery plus adjuvant therapy versus surgery only

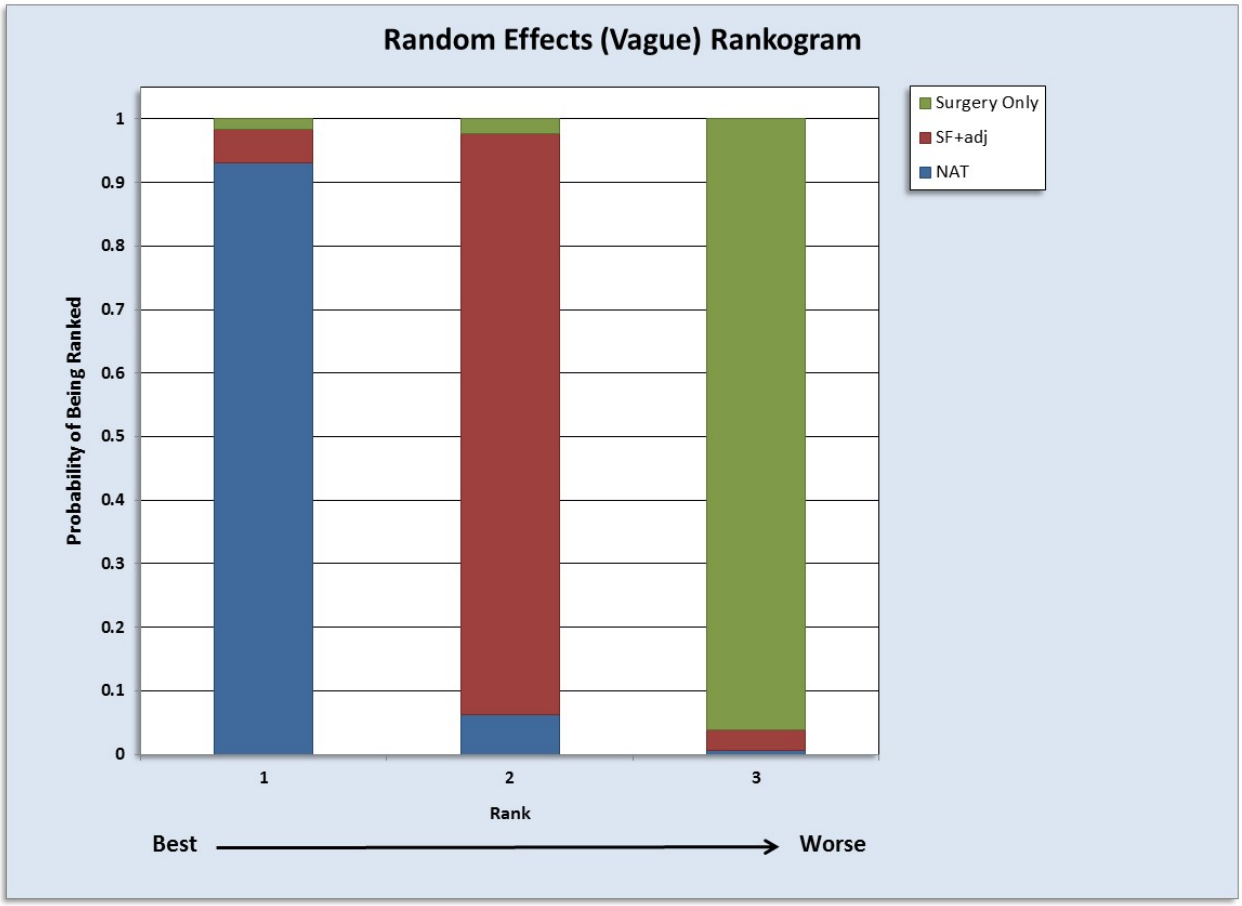


Drug	Abbreviation
NAT	A
SF+adj	B
Surgery Only	C

Figure Lxvi: League table based on results of fixed effects and random effects (vague prior) models. Where odds ratio (O.R.) is greater than 1 treatment at top left is superior

NAT		
1.27 (1.15 – 1.39)	**SF+adj**	
2.55 (1.56 – 4.17)	2.01 (1.24 – 3.26)	**Surgery Only**

Figure Lxvii: Rankogram summarising surface under the cumulative ranking (SUCRA).



4-year Survival Rates

Figure Lxviii: Results of fixed effects and random effects (vague prior) models

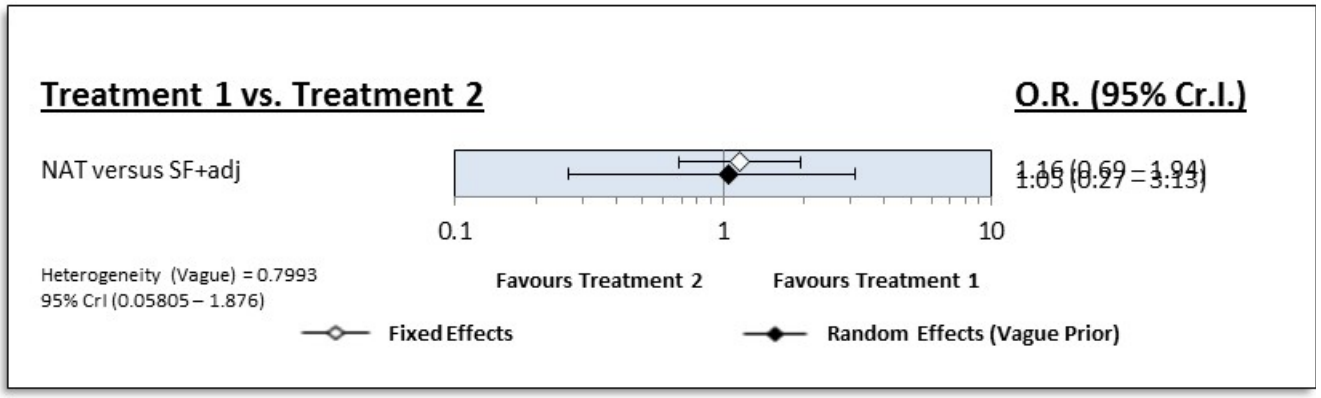
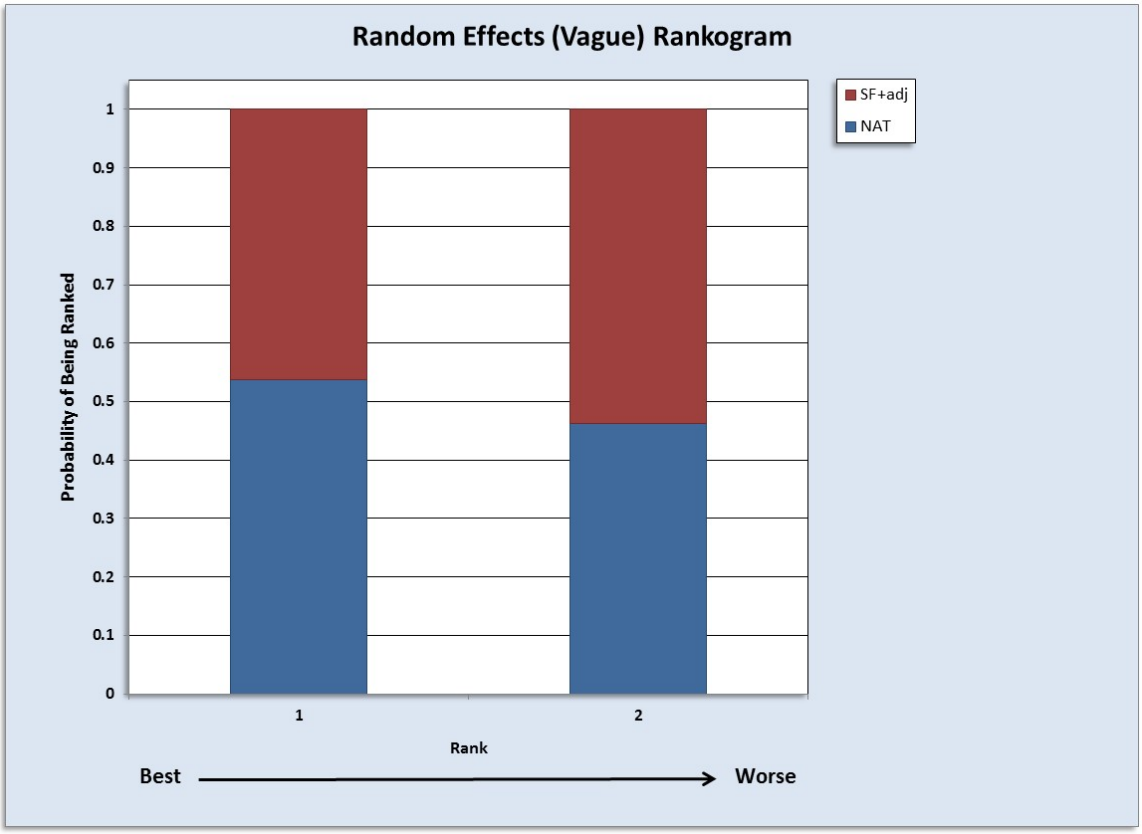


Figure Lxix: League table based on results of fixed effects and random effects (vague prior) models. Where odds ratio (O.R.) is greater than 1 treatment at top left is superior

NAT	
1.16 (0.69 – 1.94)	**SF+adj**

Figure Lxx: Rankogram summarising surface under the cumulative ranking (SUCRA).



5-year Survival Rates

Figure Lxxi: Results of fixed effects and random effects (vague prior) models

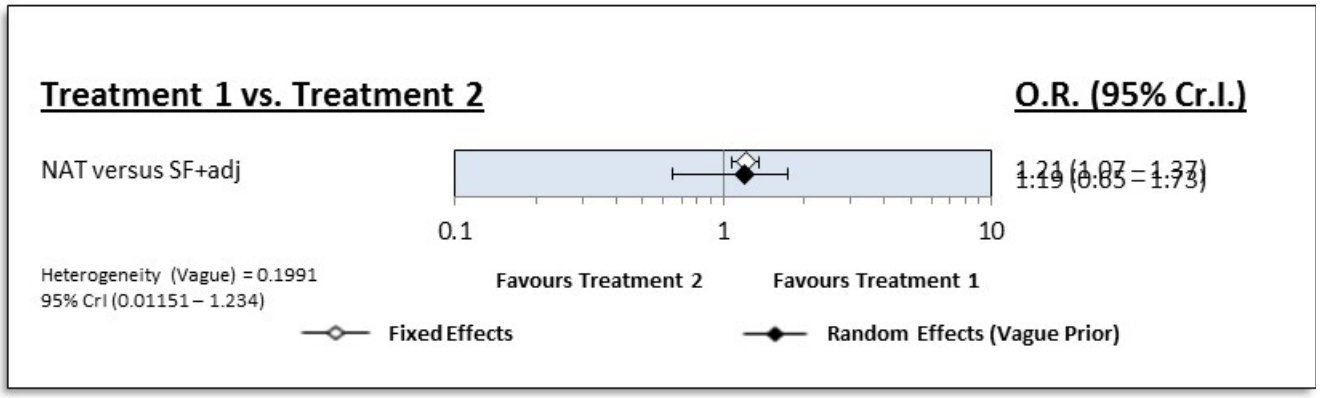
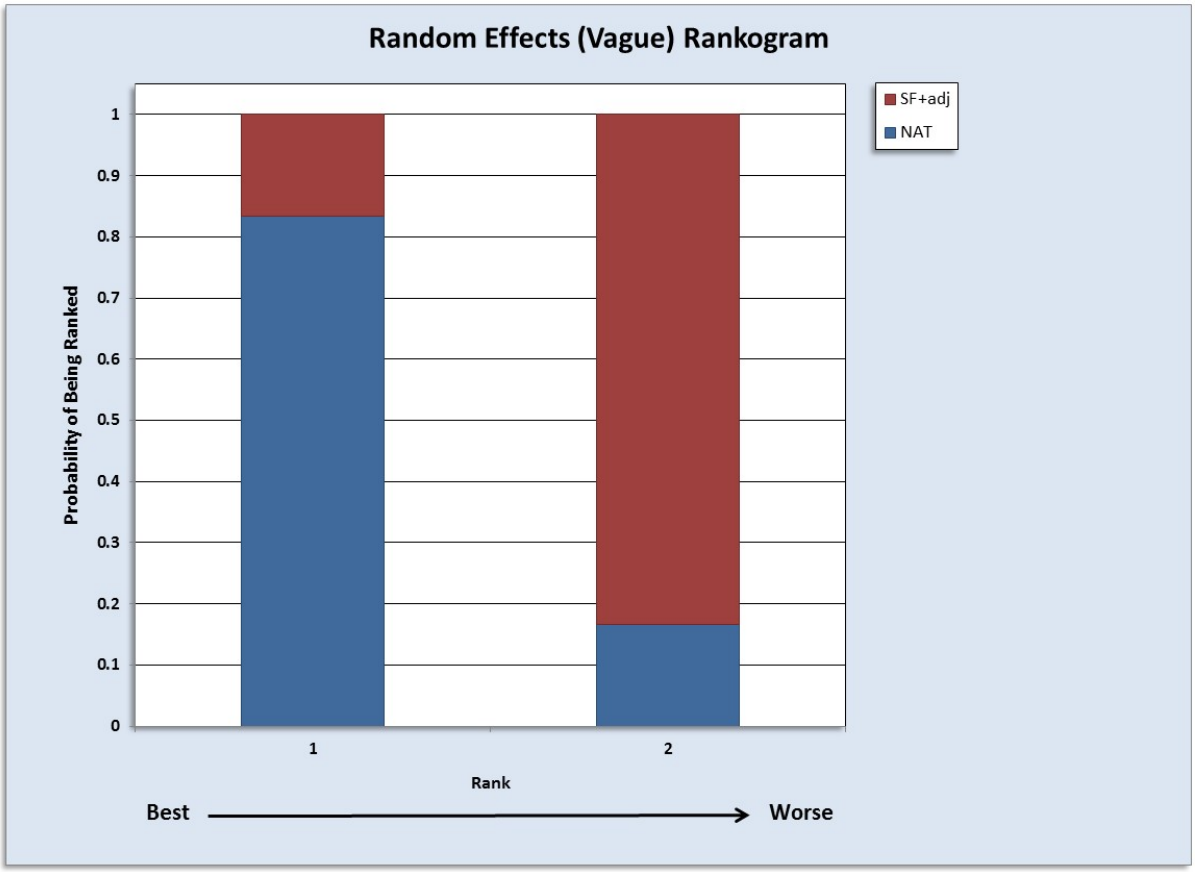


Figure Lxxii: League table based on results of fixed effects and random effects (vague prior) models. Where odds ratio (O.R.) is greater than 1 treatment at top left is superior

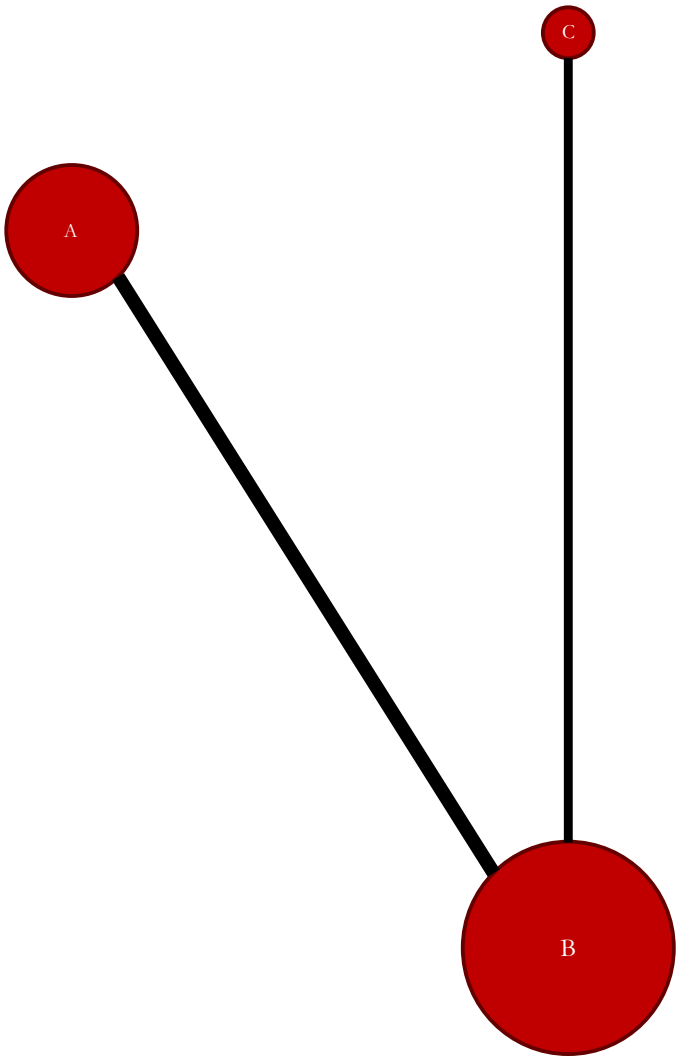
NAT	
1.21 (1.07 – 1.37)	**SF+adj**

Figure Lxxiii: Rankogram summarising surface under the cumulative ranking (SUCRA).



5-year survival: inclusion of RCTs

Figure Lxxiv: Bayesian Network Meta-analysis of Neoadjuvant therapy versus Upfront surgery plus adjuvant therapy versus surgery only

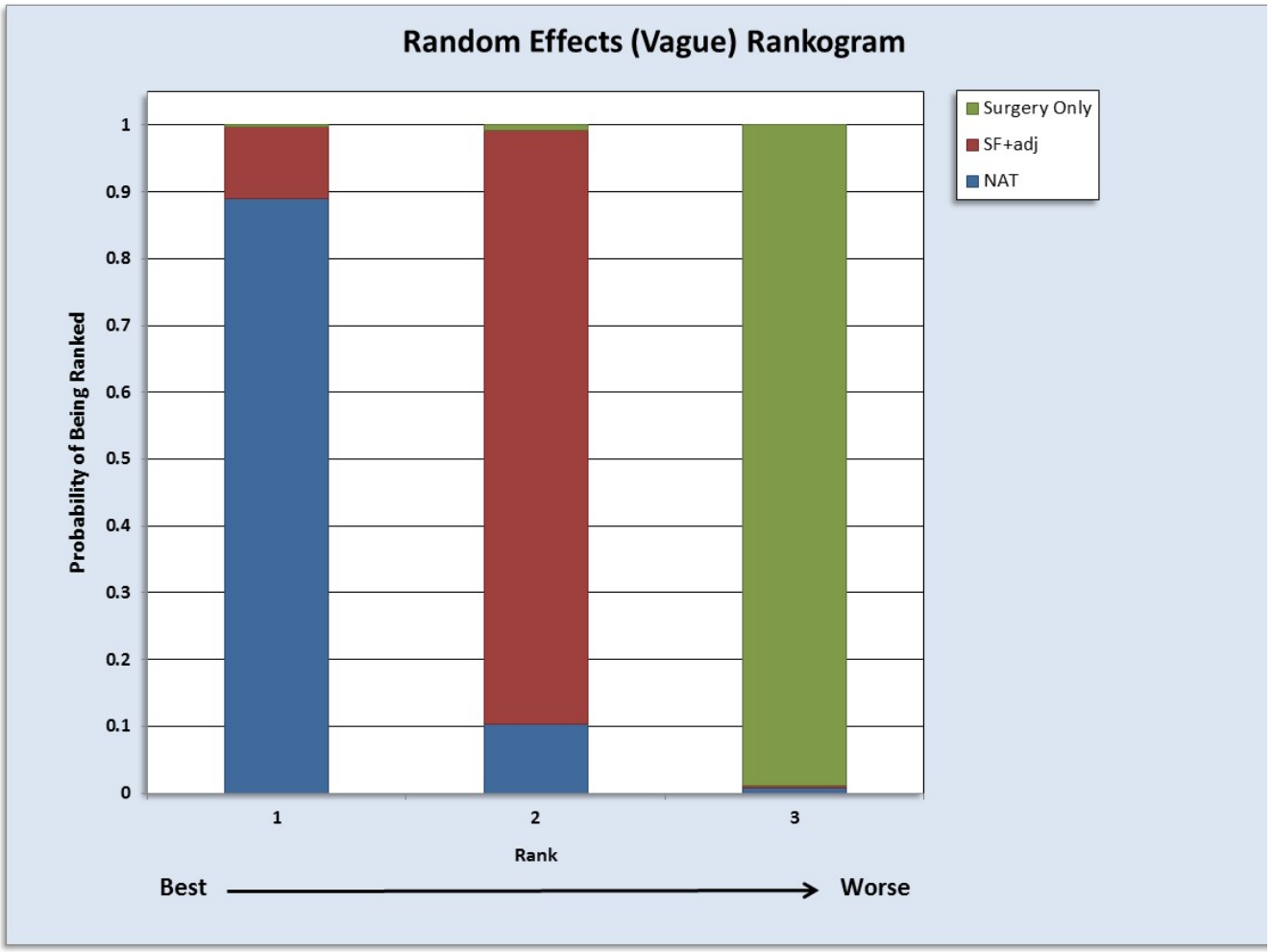


Drug	Abbreviation
NAT	A
SF+adj	B
Surgery Only	C

Figure Lxxv: League table based on results of fixed effects and random effects (vague prior) models. Where odds ratio (O.R.) is greater than 1 treatment at top left is superior

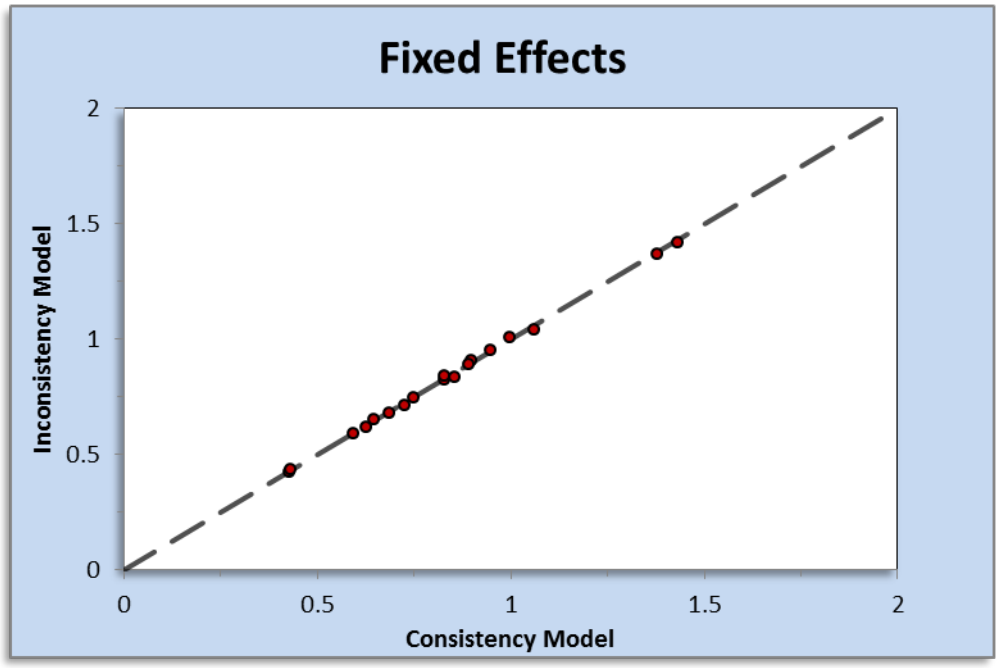
NAT		
1.21 (1.07 – 1.36)	**SF+adj**	
2.22 (1.50 – 3.30)	1.83 (1.27 – 2.67)	**Surgery Only**

Figure Lxxvi: Rankogram summarising surface under the cumulative ranking (SUCRA).

