University of Strathclyde

Department of Mathematics and Statistics

Doctor of Philosophy

A statistical analysis of spatially linked time series using data from the TOXBASE database to study Emergency Department and NHS phone line management of poisoned patients

---

Author:

Kate Pyper

Declaration of Authenticity and Author's Rights

This thesis is the result of the author's original research. It has been composed by the author and has not been previously submitted for examination which has led to the award of a degree.

Date:

Signed:

Acknowledgements

I would first like to thank my superviser Professor Chris Robertson for his valuable critiques of my work throughout this whole project, which have helped me to develop my skills as a reseracher. I would also like to thank my external project supervisers Professor Michael Eddleston and Professor Nick Bateman, who provided me with their invaluable knowledge and experience of Toxicology and poisoning epidemiology. I would like to thank David Lupton for giving his time to help with data cleaning, and Gill Jackson for her wisdom about the use of the TOXBASE system.

I would like to extend my appreciation to the admin and support staff in the department, all of whom were supportive and happy to answer any questions about procedure that I may have had.

I would not have been able to complete the project without the support provided by the other PhD students within the department, who were always happy to lend a sympathetic ear when it was needed.

Finally I would like to thank my family and friends for their support throughout this project.

Abstract

The aim of this thesis is to examine the overall trends in poisoning, which will lead to an assessment of the relationship between poisons information database TOXBASE use and the management of poisoned patients within UK based emergency departments.

Previous studies on the demographics of poisoning will be presented. This will cover important concepts in the study of toxicology before examining how specific demographic variables are linked to incidences of poisoning. A discussion of the agents used in cases of poisoning in recent history will be followed by a summary of legislation pertaining to poisoning and dangerous substances. The various services examined throughout the thesis will also be introduced.

The first of these services, NHS 24, is described as an out of hours service for use by members of the public. The NHS 24 operator can advise on appropriate action given a description of the symptoms. These descriptions can be categorised as poisoning based on the presence of specific key words or phrases. Using generalised additive models, a consistent seasonal trend in poisoning calls to NHS 24 was found.

The second service examined, TOXBASE, is a database provided by the National Poisons Information Service which provides information to clinical professionals on how to treat poisoning by a variety of substances. As in the NHS 24 analysis, generalised additive models have been used in order to assess the trends

present in accesses made to the TOXBASE database by clinicians. The results from this analysis found that there was a consistent seasonal trend in TOXBASE accesses which peaked over summer and was similar to that seen in the NHS 24 call data.

A third temporal analysis was carried out on data obtained from NHS information services pertaining to admissions and attendances due to poisoning, again showing similar results to the previous two analyses. These analyses combined suggest an underlying trend in poisonings.

Both the admissions data and TOXBASE access data were examined using funnel plots in order to determine whether there were any hospitals which were unusual in their admissions or their TOXBASE use. This analysis found some commonalities in those hospitals which are unusual in either their admission rates or TOXBASE use.

The final step in this project was to link the TOXBASE access data with the attendances and admission data in order to examine whether there was any link between TOXBASE use and admission due to poisoning. The results of this indicate that there are associations between TOXBASE use and admissions, such that an increase in TOXBASE use indicates an increase in admissions due to drugs poisoning. However, it became clear that the data used were limited in their ability to show any direct impact of TOXBASE on admissions due to drug poisoning, and that more specific data, for example on toxicants involved and case severity, would potentially be useful in mitigating the obvious confounding present in these data.

This thesis has provided new insight into patterns in cases of poisoning, as well as providing a strong basis for further analysis to establish whether there is a direct impact of TOXBASE use on patient management within UK emergency departments.

Contents

Chapter 1

Introduction

Poisoning and potential toxic exposure are major causes of attendances at emergency departments across the UK. These cases vary in both cause and severity, with more toxic chemicals and larger doses causing more severe symptoms. Cases of poisoning occur in all age groups, and the substances involved in these cases range from plants and animals, to household products, pharmaceuticals and recreational drugs.

Toxic exposure poses a significant workload for healthcare professionals. In all, these cases account for approximately 170,000 admissions to hospital per year in the UK, in addition to those patients who are discharged after treatment at an emergency department and those whose enquiries go to either of the NHS telephone services [1]. Each admission comes at a cost to the National Health Service. The estimated annual cost for the management of paracetamol poisoning, which is currently the most common in the UK, is £48.3 million [2]. It is estimated that each attendance at an A&E department for paracetamol overdose costs £137, with each admission costing an additional £631 [2]. In addition, a study carried out in Queens Medical centre in Nottingham estimated the annual cost of self-poisoning episodes at that hospital to be over £1.6 million [3].