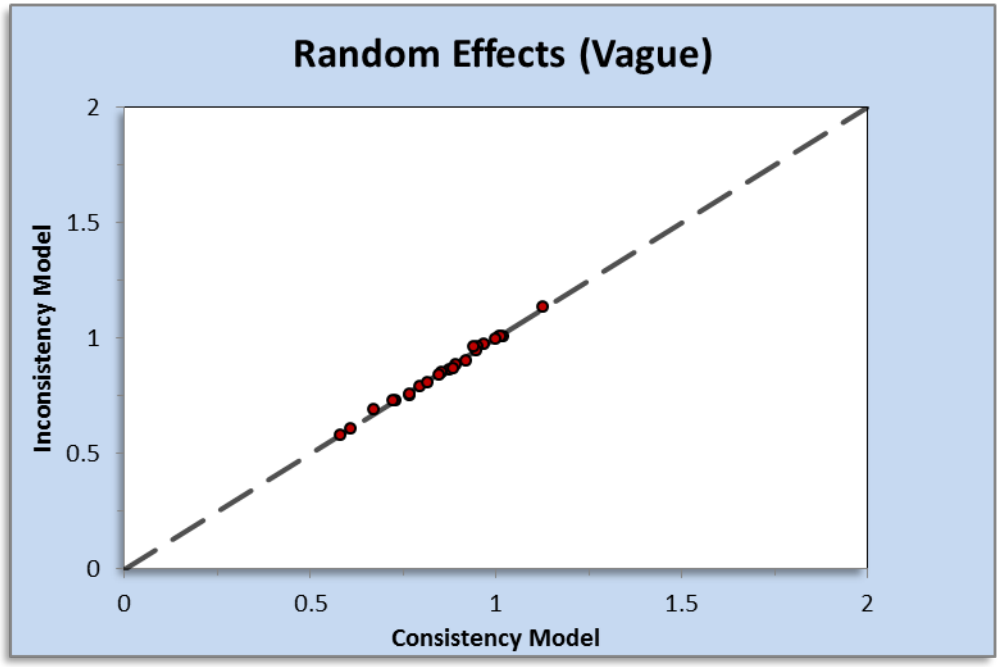


Convergence and Inconsistency

Figure Lxxvii: 1-year Survival Model

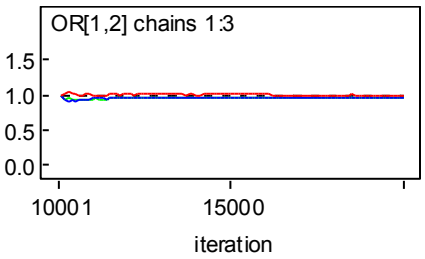


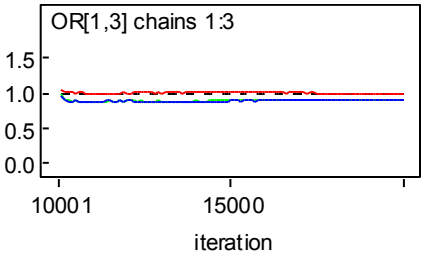
	Fixed Effects Odds Ratios		
	Treatment Comparisons	Consistency Model	InConsistency Model
OR[1,2]	SF+adj versus NAT	0.4077	0.4086
OR[1,3]	Surgery Only versus NAT	0.3946	N/A
OR[2,3]	Surgery Only versus SF+adj	0.9681	0.9636



	Random Effects Odds Ratios		
	Treatment Comparisons	Consistency Model	InConsistency Model
OR[1,2]	SF+adj versus NAT	0.4777	0.4774
OR[1,3]	Surgery Only versus NAT	0.4519	N/A
OR[2,3]	Surgery Only versus SF+adj	0.9471	0.9401

Gelman Rubin statistic





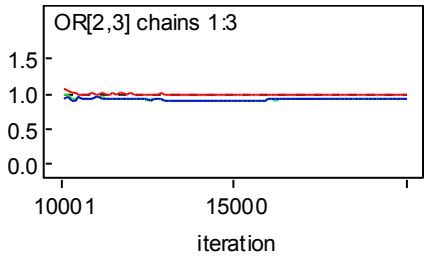
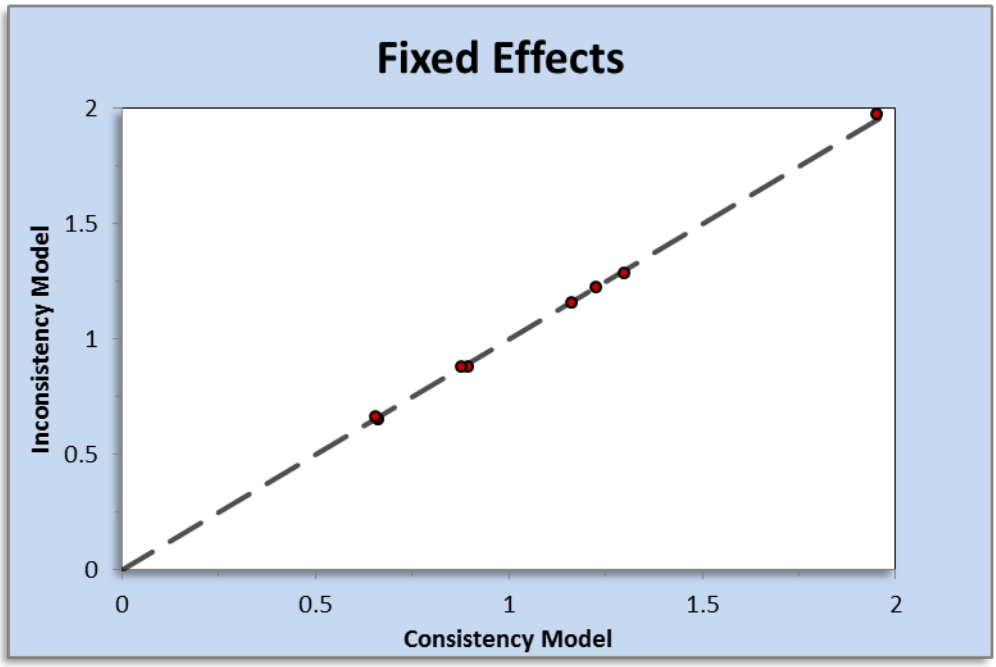
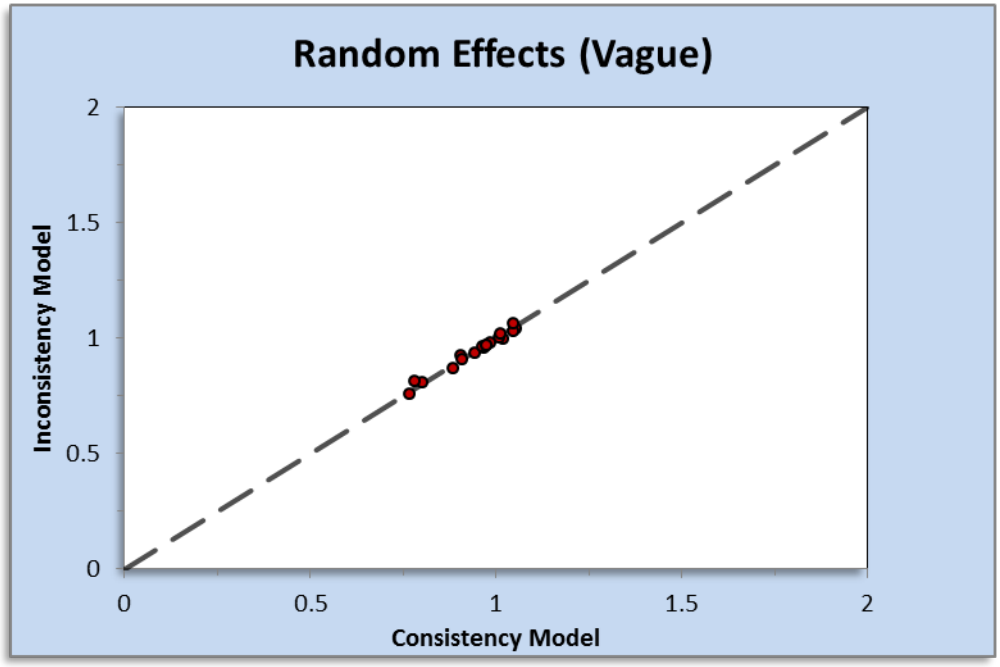


Figure Lxxviii: 2-year Survival Model

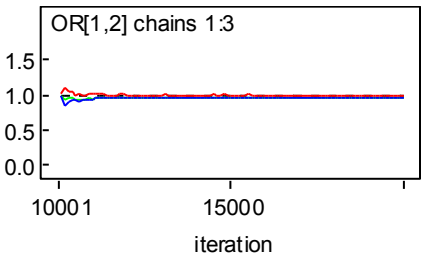


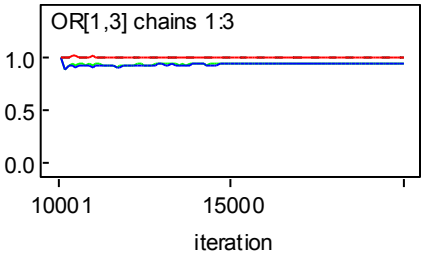
	Fixed Effects Odds Ratios		
	Treatment Comparisons	Consistency Model	InConsistency Model
OR[1,2]	Sfadj versus NAT	0.3036	0.3042
OR[1,3]	Surgery Only versus NAT	0.2356	N/A
OR[2,3]	Surgery Only versus Sfadj	0.7764	0.7756



	Random Effects Odds Ratios		
	Treatment Comparisons	Consistency Model	InConsistency Model
OR[1,2]	Sfadj versus NAT	0.604	0.6030
OR[1,3]	Surgery Only versus NAT	0.4596	N/A
OR[2,3]	Surgery Only versus Sfadj	0.7613	0.7612

Gelman Rubin statistic





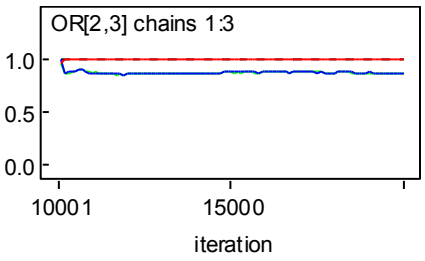
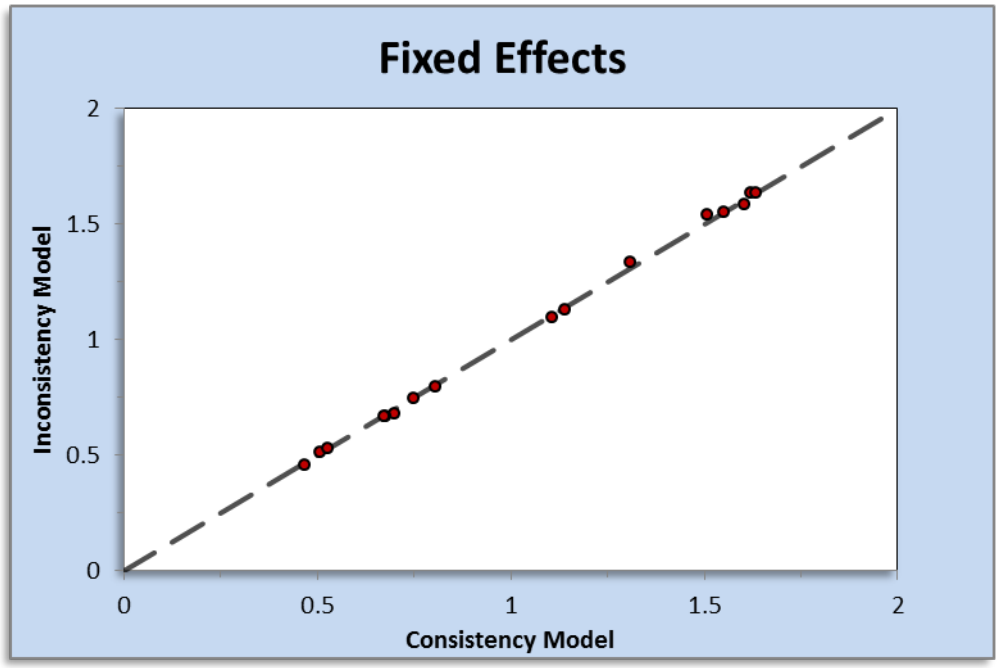
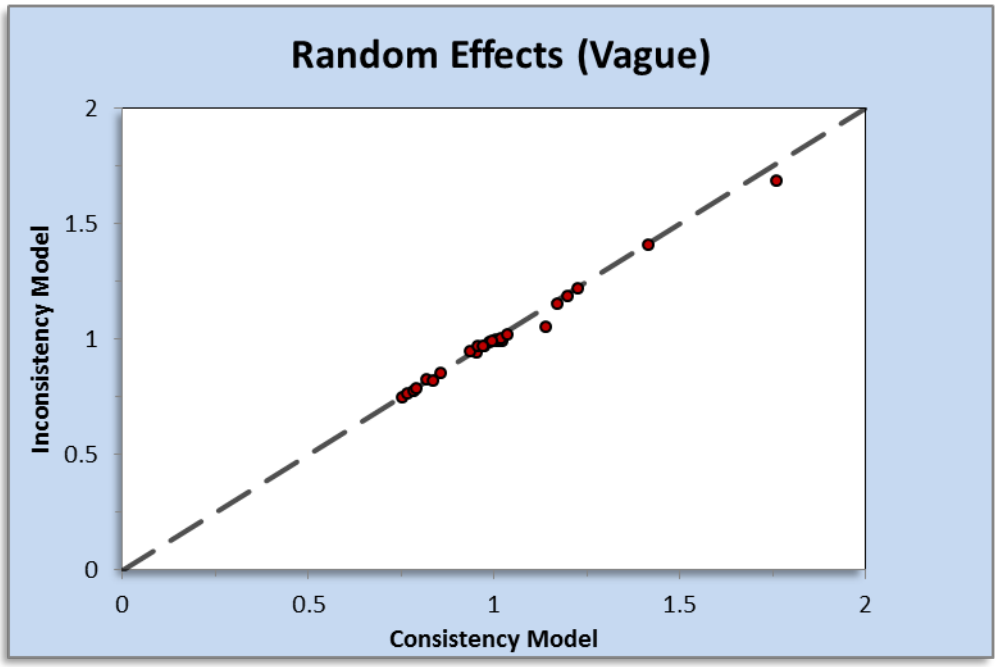


Figure Lxxix: 3-year Survival Model

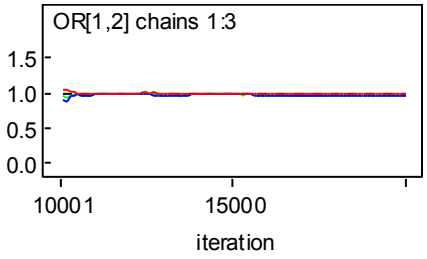


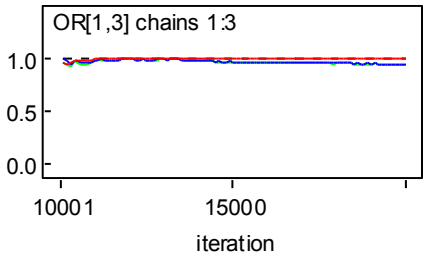
	Fixed Effects Odds Ratios		
	Treatment Comparisons	Consistency Model	InConsistency Model
OR[1,2]	Sfadj versus NAT	0.6204	0.6211
OR[1,3]	Surgery Only versus NAT	0.3085	N/A
OR[2,3]	Surgery Only versus Sfadj	0.4968	0.4980



	Random Effects Odds Ratios		
	Treatment Comparisons	Consistency Model	InConsistency Model
OR[1,2]	Sfadj versus NAT	0.6669	0.6673
OR[1,3]	Surgery Only versus NAT	0.3326	N/A
OR[2,3]	Surgery Only versus Sfadj	0.4973	0.4989

Gelman Rubin statistic


OR[1,2] chains 1:3
1.5
1.0
0.5
0.0
10001
15000
iteration


OR[1,3] chains 1:3
1.0
0.5
0.0
10001
15000
iteration

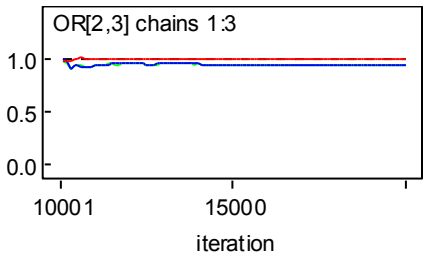
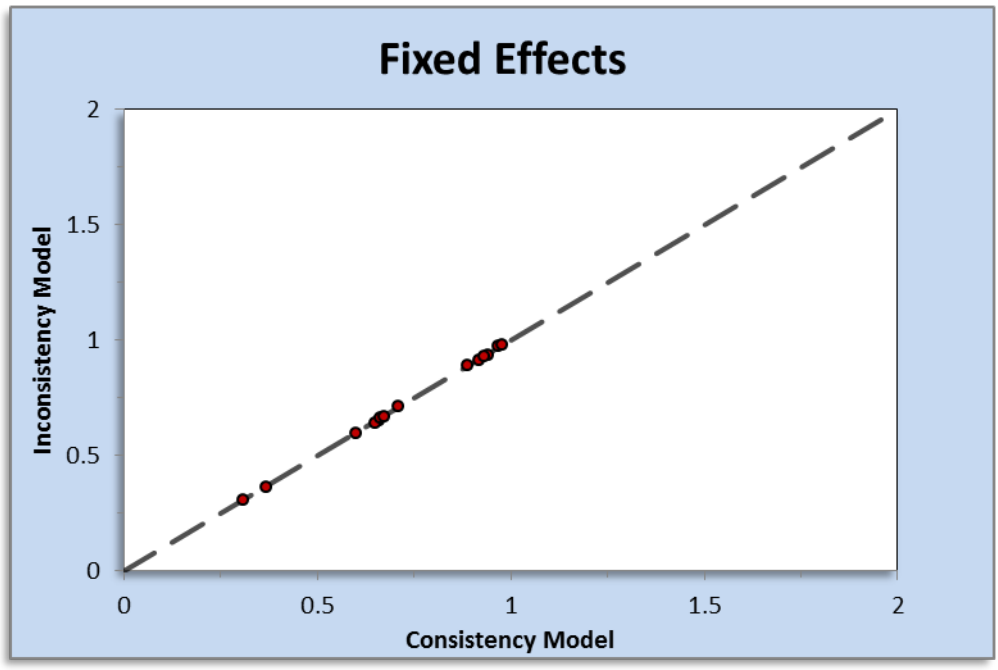
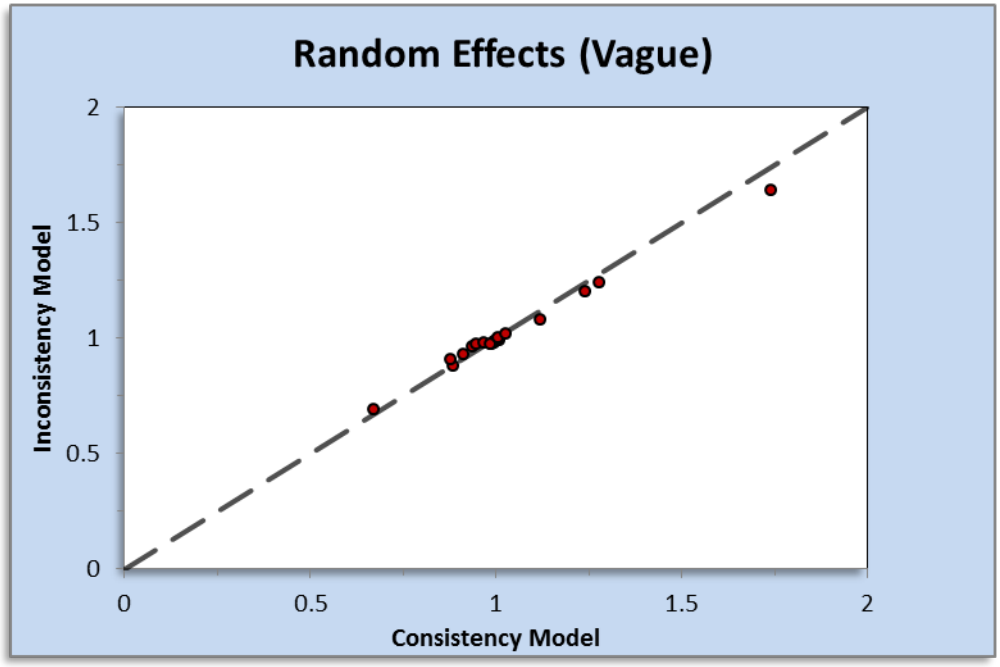

OR[2,3] chains 1:3
1.0
0.5
0.0
10001
15000
iteration

Figure Lxxx: 5-year Survival

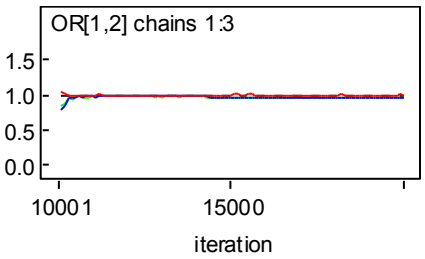


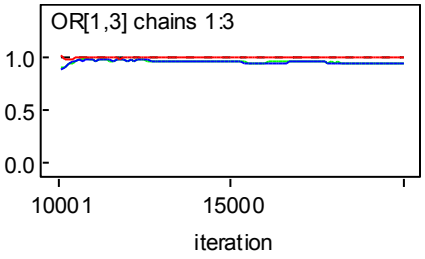
	Fixed Effects Odds Ratios		
	Treatment Comparisons	Consistency Model	InConsistency Model
OR[1,2]	SF+adj versus NAT	0.8275	0.8284
OR[1,3]	Surgery Only versus NAT	0.3935	N/A
OR[2,3]	Surgery Only versus SF+adj	0.4756	0.4752

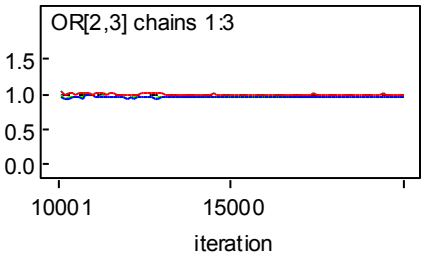


	Random Effects Odds Ratios		
	Treatment Comparisons	Consistency Model	InConsistency Model
OR[1,2]	SF+adj versus NAT	1.081	1.1117
OR[1,3]	Surgery Only versus NAT	0.3917	N/A
OR[2,3]	Surgery Only versus SF+adj	0.363	0.3620

Gelman Rubin statistic




OR[1,3] chains 1:3
1.0
0.5
0.0
10001
15000
iteration


OR[2,3] chains 1:3
1.5
1.0
0.5
0.0
10001
15000
iteration

Appendix M

Markov Decision Analysis

Figure Mi: PRISMA flowchart for phase II/III Neoadjuvant Therpy trials.

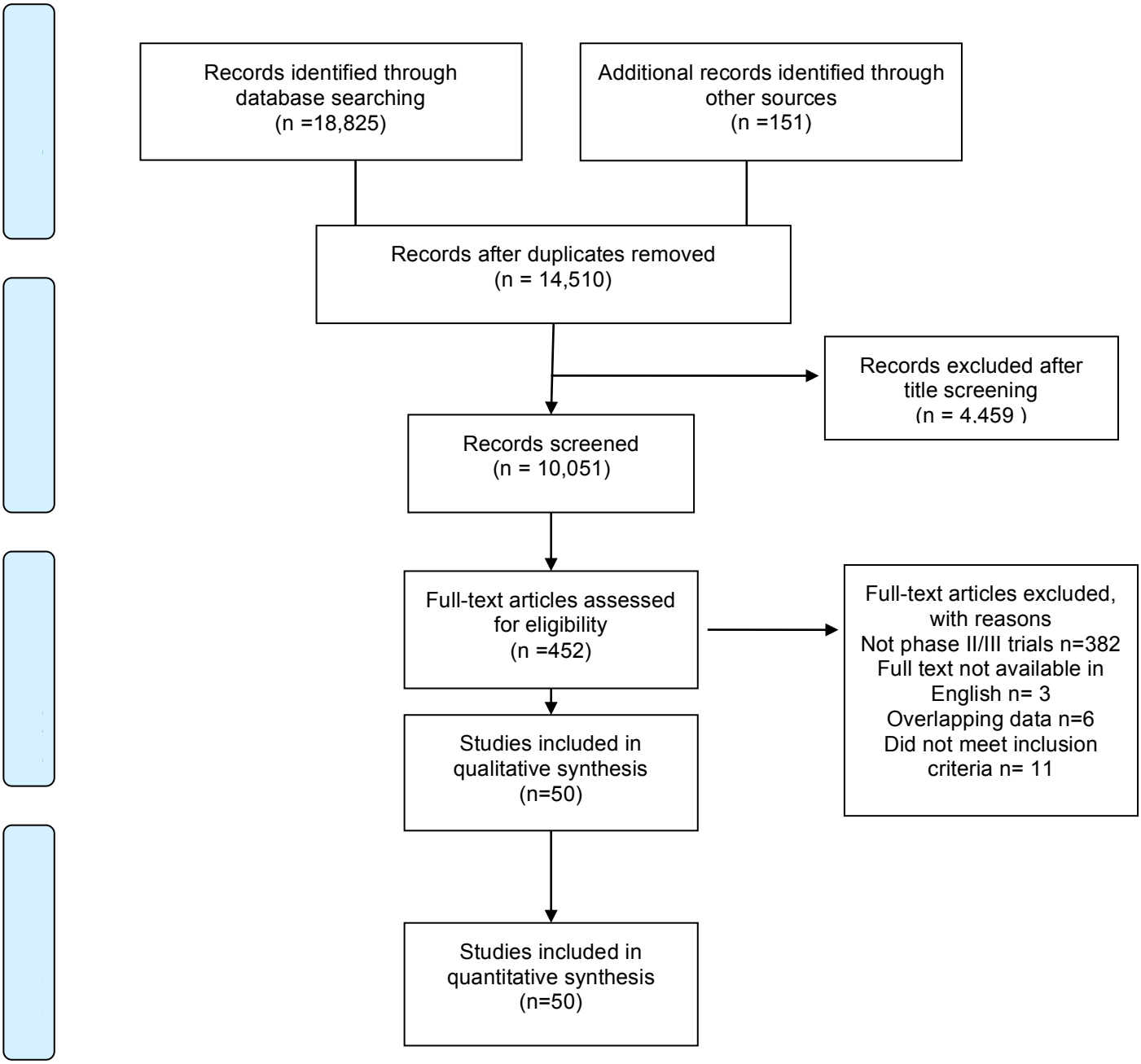


Figure Mii: PRISMA flowchart for RCT trials of Upfront Surgery approach.

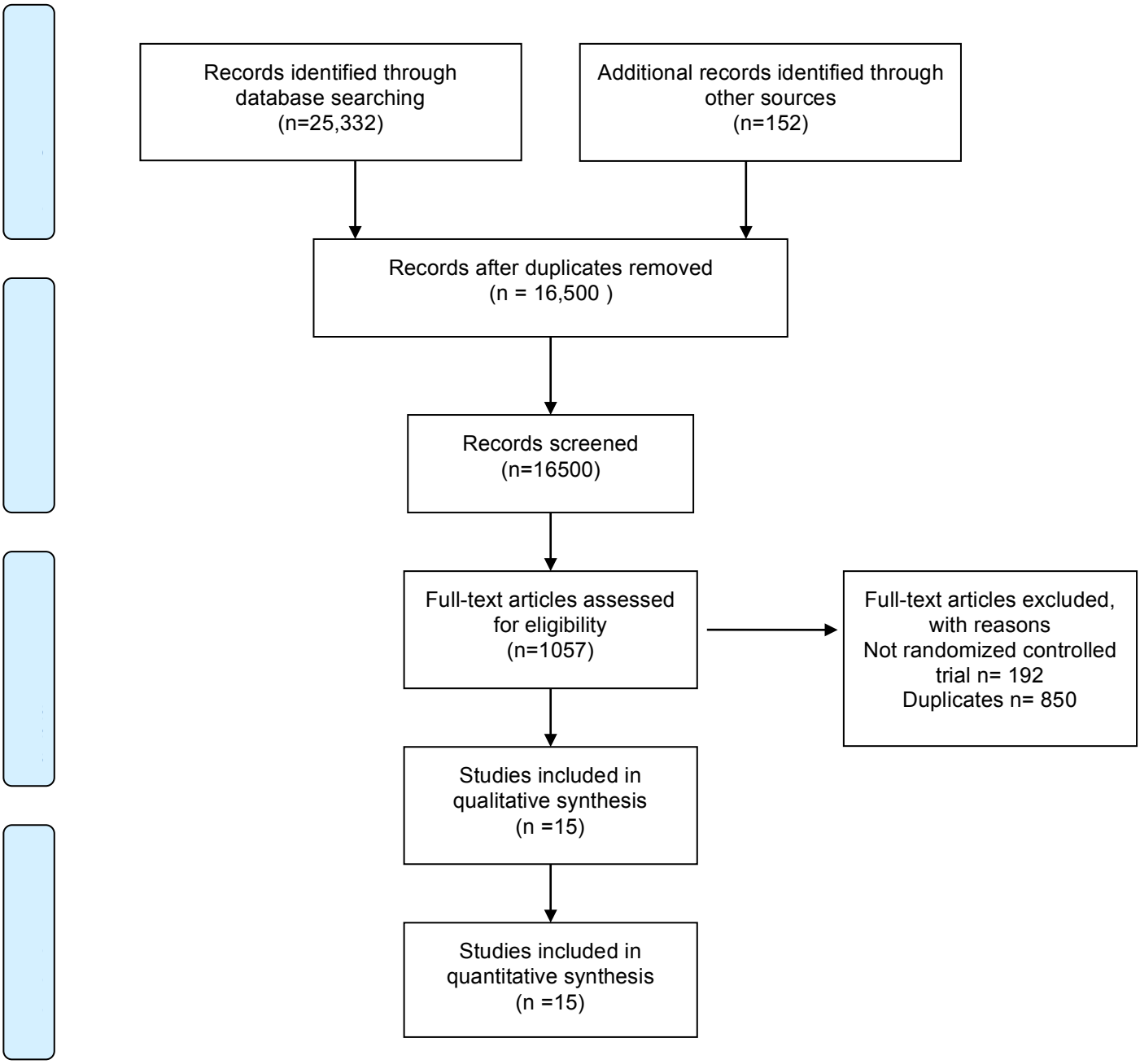


Figure Miii: PRISMA flow chart for non-RCT trials used in Upfront Sugrgery arm of the Markov model.

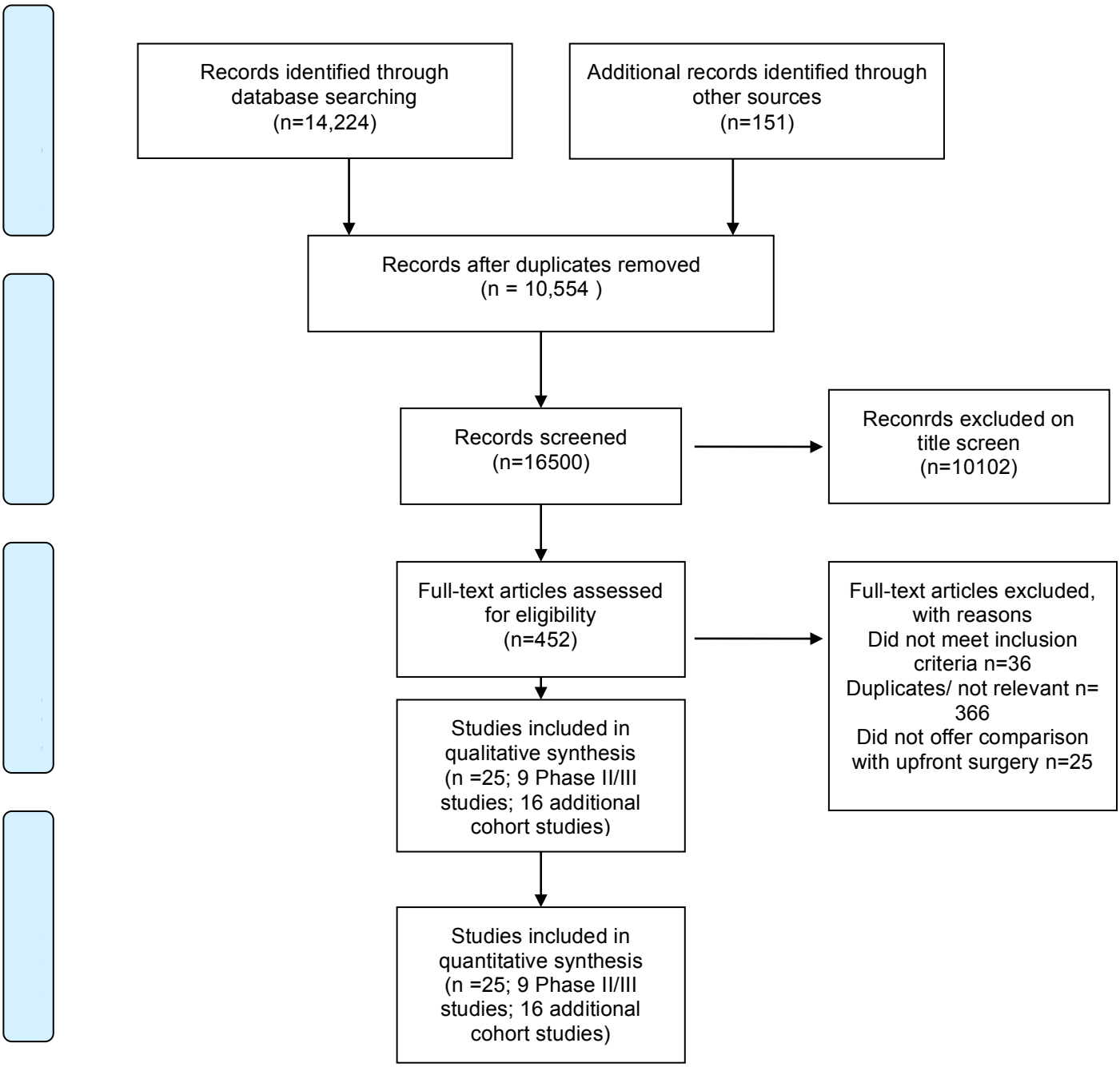


Table Mi: Summary of included trials.

Reference	Type of Study	Treatment Regime	N=	Disease Free Survival in months	Overall Survival in months
Al-Sukhun *et al.*, 2003	Prospective Phase II Trial	CRT + surgery	20		13.4
Cardenes *et al.*, 2011	Prospective Phase II Trial	CRT + surgery	28		10.3
Casadei *et al.*, 2015	Prospective Phase II Trial	CRT + surgery	18		28.3
Cetin *et al.*, 2013	Prospective Phase II Trial	CRT + surgery	11		
Chakraborty *et al.*, 2014	Prospective Phase II Trial	CRT + surgery	13	2.4	9.1
Crane *et al.*, 2011	Prospective Phase II Trial	CRT + surgery	69		19.2
Epelbaum *et al.*, 2002	Prospective Phase II Trial	CRT + surgery	20		8
Esnaola *et al.*, 2014	Prospective Phase II Trial	CRT + surgery	37	10.4	11.8
Evans *et al.*, 2008	Prospective Phase II Trial	CRT + surgery	86	15.4	22.7
Fiore *et al.*, 2017	Prospective Phase II Trial	CRT + surgery	34	20	19.2
Golcher *et al.*, 2008	Prospective Phase II Trial	CRT + surgery	121		
Golcher *et al.*, 2015	Prospective Phase II Trial	CRT + surgery	33	8.4	17.4
Heinrich *et al.*, 2008	Prospective Phase II Trial	CT + surgery	28	9.2	26.5
Herman *et al.*, 2015	Prospective Phase II Trial	CRT + surgery	49	7.8	13.9
Hong *et al.*, 2014	Prospective Phase II Trial	CRT + surgery	50	10.4	17.3
Jang *et al.*, 2018	Prospective Phase II Trial	CRT + surgery	27		21
Jensen *et al.*, 2014	Prospective Phase II Trial	CRT + surgery	23		11.5
Joensuu *et al.*, 2004	Prospective Phase II Trial	CRT + surgery	33	18	25
Kim *et al.*, 2013a	Prospective Phase II Trial	CRT + surgery	68		18.2
Landry *et al.*, 2010	Prospective Phase II Trial	CRT + surgery	21	14.2	19.4
Laurent *et al.*, 2009	Prospective Phase II Trial	CRT + surgery	22	8	17
Le Scodan *et al.*, 2009	Prospective Phase II Trial	CRT + surgery	41		9.4
Lee *et al.*, 2012	Prospective Phase II Trial	CT + surgery	43	10	16.6
Leone *et al.*, 2013	Prospective Phase II Trial	CRT + surgery	39	10.2	16.7
Lin *et al.*, 2005	Prospective Phase II Trial	CRT + surgery	42		10.3
Lind *et al.*, 2008	Prospective	CRT +	17		19

	Phase II Trial	surgery			
Magnin *et al.*, 2003	Prospective Phase II Trial	CRT + surgery	32		16
Magnino *et al.*, 2005	Prospective Phase II Trial	CRT + surgery	23		14
Marti *et al.*, 2008	Prospective Phase II Trial	CRT + surgery	26	7	13
Mattiucci *et al.*, 2010	Prospective Phase II Trial	CRT + surgery	40		15.5
Massucco *et al.*, 2006	Prospective Phase II Trial	CRT + surgery	28	10	15.4
Maximous *et al.*, 2009	Prospective Phase II Trial	CRT + surgery	25		12
Mornex *et al.*, 2006	Prospective Phase II Trial	CRT + surgery	41		9.4
Motoi *et al.*, 2013	Prospective Phase II Trial	CT+ surgery	35		19.7
Moutardier *et al.*, 2002	Prospective Phase II Trial	CRT + surgery	19		20
O'Reilly *et al.*, 2014	Prospective Phase II Trial	CT + surgery	38		27.2
Palmer *et al.*, 2007	Prospective Phase II Trial	CT + surgery	50		13.6
Pipas *et al.*, 2012	Prospective Phase II Trial	CRT + surgery	37		17.3
Pister *et al.*, 2002	Prospective Phase II Trial	CRT + surgery	37		12
Sahora *et al.*, 2011	Prospective Phase II Trial	CT + surgery	25		16
Satoi *et al.*, 2009	Prospective Phase II Trial	CRT + surgery	35		24.5
Sherman *et al.*, 2015	Prospective Phase II Trial	CRT + surgery v CT + surgery	45	34	29/42
Small *et al.*, 2011	Prospective Phase II Trial	CRT + surgery	29	9.9	11.8
Talamonti *et al.*, 2006	Prospective Phase II Trial	CRT + surgery	20		
Tinchon *et al.*, 2013	Prospective Phase II Trial	CT + surgery	12		
Turrini *et al.*, 2009	Prospective Phase II Trial	CRT + surgery	34		15.5
Van Buren *et al.*, 2013	Prospective Phase II Trial	CRT + surgery	59	6.6	16.8
Varadhachary *et al.*, 2008	Prospective Phase II Trial	CRT + surgery	90	13.2	17.4
Vento *et al.*, 2007	Prospective Phase II Trial	CRT + surgery	22		30.2
Wilkowski *et al.*, 2009	Prospective Phase II Trial	CRT + surgery	93	5.6	9.3
Regine *et al.*, 2011	RCT	Surgery + CRT	230		17.1
Neoptolemos *et al.*, 2010	RCT	Surgery +CT	551	14.1	23
VanLaethem *et al.*, 2010	RCT	Surgery +CRT	45	11.8	24.3

Schmidt *et al.*, 2012	RCT	Surgery +CRT	53	15.2	26.5
Reni *et al.*, 2012	RCT	Surgery +CRT	51	11.7	26.2
Yoshitomi *et al.*, 2008	RCT	Surgery +CT	49	12	29.8
Shimoda *et al.*, 2015	RCT	Surgery +CT	29	14.6	21.5
Uesaka *et al.*, 2016	RCT	Surgery +CT	187	22.9	46.5
Neoptolemos *et al.*, 2004	RCT	Surgery +CRT	145	10.7	15.9
Ueno *et al.*, 2009	RCT	Surgery +CT	58	11.4	22.3
Oettle *et al.*, 2013	RCT	Surgery +CT	179	13.4	22.8
Kosuge *et al.*, 2006	RCT	Surgery +CT	45	8.6	12.5
Smeenk *et al.*, 2007	RCT	Surgery +CRT	110	18	21.6
Morak *et al.*, 2008	RCT	Surgery +CR	59	12	19
Neoptolemo *et al.*, 2017	RCT	Surgery + CT	366		25.5
Regine *et al.*, 2011	RCT	Surgery +CRT	221		20.5
Neoptolemos *et al.*, 2010	RCT	Surgery +CT	537	14.3	23.6
VanLaethem *et al.*, 2010	RCT	Surgery +CT	45	10.9	24.4
Schmidt *et al.*, 2012	RCT	Surgery +CT	57	11.5	28.5
Reni *et al.*, 2012	RCT	Surgery + CT	49	15.2	31.6
Yoshitomi *et al.*, 2008	RCT	Surgery + CT	50	2.3	21.2
Shimoda *et al.*, 2015	RCT	Surgery +CT	28	10.5	18
Uesaka *et al.*, 2016	RCT	Surgery + CT	190	11.3	25.5
Neoptolemos *et al.*, 2004	RCT	Surgery +CT	147	15.3	20.1
Ueno *et al.*, 2009	RCT	Surgery Only	60	5	18.4
Oettle *et al.*, 2013	RCT	Surgery Only	175	6.7	20.2
Kosuge *et al.*, 2006	RCT	Surgery Only	44	10.2	15.8
Smeenk *et al.*, 2007	RCT	Surgery Only	108	14.4	19.2
Morak *et al.*, 2008	RCT	Surgery Only	61	7	18
Neoptolemo *et al.*, 2017	RCT	Surgery + CT	364		28
Al-Sukhun *et al.*, 2003	Prospective Phase II Trial	Surgery + adjuvant therapy	21		18.1

Casadei *et al.*, 2015	Prospective Phase II Trial	Surgery + adjuvant therapy	20		27.5
Golcher *et al.*, 2008	Prospective Phase II Trial	Surgery + adjuvant therapy	58		21
Golcher *et al.*, 2015	Prospective Phase II Trial	Surgery + adjuvant therapy	33	8.7	14.4
Lind *et al.*, 2008	Prospective Phase II Trial	Surgery + adjuvant therapy	35		11
Massucco *et al.*, 2006	Prospective Phase II Trial	Surgery + adjuvant therapy	44	8	14
Satoi *et al.*, 2009	Prospective Phase II Trial	Surgery + adjuvant therapy	41		18.5
Vento *et al.*, 2007	Prospective Phase II Trial	Surgery + adjuvant therapy	25		35.9
Jang *et al.*, 2018	Prospective Phase II Trial	Surgery + adjuvant therapy	23		12
DeGus *et al.*, 2017a	Retrospective Cohort	Surgery + adjuvant therapy	6840		24.2
Mellon *et al.*, 2016	Retrospective Cohort	Surgery + adjuvant therapy	241		22.1
Nurmi *et al.*, 2018	Retrospective Cohort	Surgery + adjuvant therapy	150	13	26
Shubert *et al.*, 2016	Retrospective Cohort	Surgery + adjuvant therapy	216		13
Artinya *et al.*, 2011	Retrospective Cohort	Surgery + adjuvant therapy	419		19
Ielpo *et al.*, 2016	Prospective Cohort	Surgery + adjuvant therapy	36		22.1
Roland *et al.*, 2015	Prospective Cohort	Surgery + adjuvant therapy	85		
DeGus *et al.*, 2017b	Retrospective Cohort	Surgery + adjuvant therapy	11316		Resectable: 24.5 Borderline: 20.0 Locally advanced: 15.5
Mokdad *et al.*, 2017	Retrospective Cohort	Surgery + adjuvant therapy	6015		21
Chen *et al.*, 2017	Retrospective Cohort	Surgery + adjuvant therapy	98		17
Tzeng *et al.*,	Prospective	Surgery +	52		25.3

2014	Cohort	adjuvant therapy			
Fujii *et al*., 2015	Prospective Cohort	Surgery + adjuvant therapy	71		13.1
Fujii *et al*., 2017	Prospective Cohort	Surgery + adjuvant therapy	416		Resectable: 23.5 Borderline: 20.1
Papalezova *et al*., 2012	Retrospective Cohort	Surgery + adjuvant therapy	92		13
Hirono *et al*., 2016	Prospective Cohort	Surgery + adjuvant therapy	124		13.7
Murakami *et al.* 2017	Retrospective Cohort	Surgery + adjuvant therapy	25		11.6

CRT= chemoradiotherapy CT= chemotherapy

Markov Decision Analysis: Resectable Only Casaes

Figure Miv: PRISMA Flowchart for Phase II/III Trials of Neoadjuvant Therapy for Resectable Pancreatic Cancer

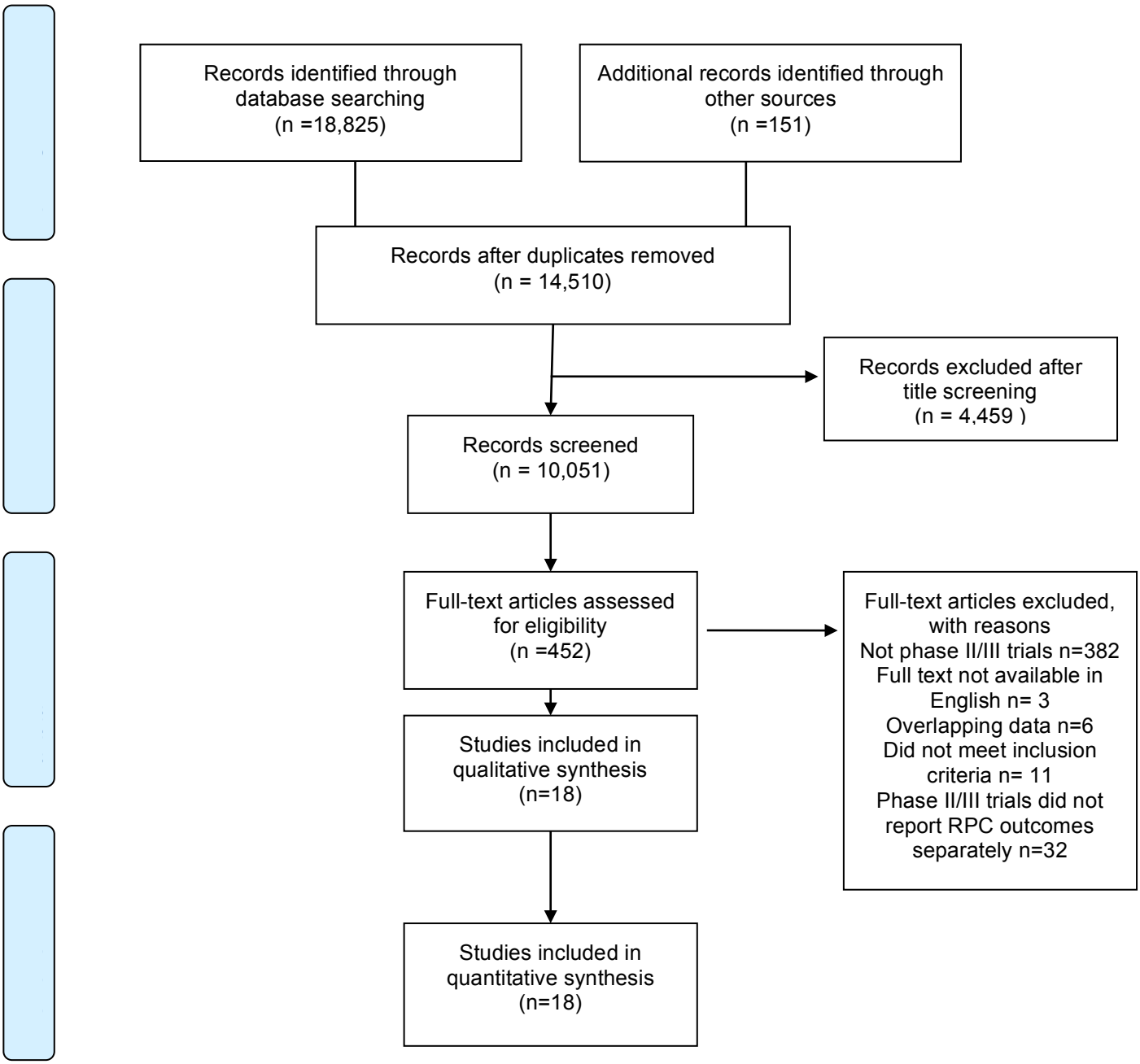


Table Mii: Summary of included trials in neoadjuvant arm of Markov model

Reference	**No.**	**Single or Multicentre trial**	**Randomization**	**NAT Treatment Regime: CRT= chemoradiotherapy; CT= chemotherapy**	**ROBINS-I risk of bias assessment[22]**
Evans *et al.*, 2008	86	Single	No	CRT: 7 weekly intravenous infusions of gemcitabine 400 mg/m^2 plus radiation therapy (30 Gy in 10 fractions over 2 weeks).	Moderate
Golcher *et al.*, 2015	33	Multicentre	Yes	CRT: 300 mg/m^2 gemcitabine and 30 mg/m^2 cisplatin on days 1, 8, 22, and 29 of radiotherapy (1.8 Gy to 55.8 Gy (tumor) or 50.4 Gy (regional lymph nodes), planning target volume ≤ 800 ml). Adjuvant chemotherapy also given as per CONKO-001 study protocol.	Low
Heinrich *et al.*, 2008	28	Single	No	CT: four biweekly cycles of gemcitabine 1,000 mg/m^2 and cisplatin 50 mg/m^2	Moderate
Hong *et al.*, 2014	50	Multicentre	No	CRT: Proton beam therapy 240-MeV protons generated from a cyclotron delivered using 3D passively scattered protons. Most commonly, 3 fields were used, with 2 fields being treated per day. Capecitabine (1650	Moderate

				mg/m^2 divided twice daily) given Monday to Friday for 2 weeks for each dose level. Adjuvant gemcitabine chemotherapy for 6 months starting 4 to 10 weeks post surgery.	
Joensuu *et al.,* 2004	33	Single	No	CRT: Gemcitabine intravenous infusion twice weekly was tested at 3 dose levels: 20, 50, and 100 mg/m^2. Radiation dose 50.4 Gy in 28 fractions.	Moderate
Kim *et al.,* 2013a	68	Multicentre	No	CRT: two 28-day cycles of gemcitabine (1 g/m^2 on days 1, 8, and 15) and oxaliplatin (85 mg/m^2 on days 1 and 15) with raditaion during cycle 1 (30 Gy in 2-Gy fractions). Adjuvant chemotherapy within 12 weeks of surgery; further regime details not provided.	Low/Moderate
LeScodan *et al.,* 2009	41	Multicentre	No	CRT: concurrent radiotherapy (50 Gy within 5 weeks) and chemotherapy 5-fluorouracil (300 mg/m^2/day, 5 days/week, weeks 1-5) and cisplatin (20 mg/m^2/day, days 1-5 and 29-33).	Low/Moderate
Maximous *et al.,* 2009	25	Single	No	CRT: 54 Gy in 30 fractions over 6 weeks. Gemcitabine intravenous infusion (300 mg/m^2) given prior to radiation on a weekly basis.	Moderate

Mornex *et al.,* 2006	41	Multicentre	No	CRT: concurrent radiotherapy (50 Gy within 5 weeks), and chemotherapy: 5-fluorouracil (300 mg/m^2/day, 5 days/week, 5 consecutive weeks) and cisplatin (20 mg/m^2/day, Days 1-5 and 29-33)	Low/Moderate
O'Reilly *et al.,* 2014	38	Single	No	CT: four cycles of intravenous infusion gemcitabine 1000mg/m^{2}and oxaliplatin 80 mg/m^2, every 2 weeks. Adjuvant gemcitabine intravenous infusion: 5 cycles (1000 mg/m^2 day 1, 8, 15 every 4 weeks).	Moderate
Palmer *et al.,* 2007	50	Single	Yes	CT: 24 patients were randomized to gemcitabine (1000 mg/m^2) every 7 days for 43 days; 26 patients were randomized to gemcitabine (1000 mg/m^2) and cisplatin (25 mg/m^2)	Low
Pipas *et al., 2012*	37	Single	No	CRT: cetuximab load at 400 mg/m^2 followed by cetuximab 250 mg/m^2 weekly and gemcitabine 50 mg/m^2 twice-weekly given concurrently with IMRT to 54 Gy.	Moderate
Pister *et al.,* 2002	37	Single	No	CRT: 30 Gy external-beam radiation therapy and concomitant weekly 3-hour infusions of paclitaxel (60 mg/m^2)	Moderate

Sherman *et al.,* 2015	45	Single	No	CRT *versus* CT: CRT: IMRT 5040 cGy along with capecitabine 1000 mg twice daily for 5 days and gemcitabine 750 mg/m^2 on day 5 of weeks 1, 2, 4, and 5 of radiotherapy. CT: 6-cycles: Capecitabine (1500 mg/m^2, days 1-14) plus gemcitabine (750 mg/m^2, days 4 and 11) plus Docetaxel (30 mg/m^2, days 4 and 11)	Moderate
Talamonti *et al.,* 2006	20	Multicentre	No	CRT: three cycles of gemcitabine (1000 mg/m^2 intravenously), with radiation during the second cycle (36 Gy in daily 2.4-Gy fractions)	Moderate
Turrini *et al.,* 2009	34	Multicentre	No	CRT: radiation therapy (45 Gy) with continuous infusion of 5-fluorouracil accompanied by a cisplatin bolus.	Moderate
Varadhachary *et al.,* 2008	90	Single	No	CT: gemcitabine (750 mg/m^2) and cisplatin (30 mg/m^2) every 2 weeks for 4 doses. CRT: 4 weekly infusions of gemcitabine (400 mg/m^2) combined with radiation therapy (30 Gy in 10 fractions administered over 2 weeks) delivered 5 days per week.	Moderate
Vento *et al.,* 2007	22	Single	No	CRT: gemcitabine intravenous infusion twice weekly before irradiation at three dose levels, which were 20, 50 and 100 mg/m^2	Moderate

				for an average of 10 cycles. Tumour radiation dose 50.4 Gy given in 28 fractions of 1.8 Gy per day, five days per week.	

*NAT= Neoadjuvant therapy

Table Miii: Summary of included Randomised Controlled Trials in Upfront Surgery pathway

Reference	**Adjuvant Regime: CRT = chemoradiotherpy; CT= chemotherapy**	**No. in Upfront Surgery and Adjuvant Arm**	**Comparison Arm**	**No. in Comparison Arm**
Regine *et al.*, 2011	CRT	230	Adjuvant CRT	221
Neoptolemos *et al.,* 2010	CT	551	Adjuvant CT	537
Van Laethem *et al.,* 2010	CRT	45	Adjuvant CT	45
Schmidt *et al.,* 2012	CRT	53	Adjuvant CT	57
Reni *et al.,* 2012	CRT	51	Adjuvant CT	49
Yoshitomi *et al.,* 2008	CT	49	Adjuvant CT	50
Shimoda *et al.,* 2015	CT	29	Adjuvant CT	28
Uesaka *et al.,* 2016	CT	187	Adjuvant CT	190
Neoptolemos *et al.,* 2004	CRT	145	Adjuvant CT	147
Ueno *et al.,* 2009	CT	58	Surgery Only	60
Oettle *et al.,* 2013	CT	179	Surgery Only	175
Kosuge *et al.,* 2006	CT	45	Surgery Only	44
Smeenk *et al.,* 2007	CRT	110	Surgery Only	108
Morak *et al.,* 2008	CR	59	Surgery Only	61
Neoptolemo *et al.,* 2017	CT	366	Adjuvant CT	364

Table Miv: Summary of included cohort studies in Upfront Surgery/ Surgery First (SF) pathway

Reference	Study Type	Multi or Single Centre	Randomization	NAT cohort	No. SF cohort	ROBINS-I risk of bias assessment[21]
Al-Sukhun *et al*. 2003	Phase II Trial	Single	No	20	21	Moderate
Casadei *et al.,* 2015	Phase II Trial	Single	Yes	18	20	Moderate
Golcher *et al.,* 2008	Phase II Trial	Single	No	121	58	Moderate
Golcher *et al.,* 2015	Phase II Trial	Multiple	Yes	33	33	Low
Lind *et al.,* 2008	Phase II Trial	Single	No	17	35	Moderate
Massucco *et al.,* 2006	Phase II Trial	Single	No	28	44	Moderate
Satoi *et al.,* 2009	Phase II Trial	Single	No	35	41	Moderate
Vento *et al.,* 2007	Phase II Trial	Single	No	22	25	Moderate
Jang et al., 2018	Phase II/III Trial	Multiple	Yes	27	23	Low
DeGus *et al.* 2017a	Retrospective	Multicentre	No	1077	6840	Serious
Mellon *et al.,* 2016	Retrospective	Single centre	No	159	241	Moderate/Serious
Nurmi *et al.,* 2018	Retrospective	Single centre	No	75	150	Serious
Shubert *et al.,* 2016	Retrospective	Multicentre	No	377	216	Moderate/Serious
Artinya *et al.,* 2011	Retrospective	Multicentre	No	39	419	Serious
Ielpo *et al.,* 2016	Prospective	Single centre	No	45	36	Serious
Roland *et al.,* 2015	Prospective	Single centre	No	222	85	Moderate/Serious
DeGus *et al.,* 2017b	Retrospective	Multicentre	No	1541	11316	Moderate/Serious
Mokdad *et al.,* 2017	Retrospective	Multicentre	No	2005	6015	Moderate/Serious
Chen *et al.,* 2017b	Retrospective	Multicentre	No	98	98	Moderate/Serious
Tzeng *et al.*, 2014	Prospective	Single centre	No	115	52	Moderate
Fujii *et al.,*	Prospective	Single	No	21	71	Moderate/Serious

2015		centre				
Fujii *et al.,* 2017	Prospective	Single centre	No	88	416	Moderate/Serious
Papalezova *et al.,* 2012	Retrospective	Single centre	No	144	92	Moderate/Serious
Hirono *et al.,* 2016	Prospective	Single centre	No	46	124	Moderate/Serious
Murakami *et al.,* 2017	Retrospective	Single centre	No	52	25	Serious

Appendix N

Discrete Event Simulation (DES) Modeling

Figure Ni: PRISMA Flow chart for neoadjuvant papers included in DES model

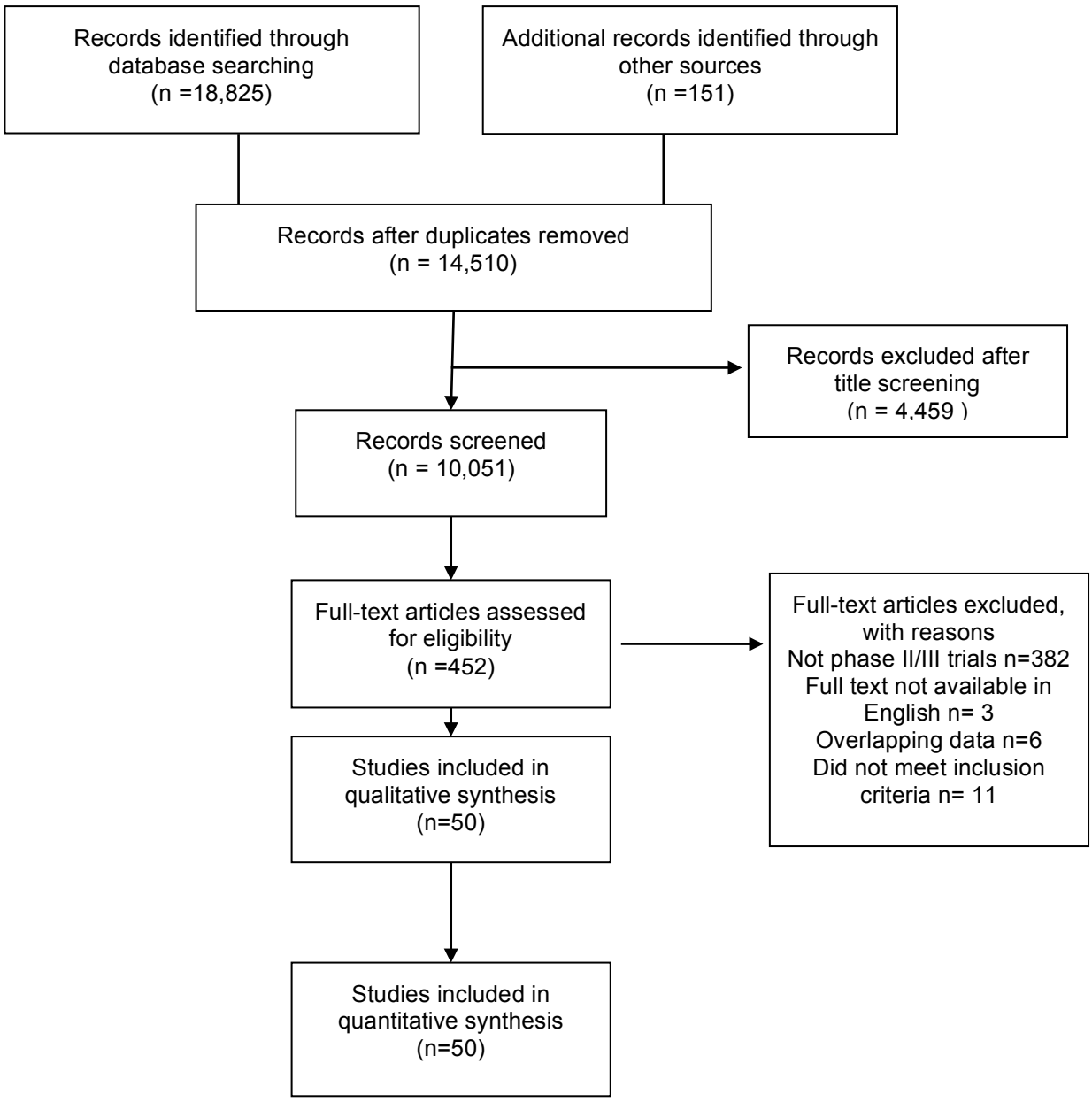


Figure Nii: PRISMA Flow chart for randomised controlled trials included in surgery-first pathway of DES model

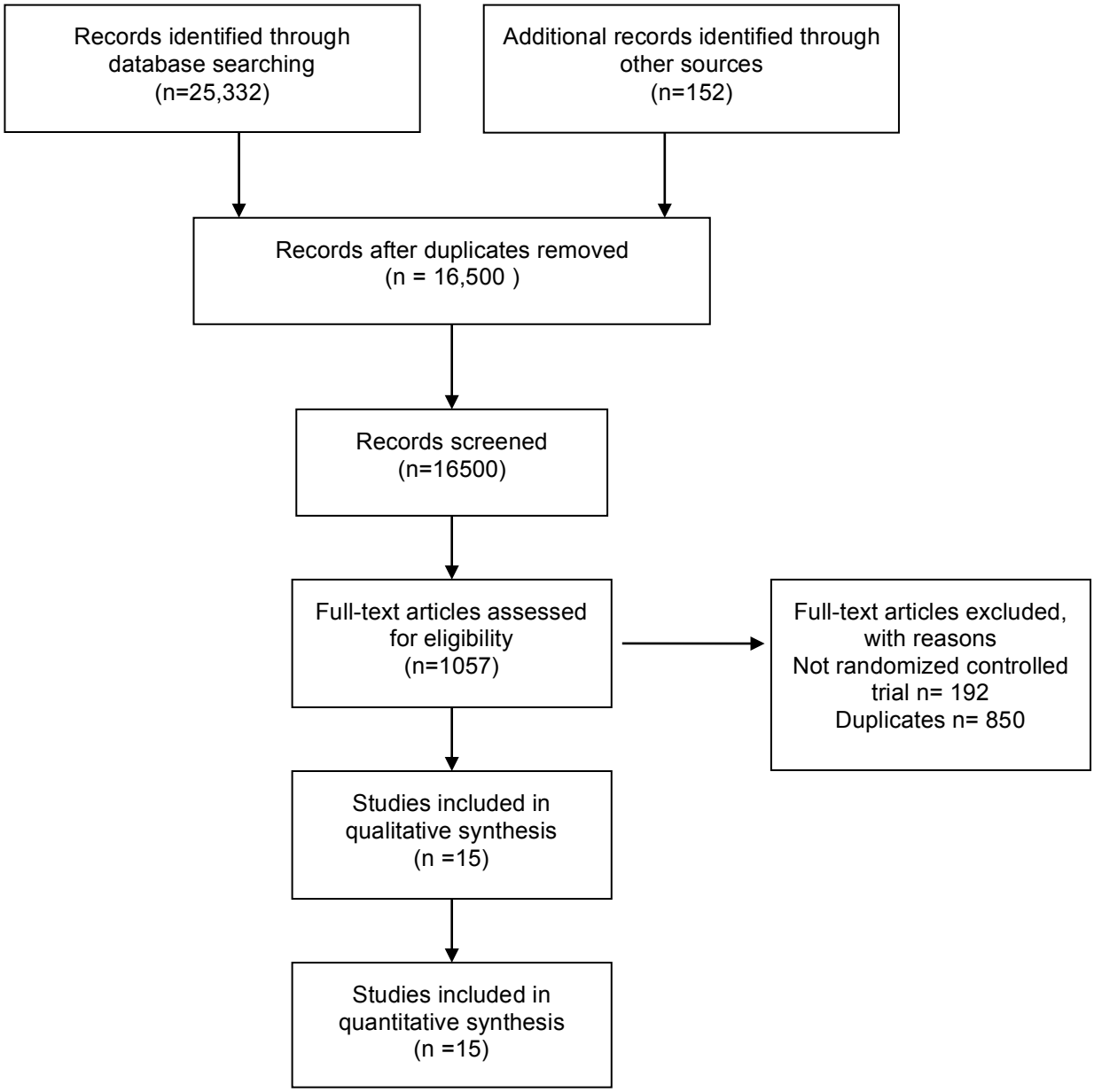


Figure Niii: PRISMA Flow chart for cohort studies included in the surgery-first pathway of DES model

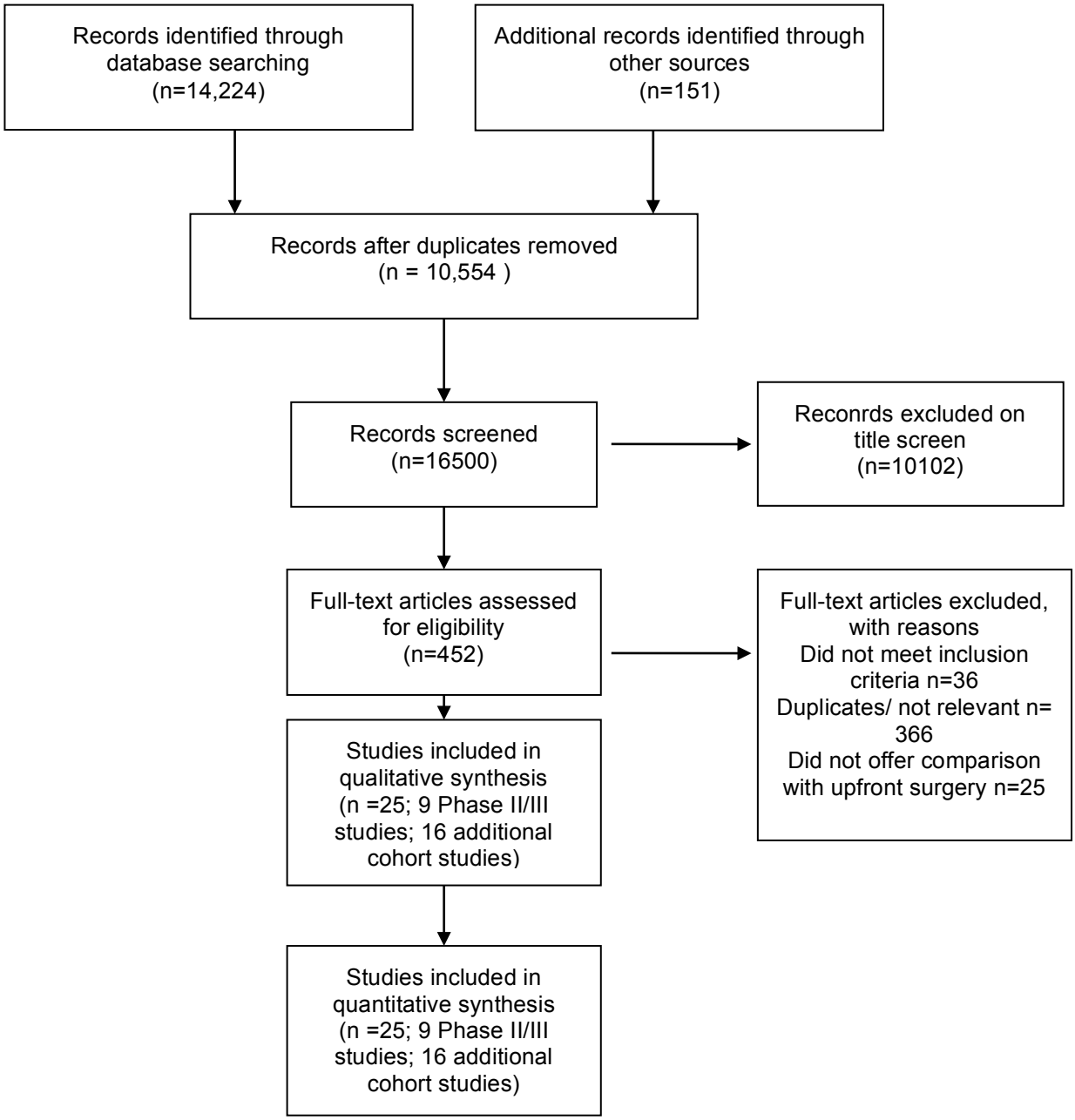


Table Ni: Summary of included studies within the DES model

Study	Type of Study	Treatment Regime	N=	Disease Free Survival in months	Overall Survival in months	ROBINS-I risk of bias assessment
Al-Sukhun *et al.*, 2003	Prospective Phase II Trial	CRT + surgery	20		13.4	Moderate
Cardenes *et al.*, 2011	Prospective Phase II Trial	CRT + surgery	28		10.3	Moderate
Casadei *et al.*, 2015	Prospective Phase II Trial	CRT + surgery	18		28.3	Low/Moderate
Cetin *et al.*, 2013	Prospective Phase II Trial	CRT + surgery	11			Moderate
Chakraborty *et al.*, 2014	Prospective Phase II Trial	CRT + surgery	13	2.4	9.1	Moderate
Crane *et al.*, 2011	Prospective Phase II Trial	CRT + surgery	69		19.2	Low/Moderate
Epelbaum *et al.*, 2002	Prospective Phase II Trial	CRT + surgery	20		8	Moderate
Esnaola *et al.*, 2014	Prospective Phase II Trial	CRT + surgery	37	10.4	11.8	Moderate
Evans *et al.*, 2008	Prospective Phase II Trial	CRT + surgery	86	15.4	22.7	Moderate
Fiore *et al.*, 2017	Prospective Phase II Trial	CRT + surgery	34	20	19.2	Moderate
Golcher *et al.*, 2008	Prospective Phase II Trial	CRT + surgery	121			Low
Golcher *et al.*, 2015	Prospective Phase II Trial	CRT + surgery	33	8.4	17.4	Moderate
Heinrich *et al.*, 2008	Prospective Phase II Trial	CT + surgery	28	9.2	26.5	Moderate
Herman *et al.*, 2015	Prospective Phase II Trial	CRT + surgery	49	7.8	13.9	Low/Moderate
Hong *et al.*, 2014	Prospective Phase II Trial	CRT + surgery	50	10.4	17.3	Moderate
Jang *et al.*,	Prospective	CRT +	27		21	Low

2018	Phase II Trial	surgery				
Jensen *et al*., 2014	Prospective Phase II Trial	CRT + surgery	23		11.5	Moderate
Joensuu *et al*., 2004	Prospective Phase II Trial	CRT + surgery	33	18	25	Moderate
Kim *et al*., 2013a	Prospective Phase II Trial	CRT + surgery	68		18.2	Low/Moderate
Landry *et al*., 2010	Prospective Phase II Trial	CRT + surgery	21	14.2	19.4	Low
Laurent *et al*., 2009	Prospective Phase II Trial	CRT + surgery	22	8	17	Moderate
Le Scodan *et al*., 2009	Prospective Phase II Trial	CRT + surgery	41		9.4	Low/Moderate
Lee *et al*., 2012	Prospective Phase II Trial	CT + surgery	43	10	16.6	Moderate
Leone *et al*., 2013	Prospective Phase II Trial	CRT + surgery	39	10.2	16.7	Moderate
Lin *et al*., 2005	Prospective Phase II Trial	CRT + surgery	42		10.3	Moderate
Lind *et al*., 2008	Prospective Phase II Trial	CRT + surgery	17		19	Moderate
Magnin *et al*., 2003	Prospective Phase II Trial	CRT + surgery	32		16	Moderate
Magnino *et al*., 2005	Prospective Phase II Trial	CRT + surgery	23		14	Moderate
Marti *et al*., 2008	Prospective Phase II Trial	CRT + surgery	26	7	13	Moderate
Mattiucci *et al*., 2010	Prospective Phase II Trial	CRT + surgery	40		15.5	Moderate
Massucco *et al*., 2006	Prospective Phase II Trial	CRT + surgery	28	10	15.4	Moderate
Maximous *et al*., 2009	Prospective Phase II Trial	CRT + surgery	25		12	Moderate
Mornex *et al*., 2006	Prospective Phase II Trial	CRT + surgery	41		9.4	Low/Moderate
Motoi *et al*., 2013	Prospective Phase II Trial	CT+ surgery	35		19.7	Low/Moderate
Moutardier *et al*., 2002	Prospective Phase II	CRT + surgery	19		20	Moderate

	Trial					
O'Reilly *et al.*, 2014	Prospective Phase II Trial	CT + surgery	38		27.2	Moderate
Palmer *et al.*, 2007	Prospective Phase II Trial	CT + surgery	50		13.6	Low
Pipas *et al.*, 2012	Prospective Phase II Trial	CRT + surgery	37		17.3	Moderate
Pister *et al.*, 2002	Prospective Phase II Trial	CRT + surgery	37		12	Moderate
Sahora *et al.*, 2011	Prospective Phase II Trial	CT + surgery	25		16	Moderate
Satoi *et al.*, 2009	Prospective Phase II Trial	CRT + surgery	35		24.5	Moderate
Sherman *et al.*, 2015	Prospective Phase II Trial	CRT + surgery v CT + surgery	45	34	29/42	Moderate
Small *et al.*, 2011)	Prospective Phase II Trial	CRT + surgery	29	9.9	11.8	Moderate
Talamonti *et al.*, 2006	Prospective Phase II Trial	CRT + surgery	20			Moderate
Tinchon *et al.*, 2013	Prospective Phase II Trial	CT + surgery	12			Moderate
Turrini *et al.*, 2009	Prospective Phase II Trial	CRT + surgery	34		15.5	Moderate
Van Buren *et al.*, 2013	Prospective Phase II Trial	CRT + surgery	59	6.6	16.8	Moderate
Varadhachary *et al.*, 2008	Prospective Phase II Trial	CRT + surgery	90	13.2	17.4	Moderate
Vento *et al.*, 2007	Prospective Phase II Trial	CRT + surgery	22		30.2	Moderate
Wilkowski *et al.*, 2009	Prospective Phase II Trial	CRT + surgery	93	5.6	9.3	Low
Regine *et al.*, 2011	RCT	Surgery + CRT	230		17.1	See figure below
Neoptolemos *et al.*, 2010	RCT	Surgery +CT	551	14.1	23	
VanLaethem *et al.*, 2010	RCT	Surgery +CRT	45	11.8	24.3	
Schmidt *et al.*, 2012	RCT	Surgery +CRT	53	15.2	26.5	
Reni *et al.*,	RCT	Surgery	51	11.7	26.2	

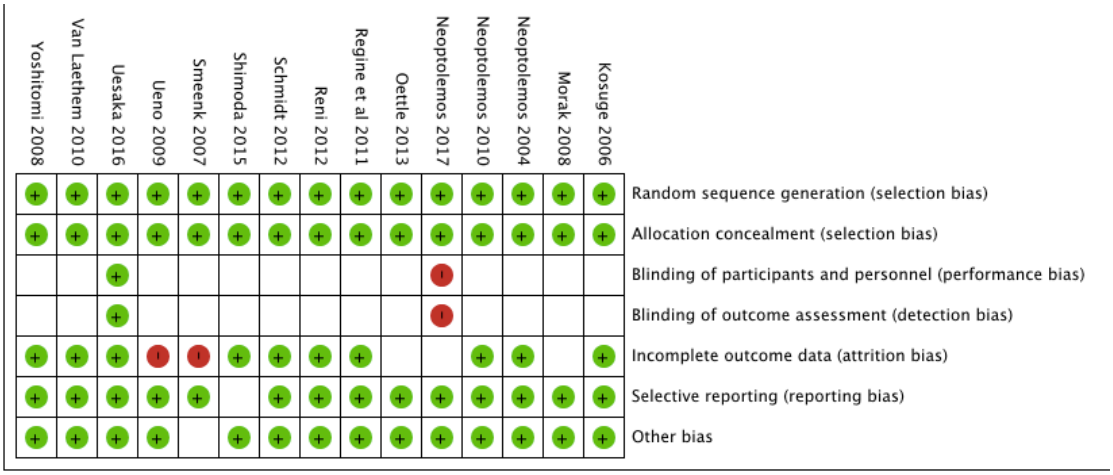
2012		+CRT				
Yoshitomi *et al.*, 2008	RCT	Surgery +CT	49	12	29.8	
Shimoda *et al.*, 2015	RCT	Surgery +CT	29	14.6	21.5	
Uesaka *et al.*, 2016	RCT	Surgery +CT	187	22.9	46.5	
Neoptolemos *et al.*, 2004	RCT	Surgery +CRT	145	10.7	15.9	
Ueno *et al.*, 2009	RCT	Surgery +CT	58	11.4	22.3	
Oettle *et al.*, 2013	RCT	Surgery +CT	179	13.4	22.8	
Kosuge *et al.*, 2006	RCT	Surgery +CT	45	8.6	12.5	
Smeenk *et al.*, 2007	RCT	Surgery +CRT	110	18	21.6	
Morak *et al.*, 2008	RCT	Surgery +CR	59	12	19	
Neoptolemo *et al.*, 2017	RCT	Surgery + CT	366		25.5	
Regine *et al.*, 2011	RCT	Surgery +CRT	221		20.5	
Neoptolemos *et al.*, 2010	RCT	Surgery +CT	537	14.3	23.6	
VanLaethem *et al.*, 2010	RCT	Surgery +CT	45	10.9	24.4	
Schmidt *et al.*, 2012	RCT	Surgery +CT	57	11.5	28.5	
Reni *et al.*, 2012	RCT	Surgery + CT	49	15.2	31.6	
Yoshitomi *et al.*, 2008	RCT	Surgery + CT	50	2.3	21.2	
Shimoda *et al.*, 2015	RCT	Surgery +CT	28	10.5	18	
Uesaka *et al.*, 2016	RCT	Surgery + CT	190	11.3	25.5	
Neoptolemos *et al.*, 2004	RCT	Surgery +CT	147	15.3	20.1	
Ueno *et al.*, 2009	RCT	Surgery Only	60	5	18.4	
Oettle *et al.*, 2013	RCT	Surgery Only	175	6.7	20.2	
Kosuge *et al.*, 2006	RCT	Surgery Only	44	10.2	15.8	
Smeenk *et al.*, 2007	RCT	Surgery Only	108	14.4	19.2	
Morak *et al.*, 2008	RCT	Surgery Only	61	7	18	
Neoptolemo *et al.*, 2017	RCT	Surgery + CT	364		28	

Al-Sukhun *et al.*, 2003	Prospective Phase II Trial	Surgery + adjuvant therapy	21		18.1	Moderate
Casadei *et al.*, 2015	Prospective Phase II Trial	Surgery + adjuvant therapy	20		27.5	Moderate
Golcher *et al.*, 2008	Prospective Phase II Trial	Surgery + adjuvant therapy	58		21	Moderate
Golcher *et al.*, 2015	Prospective Phase II Trial	Surgery + adjuvant therapy	33	8.7	14.4	Low
Lind *et al.*, 2008	Prospective Phase II Trial	Surgery + adjuvant therapy	35		11	Moderate
Massucco *et al.*, 2006	Prospective Phase II Trial	Surgery + adjuvant therapy	44	8	14	Moderate
Satoi *et al.*, 2009	Prospective Phase II Trial	Surgery + adjuvant therapy	41		18.5	Moderate
Vento *et al.*, 2007	Prospective Phase II Trial	Surgery + adjuvant therapy	25		35.9	Moderate
Jang *et al.*, 2018	Prospective Phase II Trial	Surgery + adjuvant therapy	23		12	Low
DeGus *et al.*, 2017a	Retrospective Cohort	Surgery + adjuvant therapy	6840		24.2	Moderate/Serious
Mellon *et al.*, 2016	Retrospective Cohort	Surgery + adjuvant therapy	241		22.1	Moderate/Serious
Nurmi *et al.*, 2018	Retrospective Cohort	Surgery + adjuvant therapy	150	13	26	Serious
Shubert *et al.*, 2016	Retrospective Cohort	Surgery + adjuvant therapy	216		13	Moderate/Serious
Artinya *et al.*, 2011	Retrospective Cohort	Surgery + adjuvant therapy	419		19	Serious
Ielpo *et al.*, 2016	Prospective Cohort	Surgery + adjuvant therapy	36		22.1	Serious
Roland *et al.*, 2015	Prospective Cohort	Surgery + adjuvant therapy	85			Moderate/Serious
DeGus *et al.*, 2017b	Retrospective Cohort	Surgery + adjuvant therapy	11316		Resectable: 24.5 Borderline: 20.0 Locally advanced: 15.5	Moderate/Serious
Mokdad *et al.*, 2017	Retrospective Cohort	Surgery + adjuvant therapy	6015		21	Moderate/Serious

Chen *et al.*, 2017	Retrospective Cohort	Surgery + adjuvant therapy	98		17	Moderate/Serious
Tzeng *et al.*, 2014	Prospective Cohort	Surgery + adjuvant therapy	52		25.3	Moderate
Fujii *et al.*, 2015	Prospective Cohort	Surgery + adjuvant therapy	71		13.1	Moderate/Serious
Fujii *et al.*, 2017	Prospective Cohort	Surgery + adjuvant therapy	416		Resectable: 23.5 Borderline: 20.1	Moderate/Serious
Papalezova *et al.*, 2012	Retrospective Cohort	Surgery + adjuvant therapy	92		13	Moderate/Serious
Hirono *et al.*, 2016	Prospective Cohort	Surgery + adjuvant therapy	124		13.7	Moderate/Serious
Murakami *et al.*, 2017	Retrospective Cohort	Surgery + adjuvant therapy	25		11.6	Serious

CRT= chemoradiotherapy CT= chemotherapy

Figure Niv: Risk of Bias of randomised controlled trials in DES model



Appendix O

Engaging with FUPS Data to Improve Individualised Treatment Pathway Selection

Introduction

Early complete surgical resection is the only potentially curative treatment for pancreatic cancer but in reality most patients will develop recurrence (Winter *et al.,* 2012). In the face of such challenges the need to risk-stratify potential surgical patients pre-operatively in an objective and standardised way is paramount to effective patient counseling (Lewis & Volmer, 2012). This is particularly pertinent in the high precision field of pancreatic cancer surgery where surgical volume is low, with only 10% of cases being resectable at presentation, yet operative mortality and morbidity rates are high (Lewis & Volmer, 2012).

The analysis thus far in this thesis presents an emerging picture whereby rather than a conclusively superior pathway beginning to emerge, the superior treatment pathway selection is dependent on individual patient and tumour factors. However, existing prognostic models are limited in scope and volume, falling short of differentiating patients who would, and importantly would not, benefit from competing treatment options (Lewis & Volmer, 2012) with most relying on post-operative factors to make predictions.

In summary decision making in the high-risk field of pancreatic cancer surgery is complex, involving uncertainty, and compounded by the fact that few cases are resectable at presentation. Unlike conventional statistics Bayesian statistical approach offers an opportunity to model this complexity in a way that conventional statistical analysis cannot. Such conventional models rely on maximum likelihood estimation with high dimensional integration needed to achieve this. Consequently conventional estimation is not available for many multilevel latent variable models. Bayesian estimation can obtain otherwise impossible parameters estimates (Kim *et al*., 2013), and produce more accurate parameter estimates (Depaoli, 2013), even in situations of small sample sizes (Zhang *et al*., 2007). Furthermore the Bayesian interpretation of 95% Cl is intuitively appealing within a clinical context whereby it represents the upper and lower limits where there is a 95% probability that the regression coefficient lies (Kim *et al*., 2013). Finally, as previously discussed, there is a degree of uncertainty within the existing body of literature regarding the treatment of potentially resectable PDAC stemming from its previously discussed FUPS characteristics. Bayesian statistics allows existing certainty, and uncertainty, to be incorporated and update this knowledge.

The aim of this section is use Bayesian statistical approach to further more effectively engage with available FUPS data to better risk stratify patients with potentially resectable disease into those with anticipated prognostic outcomes of 12months or less, or 36months or more, post resection. This will first be undertaken through a Bayesian analysis of the West of Scotland Pancreatic Unit database.

These results will then be triangulated with internationally available data from post resection survival analysis. Unlike previous sections this triangulation will not only be performed for comparison but taken further in section 4.6 to create prognostic Bayesian Belief Networks that can engage with the complexity of the dynamic adaptive system being modeled and utilise FUPS data to perform personalised predictions of prognostic outcome across the entire trajectory of the patient journey. This section will test the hypothesis that by viewing pancreatic management through the lens of complexity theory and utilising Bayesian statistical approach to more openly engage with FUPS data the limitations of existing predictive models can be addressed and a significant step towards the delivery of personalised realistic medicine for pancreatic cancer be achieved.

A Bayesian Analysis of the West of Scotland Pancreatic Unit Pancreatic Ductal Adenocarcinoma Database Identifying Pre-operative Factors Predicting Survival of 36 Months or More and 12 months or less.

Abstract

Background: Pancreatic cancer surgery carries high risks with potential benefits nullified by high rates of early disease reoccurrence. The growing interest in personalised predictive medicine means that there is a growing mandate to pre-operatively risk stratify patients to facilitate better patient counseling and support clinical decision making.

Methods: Bayesian statistical analysis was undertaken of the West of Scotland Pancreatic unit database of all potentially resectable cases of PDAC referred to the unit (n=418). Separate subgroup analysis of resectable only cases of PDAC (RPDAC) treated in surgery-first (n=100) and neoadjuvant (n=56) treatment pathways was also undertaken. The aims of this study were to 1) identify pre-operative variables that may predict good prognosis defined as 36 months or more 2) identify pre-operative variables that predict poor prognosis defined as 12 months or less and 3) assess whether one treatment pathway had superiority in treatment of RPDAC.

Results: Bayesian One-way ANOVA and log-linear regression analysis identified AJCC stage (*P value*: 0.000), tumour size above or below 3 centimeters (*P value*: 0.005), ASA grade (*P value*: 0.002), albumin(*P value*: 0.047) and modified Glasgow Prognostic Score (*P value*: 0.031) as statistically significant in predicting survival of 36 months or more. Bayesian linear regression analysis of all pre-operative factors produced a model of good fit to the regression line (R: 0.955; R squared 0.912*; P value:* 0.006).
Bayesian one-way ANOVA and log-linear regression analysis identified modified Glasgow prognostic score and tumour size greater than 3 centimeters as being significant in predicting survival of 12 months or less (*P value* 0.505 and 0.037).

Subgroup analysis of treatment pathways surgery-first versus neoadjuvant therapy for RPDAC, did not demonstrate statistically significant superiority of one pathway (one-way ANOVA p value: 0.808 and 0.163 respectively; log-linear regression *P value*: 0.87 and

0.871 respectively). Surgery-first pathway did demonstrate superiority in achieving R0 resection (one-way ANOVA *P value*: 0.025; log-linear regression *P value*: 0.025; surgery-first posterior mean: 0.795; 95% CI 0.698-0.891 v neoadjuvant posterior mean: 0.550; 95% CI 0.360-0.740). However receipt of multimodal treatment within either pathway was found to be statistically significant in determining survival outcome (one-way ANOVA *P value*: 0.000; linear regression and log-linear regression *P value*: 0.00) although there was no statistically significant difference between pathways in achieving multimodal treatment (one-way ANOVA and linear regression *P value*: 0.150). Further analysis showed that in the surgery-first pathway completing adjuvant therapy was statistically significant to overall survival (one-way ANOVA, linear regression and log-linear regression *P value*: 0.003), but 61% of patients in the surgery-first pathway failed to complete adjuvant therapy.

Conclusion: This study shows that Bayesian statistical approach offers an opportunity to model the complexity and uncertainty surrounding the management of PDAC in a way that conventional statistical analysis cannot. Furthermore it has highlighted factors that can be identified at the pre-operative stage of patient counseling to predict chances of a favorable and unfavorable prognosis. Whilst neoadjuvant pathway compared favorably to the traditional surgery-first pathway for RPDAC the most appropriate treatment pathway depends on individual patient and tumour factors in determining receipt of multimodal treatment.

Introduction

The aims of this study are to:

1) identify pre-operative variables that may predict good post resection prognosis defined as 36 months or more survival time

2) identify pre-operative variables that predict poor post resection prognosis defined as 12 months or less survival time

3) perform subgroup analysis of all patients in the neoadjuvant treatment pathway to assess impact of response to neoadjuvant therapy in predicting survival outcomes

4) for RPDAC cases only assess whether surgery-first or neoadjuvant treatment pathway is superior in terms of overall survival.

Methods

Patient Population

West of Scotland Pancreatic Centre Glasgow Royal Infirmary is a tertiary referral centre serving a population of 2.2-2.5 million. The unit has a prospectively maintained clinical database. The unit's retrospective database contains 20 years worth of patient data. This includes 418 PDAC patients referred to the unit with potentially

resectable PDAC. Until 2012, patients received a conventional "surgery first" approach with adjuvant chemotherapy as standard of care (n=312). From 2013-2015 patients received a neoadjuvant treatment protocol as standard (n=108).

For subgroup analysis of RPDAC cases, those included in the neoadjuvant pathway were patients with RPDAC on completion of initial staging prior to commencing NAT (n=56). Borderline and locally advanced PDAC were determined according to AHPBA/SSO/SSAT guidelines (Caller *et al.* 2009). Neoadjuvant regime was FOLFIRINOX unless patients had: poor performance status, or were aged over 70 years, or FOLFIRINOX was poorly tolerated, whereby they received Gemcitabine+Capcitabine (GEMCAP). Where the subgroup analysis performed assessment between resectable PDAC in neoadjuvant and surgery first arm the surgery first cohort was selected from August 2012 working backwards, 100 sequential patients in SF pathway who had RPDAC, and were deemed fit for surgery based on performance status score and CPET populated the SF arm of the model. No patients were lost to follow-up. All data was analysed on an intention-to-treat basis.

Statistical Methods: transparency of analyses

Bayesian statistical package in SPSS version 25.0.0 was used to conduct Bayesian one-way ANOVA analysis of each pre-operative variable against the dependent variable of survival time.

Analysis of variance (ANOVA) is an important method in confirmatory and exploratory data analysis (Gelman, 2005). Bayesian one-way ANOVA analysis of each pre-operative variable against the dependent variable of survival time was therefore undertaken. In clinical practice all pre-operative variables are considered to some extent in the pre-operative phase of decision-making, sometimes informally with their impact gauged by clinical judgment, not quantification. Here one-way ANOVA was performed assuming survival outcome was observed according to the full model, $M_{F,}$ of pre-operative factors (Solari *et al.* 2008).

In classical Frequentists analysis of variance hypotheses testing is based on the statistic:

$$F= \frac{\sum^{k}_{i=1} n_i(\tilde{A}_i-\tilde{A})^2/(k-1)}{\sum^{k}_{i=1}\sum^{n_i}_{j=1}(A_{ij}-\tilde{A}_i)^2/(n-k)}$$

where $n=\sum^{k}_{i=1} n_i$, $\tilde{A}_i=n^{-1}_i \sum^{n_i}_{j=1} A_{ij}$ and $\tilde{A}=n^{-1} \sum^{k}_{j=1} Y_{ij}$

and hypothesis testing for a pair of variables (i,j), null hypothesis M_0, tested against hypothesis, M_1:

M_0: $\mu 1=\ldots=\mu k$ versus M_1: $\mu_i \neq \mu_j$

Under this approach nuisance parameters are eliminated and the full model M_F has k+1 parameters, with a scalar test statistic constructed to compare different values of the scalar non-centrality parameter. In

this sense the classical Frequentists approach acts as if only the marginal experiment is observed, disregarding information about the single mean treatments $\tilde{A}_1,\ldots\tilde{A}_k$ but it is not clear how much information is lost (Solari *et al.* 2008).

Bayesian inference about Pearson correlation coefficient was then performed to assess the linear relation between each pre-operative variable and survival time to draw Bayesian inference by estimating Bayes factors and characterising posterior distributions.

Linear regression analysis within the context of Bayesian inference (Bayesian univariate linear regression) was undertaken with pre-operative variables assessed for their ability to explain and predict values of survival time as a scaled outcome. Log-linear regression was then performed to test the independence of each pre-operative variable against the outcome of survival time.

For all the statistical analysis a default setting of least informed prior was used. Inferential statements therefore depend only on the available data and the assumed model.

Results: Pre-Operative Variables Predicting Survival Time of 36 months or More

One-way ANOVA analysis (table 0i) identified statistically significant variance between the means of the groups contained within the following variables pertaining to 36 months or more: AJCC stage (*P*

value: 0.000), tumour size above or below 3 centimeters (*P value:* 0.005), ASA grade (*P value:* 0.002), albumin level above or below 35(*P value:* 0.047) and modified Glasgow Prognostic Score (*P value:* 0.031).

Table Oi: One-Way ANOVA Analysis: Results of overall survival time equal or greater than 36months between groups contained within each pre-operative variable.

Variable	F statistic	.Sig	Bayes Factor
Albumin (<or= to 35 v >35)	3.99	0.047	0.352
Modified Glasgow Prognostic Score (mGPS)	3.534	0.031	0.151
Neutrophil Lymphocyte Ratio (NLR): (<2 v >2)	0.076	0.784	0.136
Neutrophil Lymphocyte Ratio (NLR): (<3 v >3)	0.920	0.344	0.205
Neutrophil Lymphocyte Ratio (NLR): (<5 v >5)	0.905	0.348	0.203
AJCC Stage	6.282	0.000	4.850
Tumour size </> 2cm	2.521	0.113	0.155
Tumour size </> 3cm	7.913	0.005	2.179
T stage grouped T1/T2 v T3/T4	0.393	0.531	0.054
T stage	1.013	0.387	0.001
Location: HOP v Body/Tail	0.780	0.378	0.074
Age </>70	2.298	0.130	0.123
Jaundice </> 40	0.334	0.564	0.057
ASA	5.051	0.002	0.644
Diabetes	0.006	0.937	0.061
Smoking	0.721	0.397	0.075
BMI	0.534	0.660	0.001

Bayesian Log Regression analysis (table Oii) produced Pearson Chi squared scores that corroborated statistical significance of these

variables in independently predicting 36 months survival or more (AJCC stage: *P value:* 0.00; tumour size above or below 3 centimeters: *P value:* 0.005; ASA grade: *P value:* 0.002; albumin level above or below 35: *P value:* 0.046; modified Glasgow Prognostic Score: *P value:* 0.031).

Table Oii: Bayesian Log Linear Regression: Pre-operative Variables: overall survival grouped as 36months or above or not

Variable	Bayes factor	Pearson Chi-Squared (.Sig in parenthesis)	Continuity Correction (.Sig in parenthesis)	Fishers Exact Test (Exat sig. -2 sided/ Exact Sig -1sided)
Albumin </> 35	0.855	3.970 (0.046)	3.314 (0.069)	0.058/0.033
mGPS	0.644	6.932 (0.031)		
NLR </> 2	2.283	0.080 (0.777)	0.001 (1.000)	1.000/0.564
NLR </>3	1.460	0.949 (0.330)	0.276 (0.599)	0.402/0.301
NLR </>5	2.953	0.934 (0.334)	0.069 (0.793)	1.000/0.454
AJCC Stage	0.243	23.651 (0.000)		
Tumour size </> 2cm	2.286	2.517 (0.113)	1.959 (0.162)	0.135/0.084
Tumour size </>3 cm	0.119	7.771 (0.005)	6.992 (0.008)	0.006/0.004
T stage ½ v 3/4	6.948	0.395 (0.530)	0.179 (0.672)	0.539/0.327
T stage	257.139	3.049 (0.384)		
Location; HOP v Body/Tail	6.158	0.784 (0.376)	0.411 (0.521)	0.470/0.269
Age </> 70 years	2.668	2.297 (0.130)	1.853 (0.173)	0.164/0.089
Jaundice	4.656	0.336 (0.562)	0.183 (0.669)	0.648/0.334
ASA	0.511	14.469 (0.002)		
Diabetes	6.202	0.006 (0.936)	0.001 (1.000)	1.000/0.588
Smoking	4.390	0.725 (0.395)	0.431 (0.511)	0.442/0.259
BMI	89.930	1.621 (0.655)		
Pathway: neoadjuvant v surgery first	0.580	5.214 (0.22)	4.523 (0.33)	0.027/0.014

Although Pearson Correlation did not produce any significant correlations between any of the pre-operative variables and overall survival time in months or categorized as 36 months or more (table

Oiii), Bayesian linear regression analysis of all pre-operative factors (table Oiv) produced a model of good fit to the regression line (R: 0.955; R squared 0.912; *p value* 0.006).

Table Oiii: Pearson Correlation Coefficient: results of each variable assessed against overall survival defined as 36 months or more (column 2) and overall survival time in months (column 3)

		OS	OS
OS	Pearson Correlation	1	.780
	Bayes Factor		.000
	N	411	410
OS	Pearson Correlation	.780	1
	Bayes Factor	.000	
	N	410	410
Albgroup	Pearson Correlation	.122	.089
	Bayes Factor	2.840	7.209
	N	266	266
mGPS	Pearson Correlation	-.183	-.158
	Bayes Factor	.566	1.363
	N	207	207
NLRgroup	Pearson Correlation	.048	-.039
	Bayes Factor	7.338	7.433
	N	35	35
NLR3	Pearson Correlation	-.165	-.018
	Bayes Factor	4.886	7.577
	N	35	35
NLR 5	Pearson Correlation	-.163	.090
	Bayes Factor	4.921	6.677
	N	35	35
AJCC Stage	Pearson Correlation	-.108	-.166
	Bayes Factor	3.479	.267
	N	322	321
Tumour Size	Pearson Correlation	-.088	-.045
	Bayes Factor	6.464	16.341
	N	323	322
size3	Pearson Correlation	-.155	-.160
	Bayes Factor	.459	.366
	N	323	322

T stage	Pearson Correlation	-.035	-.090
	Bayes Factor	18.548	6.194
	N	322	321
t_t	Pearson Correlation	.004	-.034
	Bayes Factor	22.548	18.802
	N	323	322
Tumour Location	Pearson Correlation	-.056	.018
	Bayes Factor	13.604	19.235
	N	254	254
agegroup1_70	Pearson Correlation	.075	.039
	Bayes Factor	8.112	18.743
	N	410	409
Jaunice	Pearson Correlation	.035	.037
	Bayes Factor	17.473	17.192
	N	269	268
(ASA)	Pearson Correlation	-.241	-.219
	Bayes Factor	.018	.063
	N	236	236
Diabetes	Pearson Correlation	-.006	.018
	Bayes Factor	16.528	16.121
	N	173	173
Smoking Hx	Pearson Correlation	-.056	-.038
	Bayes Factor	13.371	16.175
	N	231	231
BMI	Pearson Correlation	.025	.002
	Bayes Factor	16.760	17.804
	N	200	200

Table Oiv: Bayesian Linear Regression

ANOVA[a,b]

Source	Sum of Squares	df	Mean Square	F	Sig.
Regression	4.445	21	.212	4.917	.006
Residual	.430	10	.043		
Total	4.875	31			

a. Dependent Variable: OS

b. Model: (Intercept), Albgroup, NLRgroup, NLR3, NLR 5, mGPS , AJCC Stage, Tumour Size, size3, T stage, t_t , Tumour Location, agegroup1_70, Jaunice, (ASA) , Diabetes, Smoking Hx, BMI

Bayes Factor Model Summary[a,b]

Bayes Factor[c]	R	R Square	Adjusted R Square	Std. Error of the Estimate
1.783	.955	.912	.726	.2075

a. Method: JZS

b. Model: (Intercept), Albgroup, NLRgroup, NLR3, NLR 5, mGPS , AJCC Stage, Tumour Size, size3, T stage, t_t , Tumour Location, agegroup1_70, Jaunice, (ASA) , Diabetes, Smoking Hx, BMI

c. Bayes factor: Testing model versus null model (Intercept).

This finding is supported in part by the Bayes Factors produced by one-way ANOVA analysis whereby the other pre-operative variables that did not reach statistical significance in terms of *P value* produced Bayes Factors that supported rejecting the null hypothesis (table Oi). Interestingly conflict arose between *P value* and Bayes Factor for the variables AJCC stage and tumour size (when defined as either above or below 3 centimeters). However under Pearson Correlation coefficient and Bayesian log linear regression, the Bayes Factor

returned for both variables supported rejecting the null hypothesis (table Oii, table Oiii).

Some variables that did not reach statistical significance determined by *P value* in any of the analysis offered here, but had Bayes Factors supporting rejecting the null hypothesis in one-way ANOVA analysis (table Oi), went on to produced Bayes Factors under Pearson correlation and Bayesian log-linear regression that supported the null hypothesis. These included: tumour location, age, jaundice, diabetes, smoking history and BMI (table Oii; table Oiii). Tumour size when categorized as less or greater than 2 centimeters supported the null hypotheses, but when categorized as less or greater than 3 centimeters rejected the null hypotheses (table Oii; table Oiii). Pearson correlation for T stage produced Bayes Factors that seemed to strongly support the null hypotheses but this finding is likely to be skewed by the fact that most patients in the database had T stage falling into the T3/T4 category. Neutrophil lymphocyte ration (NLR), although having Bayes Factors that moderately supported rejecting the null hypothesis in one-way ANOVA and Pearson correlation, produced Bayes Factors that anecdotally supported the null hypothesis.

Pre-Operative Variables Predicting Survival Time of 12 months or Less.

Pre-operative factors were less strong in predicting survival time of 12 months or less. Although linear regression suggested good fit of modeling all pre-operative variables against survival outcome (R and R squared: 1.000) (table Ov), significance defined by *P value* was not reached.

Table Ov: Bayesian Linear Regression

ANOVA[a,b]

Source	Sum of Squares	df	Mean Square	F	Sig.
Regression	37888.403	113	335.296	26.824	.153
Residual	12.500	1	12.500		
Total	37900.903	114			

a. Dependent Variable: OS

b. Model: (Intercept), Alb_group, GPS , NLR3, NLR5, location, size , size3, Tstage, AJCC stage, age , PS , jaundice, DM , smoking, BMI

Bayes Factor Model Summary[a,b]

Bayes Factor[c]	R	R Square	Adjusted R Square	Std. Error of the Estimate
.912	1.000	1.000	.962	3.5355

a. Method: JZS

b. Model: (Intercept), Alb_group, GPS , NLR3, NLR5, location, size , size3, Tstage, AJCC stage, age , PS , jaundice, DM , smoking, BMI

c. Bayes factor: Testing model versus null model (Intercept).

One-way ANOVA and log-linear regression identified modified Glasgow Prognostic Score and tumour size less than or greater than 3 centimeters as significant (*P value*: 0.050 and 0.037 respectively) (table Ovi and table Ovii). One-way ANOVA analysis produced Bayes Factors for each variable that weakly supported rejecting the null hypothesis (table Ovi). However, in both Pearson correlation and log-linear analysis Bayes Factors produced were found to support accepting the null hypothesis with the exception of tumour size greater or less than 3 centimeters and BMI in log-linear analysis. Pearson correlation also did not identify and significant correlation between pre-operative variables and the overall survival time or survival categorized as 12 months or less versus greater than 12 months (table Oviii).

Table Ovi: One-Way ANOVA Analysis: Results of overall survival time equal or less than 12months between groups contained within each pre-operative variable.

Variable	F statistic	.Sig	Bayes Factor
Albumin (<or= to 35 v >35)	2.610	0.108	0.202
Modified Glasgow Prognostic Score (mGPS)	2.415	0.050	0.009
Neutrophil Lymphocyte Ratio (NLR): (<3 v >3)	0.237	0.789	0.007
Neutrophil Lymphocyte Ratio (NLR): (<5 v >5)	0.843	0.360	0.088
AJCC Stage	1.406	0.232	0.000
Tumour size </> 2cm	1.911	0.168	0.115
Tumour size </> 3cm	4.381	0.037	0.388
T stage grouped T1/T2 v T3/T4	3.335	0.069	0.232
Location: HOP v Body/Tail	0.652	0.420	0.069
Age </>70	1.703	0.193	0.92
Jaundice	2.010	0.157	0.132
ASA	1.619	0.186	0.005
Diabetes	0.534	0.446	0.079
Smoking	0.001	0.970	0.052
BMI	0.939	0.631	0.000
Pathway	0.097	0.755	0.041

Table Ovii: Bayesian Log Linear Regression: Pre-operative Variable: Overall survival grouped as 12 months or less or greater than 12 months

Variable	Bayes factor	Pearson Chi-Squared (.Sig in parenthesis)	Continuity Correction (.Sig in parenthesis)	Fishers Exact Test (Exat sig. -2 sided/ Exact Sig -1sided)
Albumin </> 35	1.147	2.602 (0.107)	2.136 (0.144)	0.103/0.072
mGPS	14.504	9.432 (0.050)		
NLR </>3	53.156	0.480 (0.786)		
NLR </>5	3.603	0.849 (0.357)	0.508 (0.476)	0.411/0.241
AJCC Stage	245.374	5.612 (0.230)		
Tumour size </> 2cm	2.859	1.911 (0.167)	1.478 (0.224)	0.185/0.111
Tumour size </>3 cm	0.580	4.348 (0.037)	3.873 (0.049)	0.046-0.024
T stage ½ v 3/4	1.406	3.321 (0.068)	2.729 (0.99)	0.088/0.046
Location; HOP v Body/Tail	4.514	0.656 (0.418)	0.398 (0.518)	0.477/0.267
Age </> 70 years	2.700	1.704 (0.192)	1.419 (0.234)	0.210/0.116
Jaundice	1.736	2.101 (0.156)	1.656 (0.198)	0.191/0.099
ASA	30.260	4.839 (0.184)		
Diabetes	4.146	0.539 (0.463)	0.260 (0.610)	0.510/0.311
Smoking	4.782	0.001 (0.970)	0.000 (1.000)	1.000/0.544
BMI	0.000	174.842 (0.532)		

Table Oviii: Pearson Correlation Coefficient: results of each variable assessed against overall survival defined as 12 months or less

		OS_code	OS
OS_code	Pearson Correlation	1	-.547
	Bayes Factor		.000
	N	410	410
OS	Pearson Correlation	-.547	1
	Bayes Factor	.000	
	N	410	411
Alb_group	Pearson Correlation	.113	-.046
	Bayes Factor	4.938	14.555
	N	203	204
NLR3	Pearson Correlation	-.017	-.062
	Bayes Factor	16.837	12.155
	N	189	190
NLR5	Pearson Correlation	-.067	-.026
	Bayes Factor	11.367	16.290
	N	188	189
GPS	Pearson Correlation	.167	-.128
	Bayes Factor	1.232	3.685
	N	191	192
location	Pearson Correlation	-.051	.018
	Bayes Factor	14.469	19.286
	N	253	254
size	Pearson Correlation	.077	-.037
	Bayes Factor	8.730	18.242
	N	322	323
size3	Pearson Correlation	.116	-.108
	Bayes Factor	2.579	3.480
	N	322	323
Tstage	Pearson Correlation	.102	-.072
	Bayes Factor	4.311	9.922
	N	321	322

AJCC stage	Pearson Correlation	.122	-.160
	Bayes Factor	2.091	.356
	N	321	322
age	Pearson Correlation	-.065	.016
	Bayes Factor	10.886	24.175
	N	409	410
PS	Pearson Correlation	.032	-.213
	Bayes Factor	17.148	.085
	N	235	236
jaundice	Pearson Correlation	.087	.016
	Bayes Factor	7.594	19.933
	N	268	269
DM	Pearson Correlation	-.056	.019
	Bayes Factor	12.680	16.086
	N	172	173
smoking	Pearson Correlation	.002	-.038
	Bayes Factor	19.078	16.192
	N	230	231
BMI	Pearson Correlation	-.083	.090
	Bayes Factor	8.668	7.423
	N	239	240

Subgroup Analysis of Response to Neoadjuvant Therapy in Neoadjuvant Pathway Related to Survival Time

Analysis of CT response to neoadjuvant therapy was not able to produce any statistically significant relationship with survival time (table Oix-Oxi).

Table Oix: One-way ANOVA Neoadjuvant Sub-Group Analysis: Assessment of response to neoadjuvant treatment for all patients treated in neoadjuvant pathway

Variable	F statistic	.Sig	Bayes Factor
Response to NAT	2.766	0.68	0.140

Table 0x: Pearson Correlation co-efficient Neoadjuvant Sub-Group Analysis

Bayes Factor Inference on Pairwise Correlations[a]

		OScode	Radiologicalresponsecode
OScode	Pearson Correlation	1	-.129
	Bayes Factor		5.578
	N	108	100
Radiologicalresponsecode	Pearson Correlation	-.129	1
	Bayes Factor	5.578	
	N	100	100

a. Bayes factor: Null versus alternative hypothesis.

Table 0xi: Bayesian Log Linear Regression Neoadjuvant Sub-Group Analysis

Test of Independence[a]

	Value	df	Asymptotic Sig.(2-sided)
Bayes Factor	1.523[b]		
Pearson Chi-Square	5.395[c]	2	.067

a. The total sum is fixed in the contingency table.

b. This analysis tests independence versus association, and assumes a multinomial model and conjugate priors.

c. 3 cells(50.0%) have expected count less than 5. The minimum expected count is 1.350.

Bayesian Subgroup Analysis Comparing Neoadjuvant and Surgery-First Treatment Pathways for Resectable Pancreatic Ductal Adenocarcinoma.

For cases of RPDAC Bayesian one-way ANOVA and Bayesian log-linear regression analysis of the treatment pathways surgery-first versus neoadjuvant therapy, did not demonstrate statistically significant superiority of one pathway related to survival time whether categorized as 36 months or more, or overall survival time in months (one-way ANOVA *P value:* 0.808 and 0.163 respectively; log-linear regression *P value*: 0.87 and 0.871 respectively) (table Oxii; table Oxiii). Pearson correlation coefficient also failed to produce and significant correlation between survival and treatment pathways and Bayes Factor supported the null hypothesis that there was no statistically significant difference in outcome between treatment pathways (table Oxiv).

Statistical significance was found between treatment pathways pertaining to achievement of R0 resection (table Oxii iii and table Oxiii iii) in favour of surgery-first pathway (one-way ANOVA *P value:* 0.025; log-linear regression *P value*: 0.025; surgery-first posterior mean: 0.795; 95% CI 0.698-0.891 v neoadjuvant posterior mean: 0.550; 95% CI 0.360-0.740).

Achievement of multimodal treatment (either neoadjuvant therapy and resection or upfront surgical resection and adjuvant therapy) was found to be statistically significant to overall survival time (one-way ANOVA *P value:* 0.000; linear regression and log-linear

regression *P value:* 0.00). There was no statistically significant difference between pathways in achieving multimodal treatment (one-way ANOVA and linear regression *P value:* 0.150).

Within the surgery-first pathway completing adjuvant therapy was statistically significant to overall survival (one-way ANOVA, linear regression and log-linear regression *P value:* 0.003). However, 61% of patients within this pathway did not complete adjuvant therapy of which 53 out of 77 lived under 36 months. 8 out of 23 patients did not complete adjuvant therapy but achieved survival time of 36 months or more.

Table Oxii: One-way ANOVA RPDAC Sub-Group Analysis

i) pathway related to overall survival time of 36months or more

ANOVA

OS code	Sum of Squares	df	Mean Square	F	Sig.	Bayes Factor[a]
Between Groups	.010	1	.010	.059	.808	.069
Within Groups	24.026	136	.177			
Total	24.036	137				

a. Bayes factor: JZS

ii) pathway related to overall survival time in months

ANOVA

overall_surv_months	Sum of Squares	df	Mean Square	F	Sig.	Bayes Factor[a]
Between Groups	918.099	1	918.099	1.965	.163	.178
Within Groups	63538.401	136	467.194			
Total	64456.500	137				

a. Bayes factor: JZS

iii) pathway related to achieving R0 resection

ANOVA

t_margin	Sum of Squares	df	Mean Square	F	Sig.	Bayes Factor[a]
Between Groups	.955	1	.955	5.186	.025	.974
Within Groups	17.668	96	.184			
Total	18.622	97				

a. Bayes factor: JZS

Bayesian Estimates of Coefficients[a,b,c]

	Posterior			95% Credible Interval	
Parameter	Mode	Mean	Variance	Lower Bound	Upper Bound
Surgery First	.795	.795	.002	.698	.891
Neoadjuvant	.550	.550	.009	.360	.740

a. Dependent Variable: t_margin

b. Model: neoadj_1

c. Assume standard reference priors.

Table Oxiii: Bayesian Log Linear Regression RPDAC Sub-Group Analysis

i) Pathway compared to survival of 36 months or more

Test of Independence[a]

	Value	df	Asymptotic Sig.(2-sided)	Exact Sig.(2-sided)	Exact Sig.(1-sided)
Bayes Factor	4.196[b]				
Pearson Chi-Square	.060[c]	1	.807		
Continuity Correction	.000	1	.987		
Fisher's Exact Test				1.000	.501

ii) Pathway compared to survival in months

	Value	df	Asymptotic Sig.(2-sided)
Bayes Factor	1402.978[b]		
Pearson Chi-Square	40.728[c]	52	.871

iii) Pathway compared to R0 resection

	Value	df	Asymptotic Sig.(2-sided)	Exact Sig.(2-sided)	Exact Sig.(1-sided)
Bayes Factor	.361[b]				
Pearson Chi-Square	5.023[c]	1	.025		
Continuity Correction	3.817	1	.051		
Fisher's Exact Test				.042	.029

Table Oxiv: Pearson Correlation co-efficient RPDAC Sub-Group Analysis

		neoadj_1
neoadj_1	Pearson Correlation	1
	Bayes Factor	
	N	138
OS >36months	Pearson Correlation	-.021
	Bayes Factor	14.400
	N	138
overall_surv_months	Pearson Correlation	-.119
	Bayes Factor	5.627
	N	138
R0 resection	Pearson Correlation	-.226
	Bayes Factor	1.026
	N	98

Discussion

Findings as a partial remnant and triangulation

Bayesian One-way ANOVA and log-linear regression analysis identified AJCC stage, tumour size above or below 3 centimeters, ASA grade, albumin and modified Glasgow Prognostic Score as statistically significant in predicting post resection survival time of 36 months or more. These findings corroborate an existing body of studies that have also identified these pre-operative factors in predicting favorable survival time (Section 4.6). Pre-operative factors identified in previous studies as being statistical significant, but not in this study, for predicting prolonged survival included: NLR, tumour location, T stage, jaundice, and history of smoking and

diabetes (Section 4.6). However the advantage of Bayesian analysis showed that these variables were found to have Bayes Factors supporting rejection of the null hypothesis in this study also, hence corroborating their link, although to a lesser extent, in predicting prolonged survival time.

Bayesian one-way ANOVA and log-linear regression analysis identified modified Glasgow prognostic score and tumour size greater than 3 centimeters as being significant in predicting survival of 12 months or less which supports the findings the existing body of studies reporting the predictive significance of these factors pertaining to poor prognostic outcome (Section 4.6). Overall modeling of pre-operative variables to predict poor prognosis performed less well than modeling post-operative factors to predict prolonged prognosis. This corroborates the existing body of evidence exploring factors to predict poor outcome with existing models focusing more on post-operatively available data to make meaningful predictions (Bradley *et al*., 2019b; Bradley *et al*., 2019c).

Response to neoadjuvant treatment was not found to significantly predict survival time of greater than or equal to 36 months or 12 months or less. This finding is not surprising considering ambiguity surrounding the existing body of research on neoadjuvant therapy particularly for RPC. Although NAT is supported by current guidelines for borderline RPC, optimal treatment of RPDAC remains controversial (Tempero *et al.* 2014; deGeus 2016).

Subgroup analysis of treatment pathways surgery-first versus neoadjuvant therapy for RPDAC, did not demonstrate statistically significant superiority of one pathway. These findings are in keeping with the few existing RCTs and meta-analysis comparing neoadjuvant therapy and surgery first which have failed to demonstrate conclusively superior survival benefit of either pathway.

Surgery-first pathway did demonstrate superiority in achieving R0 resection. This could demonstrate that NAT allowed a period of time for more aggressive tumours, in which surgery would be ultimately futile, to declare themselves. Conversely it could be argued that this demonstrates losing the window of resectability. However, receipt of multimodal treatment within either pathway was again found to be statistically significant in determining survival outcome. Although there was no statistically significant difference between pathways in achieving multimodal treatment, analysis showed that in the surgery-first pathway completing adjuvant therapy was statistically significant to overall survival time but 61% of patients in the surgery-first pathway failed to complete adjuvant therapy. Considering that on an intention-to-treat basis, a significant number in each pathway failed to receive all intended treatment modalities this study adds an important dimension to the ongoing debate regarding treatment selection for RPDAC. This highlights the need to work towards better, personalised predictive medicine to assist in selecting the most appropriate treatment strategy for individual patients.

Conclusion

This study shows that Bayesian statistical approach offers an opportunity to model the complexity and uncertainty surrounding management of PDAC, and in particular RPDAC, in a way that conventional statistical analysis cannot. Furthermore it has highlighted factors that can be identified at the pre-operative stage of patient counseling to predict chances of a favorable and unfavorable prognosis. Whilst neither treatment pathway demonstrated superiority for management of RPDAC, this study shows that the most appropriate treatment pathway therefore depends on individual patient and tumour factors in determining receipt of multimodal treatment. This highlights the important step this study makes towards achieving personalised predictive medicine in research and clinical practice. Section 4.6 will therefore focus on exploring whether Bayesian statistical approach can be taken further to engage with complexity of handle FUPS data to deliver personalised realistic medicine by supporting better decision making and patient selection through personalised predictive medicine.

Appendix P

Post-resection Prognostic Models

Table Pi. Summary of existing model development studies predicting post-resection prognosis for PDAC

Study	Data Source (n=20510)	Model Outcome	Predictor Selection	Included Variables	Missing Data	Modeling Method	Model Performance and Validation
Brennan et al., (2004)	Single institution database (n=555)	1,2,3 year survival	Cox multivaraite analysis; p-value but non significant variables included	age, sex, portal vein inclusion, splenectomy, margin, location, differentiation, posterior margin, nodes positive, nodes negative, back pain, T stage, weight loss, maximum pathological axis (n=14)	Predicted using regression models	Multivaraite Cox proportional hazards regression	Calibration plot: good fit Discrimination: *c*-statistic 0.64 External Evaluation: absent; internal validation by bootstrap method
Kanda et al., (2014)	Single institution database (n=324)	Death within 12 months of surgery	Pre-selection by Univaraiate analysis then Cox multivaraite analysis. P-value < 0.1	CEA, CA19-9 (n=2)	Complete case analysis	Multivaraite Cox proportional hazards regression	Calibration: absent Discrimination: AUC 0.702 External Evaluation: absent
Miura et al., (2014)	Single institution database (n=50)	1,3,5 year survival	Pre-selection by Univaraiate analysis then Cox multivaraite analysis. P-value <0.05	Platelet count, CRP, CA19-9 (n=3)	Complete case analysis	Multivaraite Cox proportional hazards regression	Calibration: none Discrimination: none External Evaluation: absent
Shen et al., (2018)	Multi-centre databases (n=239)	6,12,18 month survival	Pre-selection by Univaraiate	age, length of tumour contact, peripancreatic venous		Multivaraite Cox proportional hazards	Calibration plot: good fit Discrimination: *c*-

			analysis then Cox multivariate analysis. P-value <0.05	abnormality, lymph node staging (n=4)		regression	statistic 0.824 External Evaluation: performed
Xu et al., (2017)	Single institution database (n=265)	1,3,5 year survival	Pre-selection by Univariate analysis then Cox multivariate analysis. P-value <0.05	Tumour grade, pathological stage, neural invasion, vascular invasion, Neutrophil Lymphocyte Ratio, Platelet to Lymphocyte Ratio, Albumin Globulin Ratio (n=7)		Multivariate Cox proportional hazards regression	Calibration: calibration curve Discrimination: 1yr: c-index 0.86, AUC: 0.938. 3yr: c-index 0.837, AUC 0.844, 5yr: c-index: 0.809, AUC 0.884 External Evaluation: absent
Walczak & Velanovich (2017)	Single institution database (n=219)	7 month survival post surgery	Single hidden layer back propogation trained ANN.	age, sex, stage, survival time, quality of life, adjuvant therapy, resection details (n=7)	Complete case analysis	Artificial Neural Network	Calibration: absent Discrimination: AUC: 0.6576, sensitivity 91.30%, specificity 38.27% External Validation: absent; internal validation by random split method
Hsu et al., (2012)	Single institution database (n=740)	Death at 9 and 12 months	Pre-selection by Univariate analysis then Cox multivariate analysis. P-value <0.05	age, tumour size, comorbidities, tumour grade (n=4)	Complete case analysis	Multivariate Cox proportional hazards regression	Calibration: none Discrimination: none External Validation: absent; internal validation with p-value <0.05
Botsis et al., (2009)	Single institution database (n=218)	Survival time	Pre-selection by Univariate analysis then Cox	Age, differentiation, tumour size, Alk Phos, Albumin, Ca19-9 (n=6)	MICE presuming data missing at random	Multivariate Cox proportional hazards regression	Calibration: absent Discrimination: *c*-statistic 0.73 External

			multivaraite analysis. P-value <0.05				Validation: absent; internal validation by bootstrap method
Liu et al., (2018)	Multi-centre databases (n=1223)	Survival time	Pre-selection by Univaraiate analysis then Cox multivaraite analysis. P-value <0.05	American Joint Commission on Cancer stage, tumour grade, post-operative Ca19-9 (n=3)		Multivaraite Cox proportional hazards regression	Calibration: absent Discrimination: *c*-statistic 0.70, AIC: 2406.37 External Validation: absent; internal validation by random split method
Balzano et al., (2017)	Single institution database (n=296)	1 year mortality	Pre-selection by Univaraiate analysis then Cox multivaraite analysis. P-value <0.2 for univariate analysis and <0.1 for multivariate analysis	American Society of Anaesthesiologists' score, Geriatric Nutritional Risk Index, abdominal/back pain, non metastatic liver disease or insulin resistance (n=4)		Multivaraite Cox proportional hazards regression	Calibration: Hosmer-Lemeshow 0.403. Discrimination: R^2 53.5%, AUC: 88.7%. External Validation performed
Smith & Mezhir (2014)	National Registry Database (n=6400)	6 months, 1,3,5 year survival	Backward stepdown selection process	Survival: age, gender, marital status, race, grade, histology, T&M, size, radiation, Lymph Node Ratio (n=11). Lymph node ratio: grade T&M stage, size (n=4)	Complete case analysis	Bayesian Model	Calibration curve: goodness-of-fit statistic 0.4847 Discrimination: *c*-statistic: 0.65 External Validation: absent; internal validation by random split

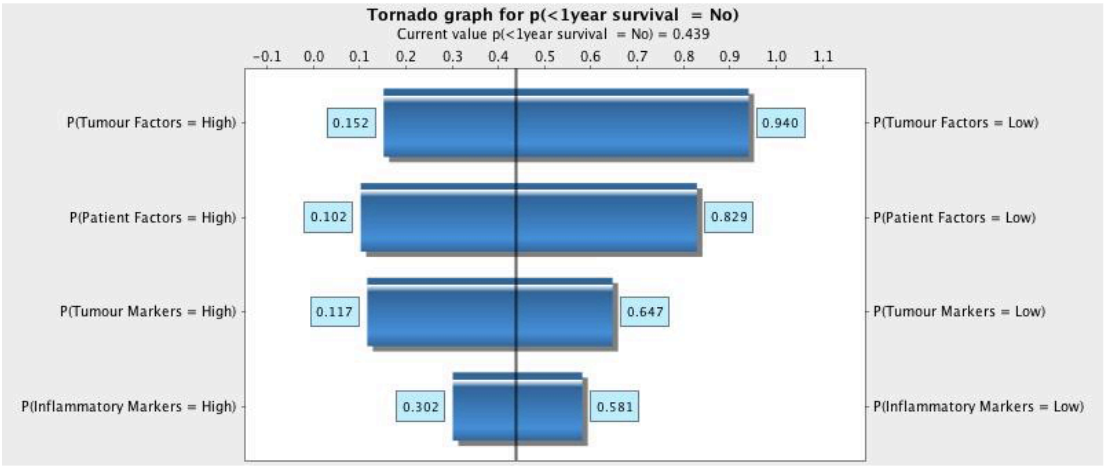
							method
Pu et al., (2017)	Single institution database (n=220)	1,2,3 year survival	Pre-selection by Univariate analysis then Cox multivariate analysis. P-value <0.05	Differentiation, American Joint Commission on Cancer stage, Alkaline Phosphate to Albumin Ratio (n=3)	Complete case analysis	Multivariate Cox proportional hazards regression	Calibration curve: optimal consistency Discrimination: training: 0.673 validation: 0.693 External validation: absent; internal validation by random split method
Dasari et al., (2015)	Single institution database (n=567)	1,3 year survival	Pre-selection by Univariate analysis then Cox multivariate analysis. P-value <0.05	Tumour site, T stage, Lymph Node Ratio (n=3)	Complete case analysis	Multivariate Cox proportional hazards regression	Calibration: none Discrimination: AUC 1yr &3yr: 0.66 & 0.74 External Validation performed
Pu et al., (2018)	National Registry Database (n=3458)	1,3,5 year survival	Pre-selection by Univariate analysis then Cox multivariate analysis. P-value <0.05	Age, grade, T stage (n=3)	Complete case analysis	Multivariate Cox proportional hazards regression	Calibration curve: optimal fit Discrimination: *c*-statistic 0.63 External and internal validation performed using bootstrap method
Katz et al., (2012b)	National Registry Database (n=5736)	3 year survival	Pre-selection by Univariate analysis then Cox multivariate analysis. P-value <0.05	Age, gender, race, site, grade, stage, radiotherapy (n=7)	Complete case analysis	Multivariate Cox proportional hazards regression	Calibration curve: results not reported Discrimination: absent External Validation: absent

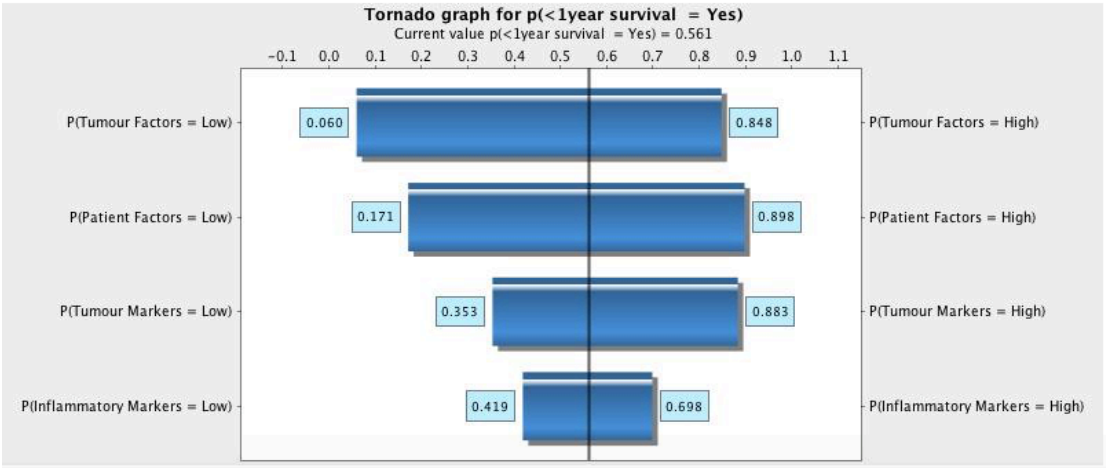
Appendix Q

Bayesian Belief Network Sensitivity Analysis

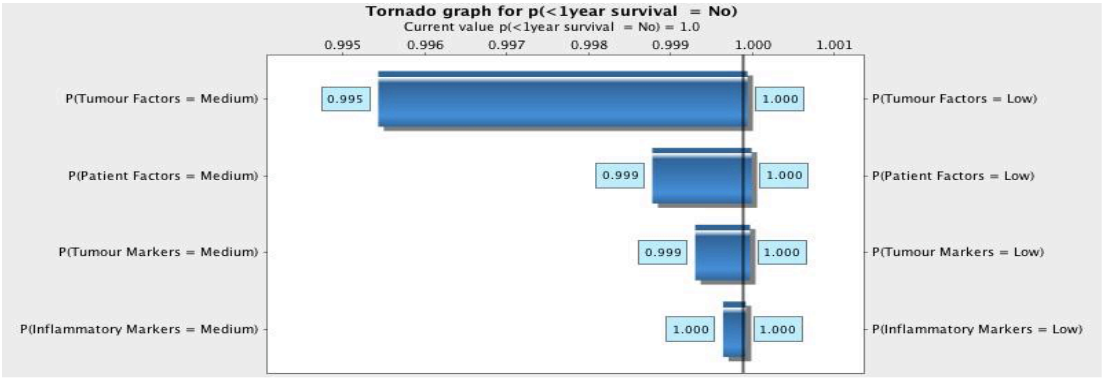
Figure Qi: Sensitivity Analysis

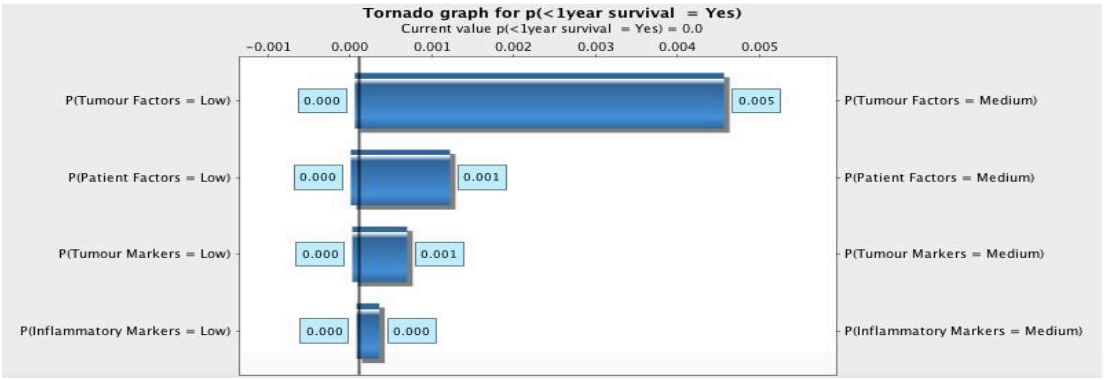
a) Pre-operative poor prognosis: Scenario 2



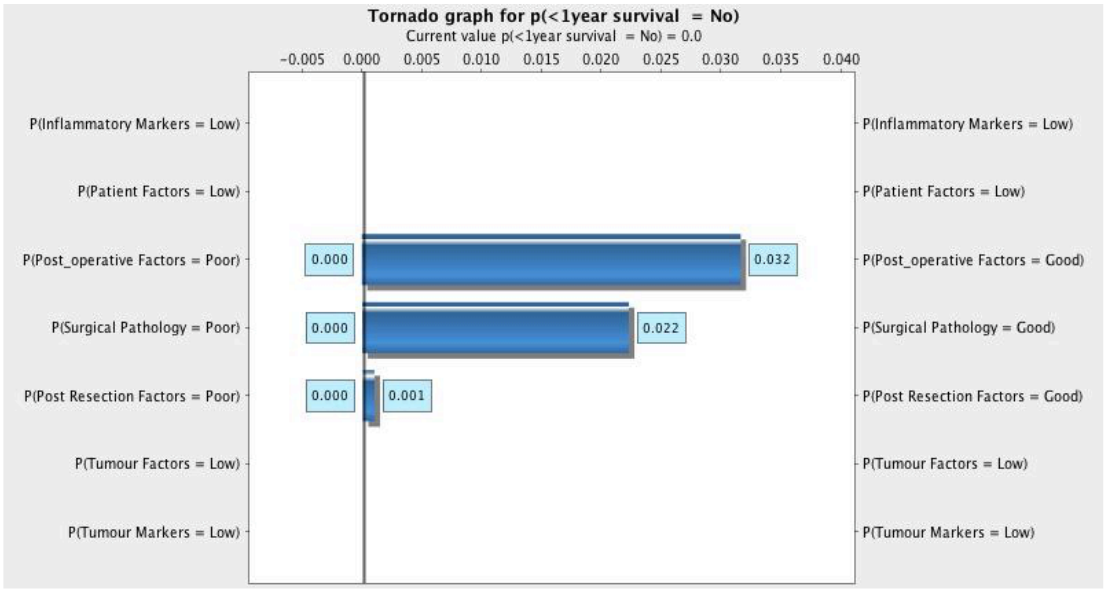


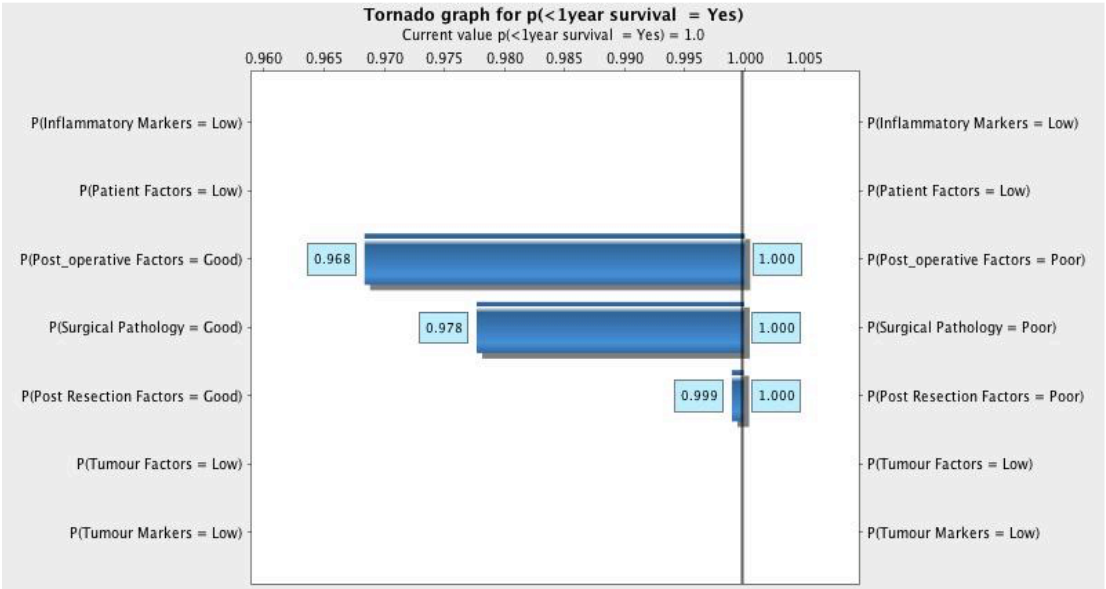
b) Pre-operative poor prognosis: Scenario 3



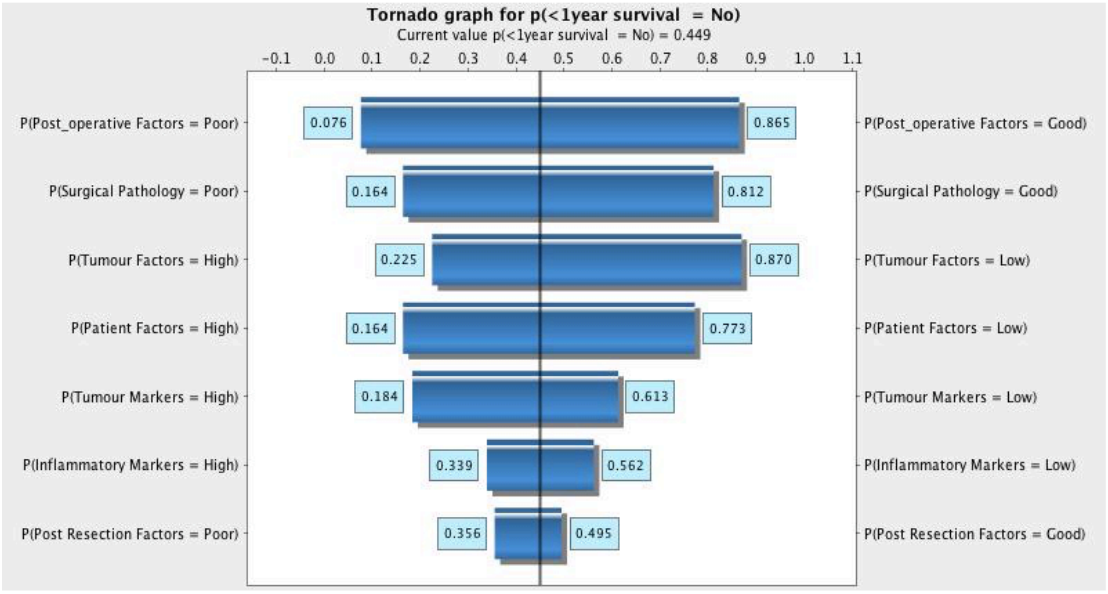


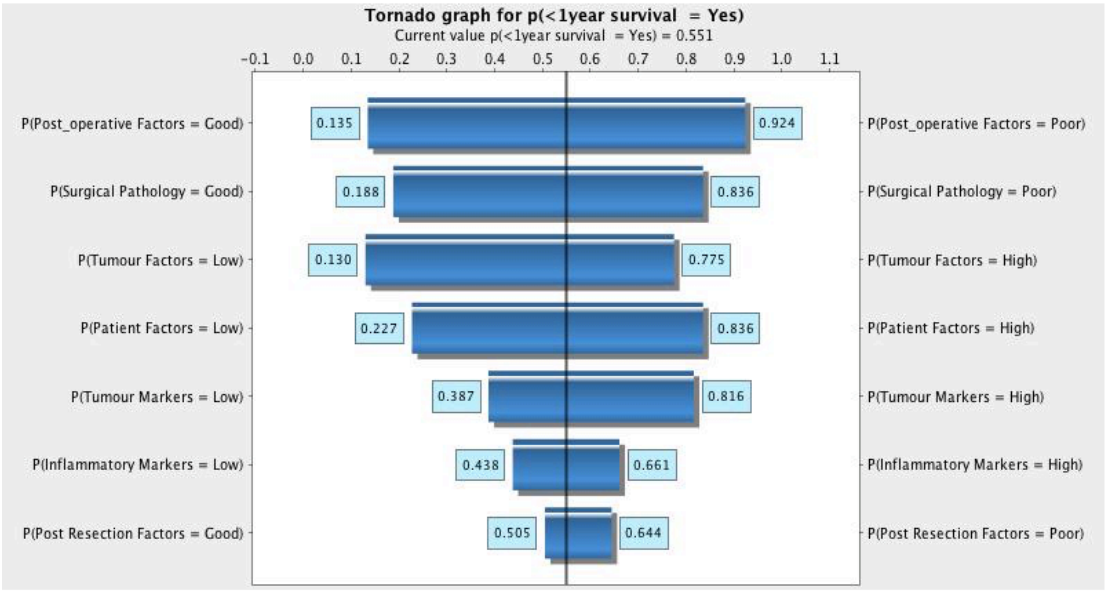
c) Post-operative poor prognosis: Scenario 1



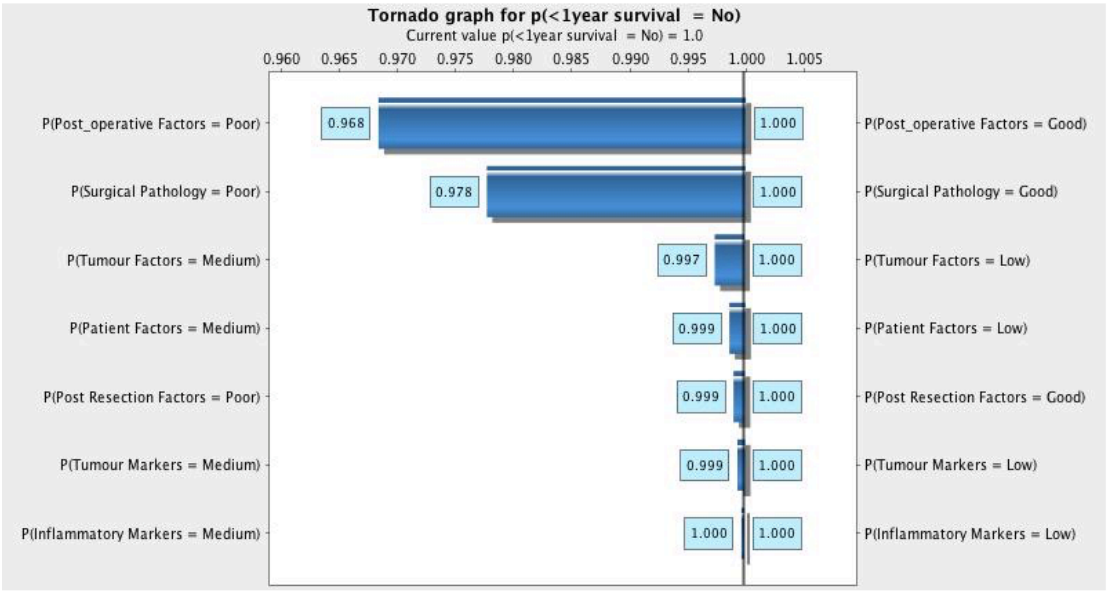


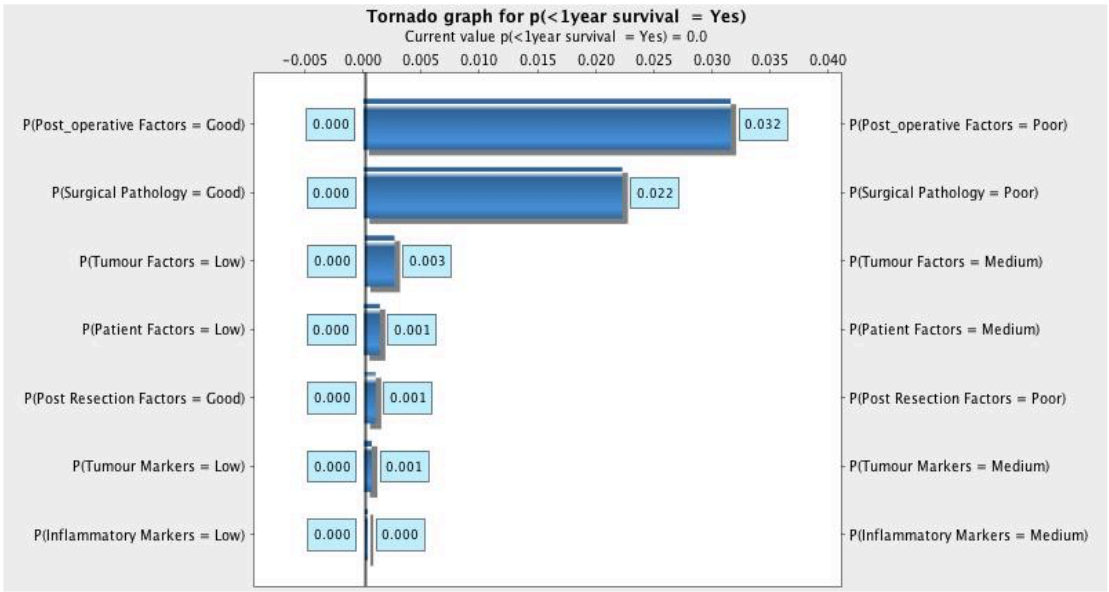
d) Post-operative poor prognosis: Scenario 2



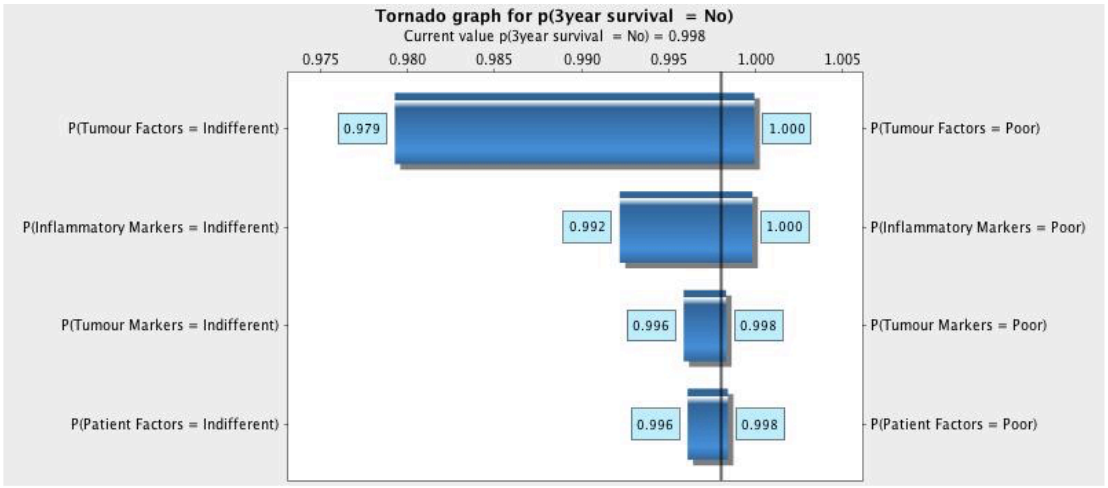


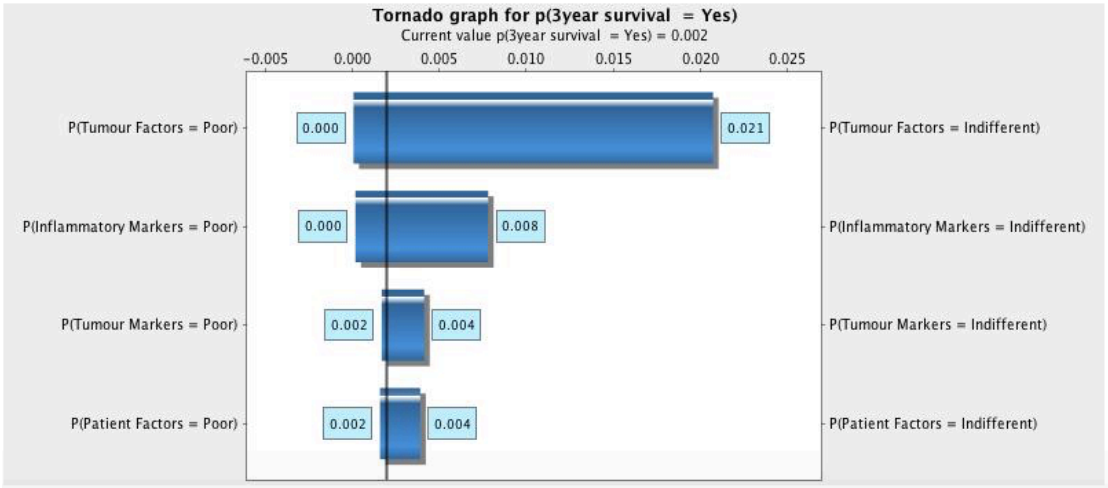
e) Post-operative poor prognosis: Scenario 3



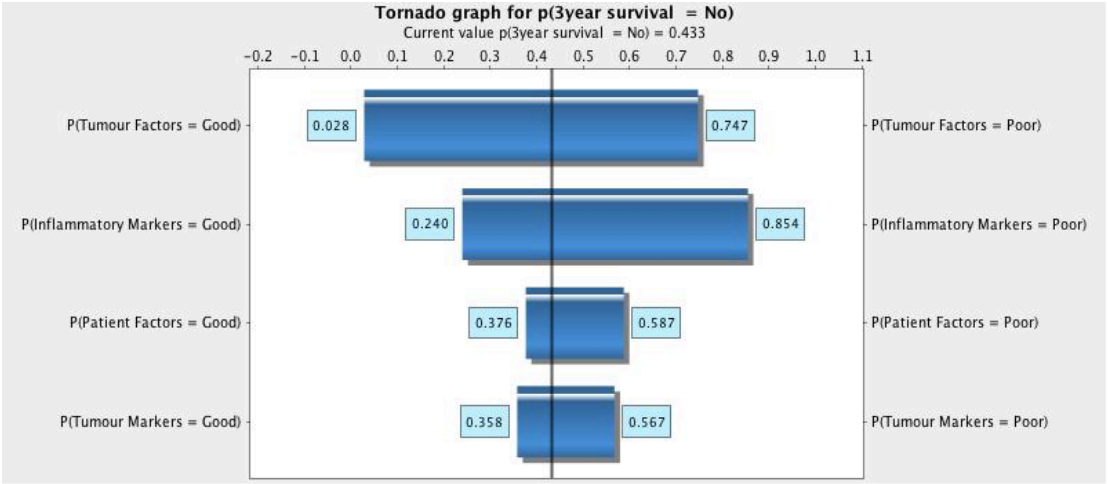


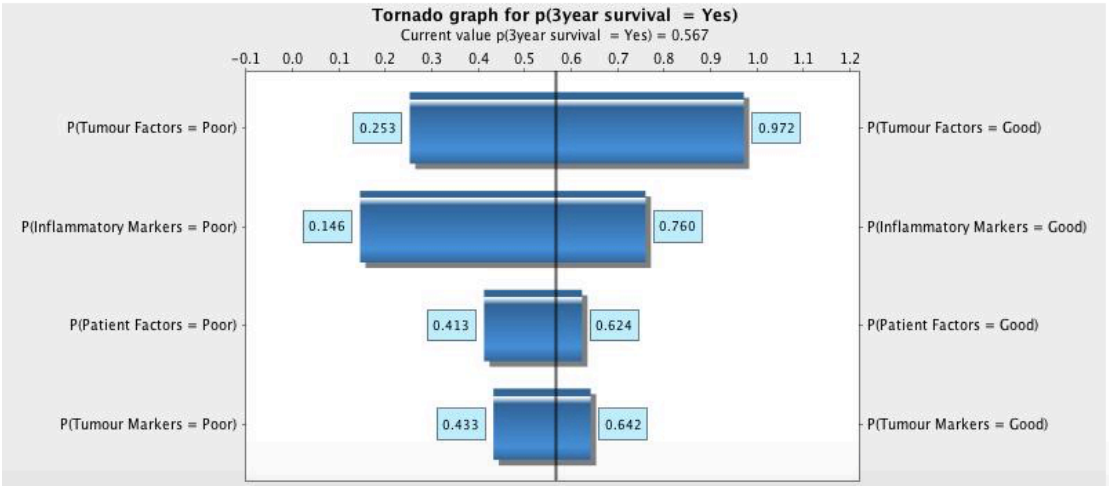
f) Pre-operative good prognosis: Scenario 1



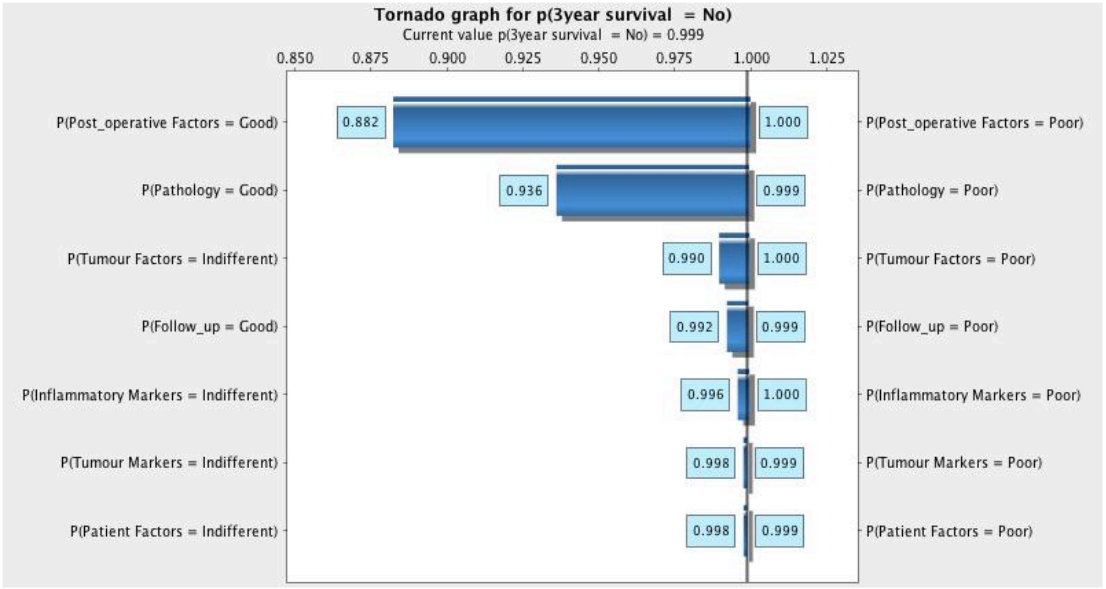


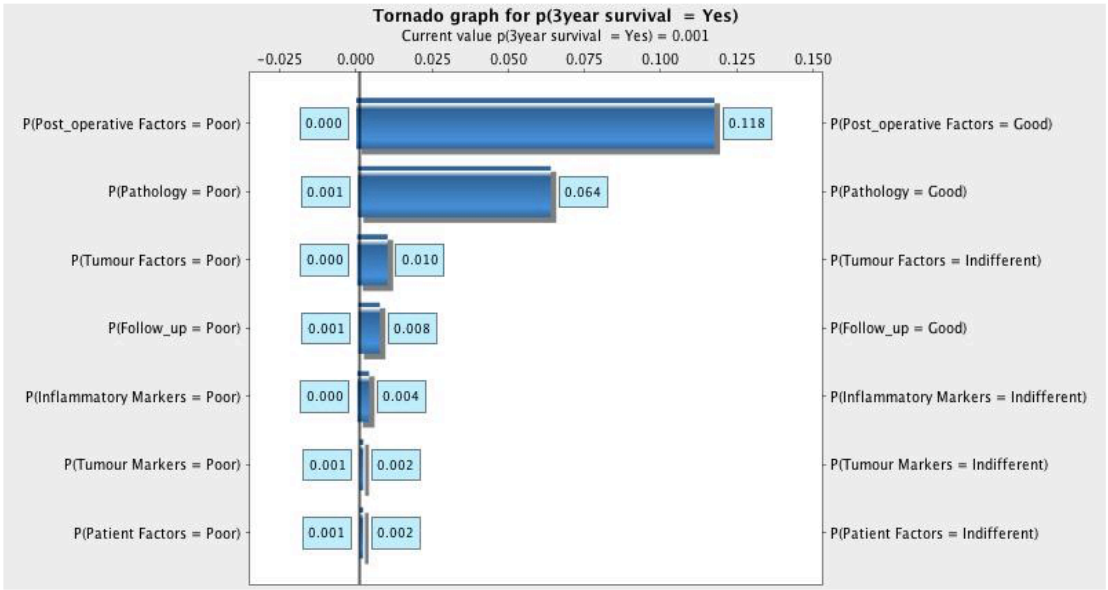
g) Pre-operative good prognosis: Scenario 2



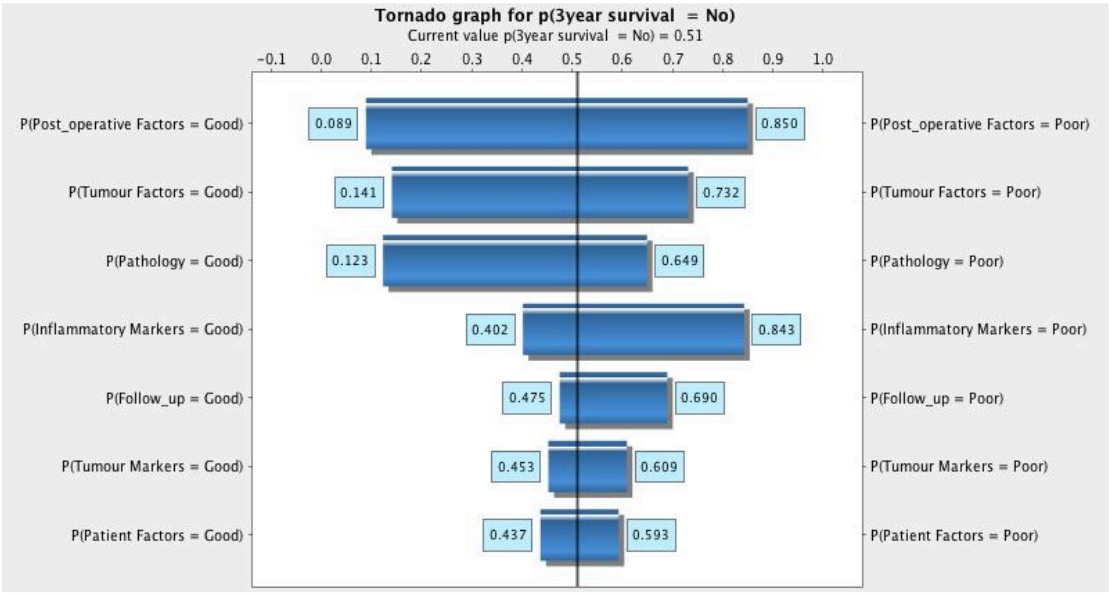


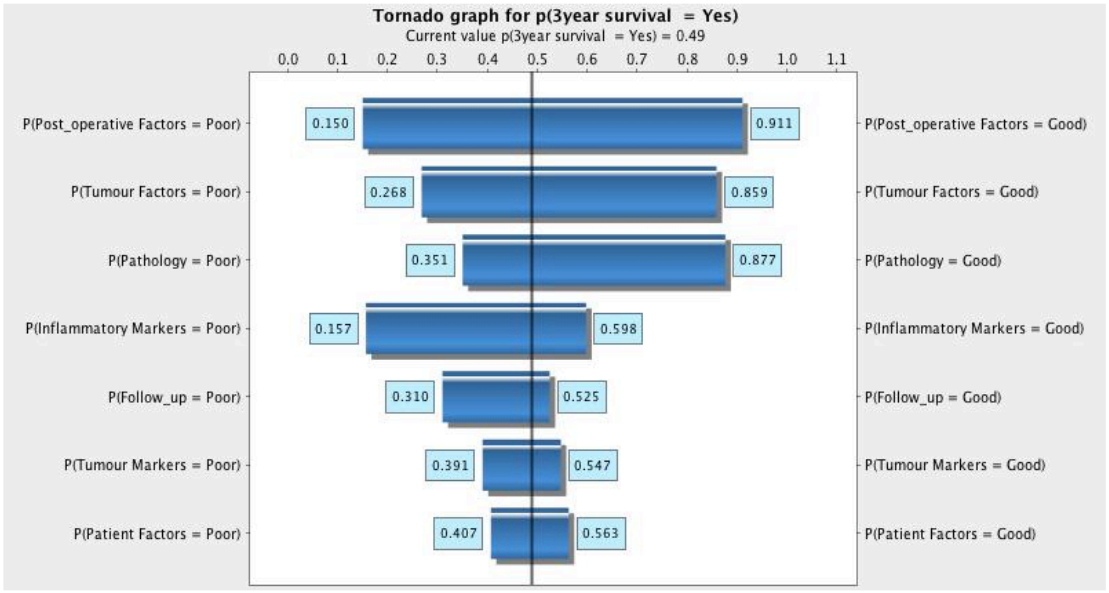
h) Post-operative good prognosis: Scenario 1



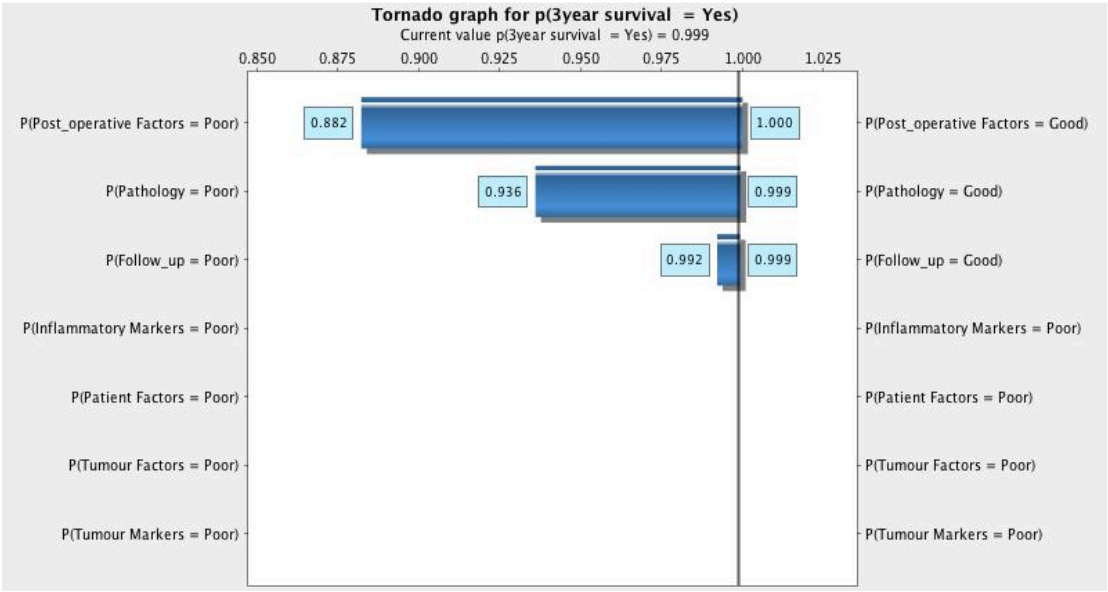


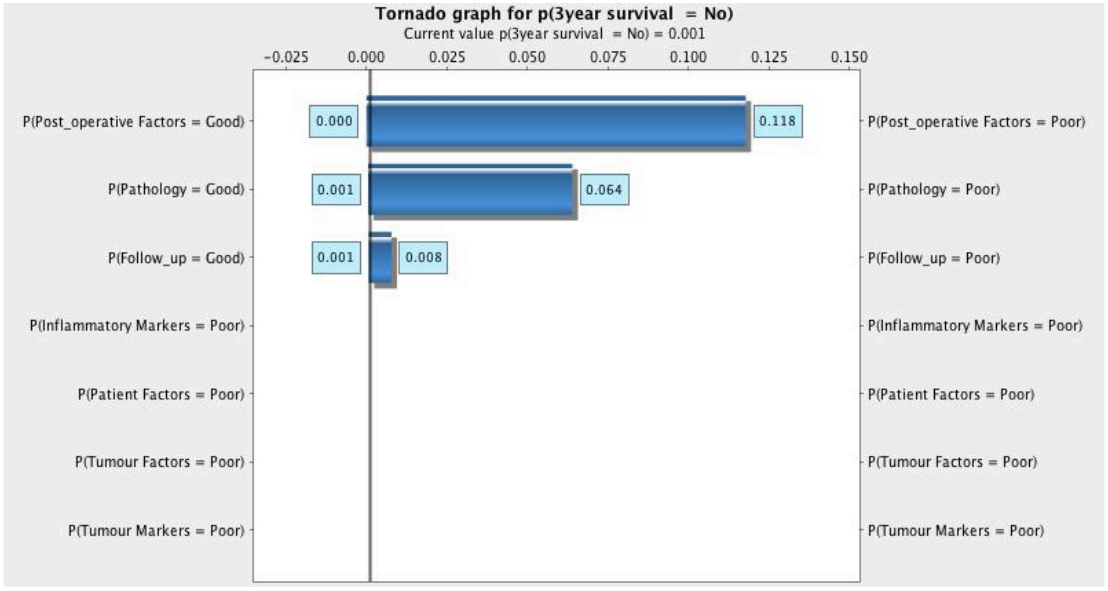
i) Post-operative good prognosis: Scenario 2





j) Post-operative good prognosis: Scenario 3





Appendix R

TRIPOD Checklist for Bayesian Belief Network

Section/Topic			Checklist Item	Page
Title and abstract				
Title	1	D;V	Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted.	430
Abstract	2	D;V	Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions.	430-432
Introduction				
Background and objectives	3a	D;V	Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models.	432-438
	3b	D;V	Specify the objectives, including whether the study describes the development or validation of the model or both.	432-438
Methods				
Source of data	4a	D;V	Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable.	439-440
	4b	D;V	Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up.	439-442
Participants	5a	D;V	Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres.	439-440
	5b	D;V	Describe eligibility criteria for participants.	439-440
	5c	D;V	Give details of treatments received, if relevant.	430-440
Outcome	6a	D;V	Clearly define the outcome that is predicted by the prediction model, including how and when assessed.	430-440
	6b	D;V	Report any actions to blind assessment of the outcome to be predicted.	NA
Predictors	7a	D;V	Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured.	440-452
	7b	D;V	Report any actions to blind assessment of predictors for the outcome and other predictors.	NA
Sample size	8	D;V	Explain how the study size was arrived at.	440-452
Missing data	9	D;V	Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method.	442
Statistical analysis methods	10a	D	Describe how predictors were handled in the analyses.	440-452
	10b	D	Specify type of model, all model-building procedures (including any predictor selection), and method for internal validation.	430-452
	10c	V	For validation, describe how the predictions were calculated.	440-456
	10d	D;V	Specify all measures used to assess model performance and, if relevant, to compare multiple models.	465-458
	10e	V	Describe any model updating (e.g., recalibration) arising from the validation, if done.	NA
Risk groups	11	D;V	Provide details on how risk groups were created, if done.	NA
Development vs. validation	12	V	For validation, identify any differences from the development data in setting, eligibility criteria, outcome, and predictors.	465-458
Results				
Participants	13a	D;V	Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful.	459-465
	13b	D;V	Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome.	459-465
	13c	V	For validation, show a comparison with the development data of the distribution of important variables (demographics, predictors and outcome).	NA
Model development	14a	D	Specify the number of participants and outcome events in each analysis.	459-465
	14b	D	If done, report the unadjusted association between each candidate predictor and outcome.	459-465
Model specification	15a	D	Present the full prediction model to allow predictions for individuals (i.e., all regression coefficients, and model intercept or baseline survival at a given time point).	440-465
	15b	D	Explain how to the use the prediction model.	440-

				465
Model performance	16	D;V	Report performance measures (with CIs) for the prediction model.	459-465
Model-updating	17	V	If done, report the results from any model updating (i.e., model specification, model performance).	NA
Discussion				
Limitations	18	D;V	Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data).	465-467 Chapter 5
Interpretation	9a	V	For validation, discuss the results with reference to performance in the development data, and any other validation data.	459-467
	9b	D;V	Give an overall interpretation of the results, considering objectives, limitations, results from similar studies, and other relevant evidence.	465-467
Implications	20	D;V	Discuss the potential clinical use of the model and implications for future research.	465-467, Chapter 5 &6
Other information				
Supplementary information	21	D;V	Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets.	Appendix P, Q, R
Funding	22	D;V	Give the source of funding and the role of the funders for the present study.	NA