In order to aid in the treatment of poisoning, the UK's National Poisons Information Service (NPIS) provides and maintains TOXBASE, a database outlining the appropriate management of cases of poisoning by a wide range of substances. The main aim of this work is to use routinely collected data in order to address the question of whether appropriate usage of this tool has a positive impact on the treatment of poisoning.

This chapter will provide some of the background information which was used to influence decisions made throughout this project. The initial section will introduce some important toxicological concepts, and make clear the distinction between some of the terminology used throughout this thesis. This will be followed by a discussion of some of the literature currently published on poisonings and exposures, focusing mainly on UK studies. This discussion will outline some of the key risk factors which have been found to correlate with incidences of poisoning. This will lead into a discussion of how toxic exposures have developed over the past century, including some of the specific legislation that has given rise to these changes. The more general legislation concerning dangerous substances will then be introduced with particular focus on European and British legislation.

All of this will lead to the introduction of the NPIS and the provision of poisons information in the UK. The TOXBASE database will then be introduced, with a brief overview of some of the research carried out on its usage. Introduction will then be given to two services which play an important role in the initial assessment of poisoned patients: the NHS telephone hotlines and emergency departments and some thought will be given to how these systems handle incidences of poisoning. The chapter will then conclude with an outline of the work presented in this thesis.

1.1 Important Concepts in Toxicology

Toxicology is defined as the study of the adverse effects of chemicals on living organisms [4]. Modern toxicology emphasises that any chemical can cause a toxic response under sufficient exposure. Even commonplace substances such as table salt and drinking water can be harmful with a large enough dose. It is for that reason that there exist multiple definitions in the toxicology literature to define varying degrees of exposure.

1.1.1 Exposure versus Poisoning

An exposure is defined to be any instance where a person comes into contact with a substance with the potential to cause harm. This can be, among other things, through ingestion, inhalation or skin contact. An exposure does not necessarily imply a negative response to said substance. There are two different types of exposure: potential and proven.

A potential exposure is defined to be a case where it is thought that an exposure has occurred, when in reality it may not have. An example of this would be finding a child with an empty pill bottle, but the contents of the bottle are later found to have rolled under a piece of furniture. Naturally the parent would panic and assume exposure had occurred when in reality it had not.

A proven exposure is where there has been definite contact with a harmful dose of some substance, but a detrimental effect is not necessarily observed. This could occur, for example when someone takes more than the recommended dose of a particular medication, but they have developed a tolerance to this medication and do not experience a toxic effect. An incidence of poisoning is defined to be a proven exposure with clear detrimental effect.

1.1.2 Types of Poisoning

Self-Poisoning

A case of poisoning occurs where a person has been exposed to an amount of a substance which would reasonably be associated with the potential to cause harm. This can occur as a result of an external source emitting a potentially harmful substance. This can also be because a person either deliberately or accidentally exposes themselves to a large enough dose of a particular substance to cause harm, which is then referred to as self-poisoning.

Deliberate self-poisoning is a form of self-harm where the individual exposes themselves to a dangerous amount of a substance with the intention of causing themselves harm. Unintentional self-poisoning can occur when a person is not aware of the harmful effect of a substance, as is likely to be the case in very young children, or where the individual is unaware of the recommended dosage instruction. These cases are generally referred to as unintentional paediatric poisoning and therapeutic excess respectively.

Chronic versus Acute Poisoning

In cases of poisoning, symptoms may vary depending on the length of time of the exposure. Toxicological literature generally distinguishes between four different categories; acute, subacute, subchronic and chronic [4].

Acute poisoning is defined as exposure of less than 24 hours. These are cases that are commonly seen in emergency departments. In fact, acute self-poisoning is one of the most common medical presentations in the UK [5].

Subacute exposures occur after repeated exposure to a chemical over several days, or up to a month. Subchronic exposures are a result of repeated exposure to a substance for a period of between 1 and 3 months. Finally, chronic exposures generally occur in repeated doses over periods longer than 3 months [4].

1.1.3 The Dose-Response Relationship

The way in which the amount of substance involved in an exposure affects the response to an exposure is described by the dose-response relationship. Specifically it is the rate of response to a substance measured as a proportion. This relationship, assuming that the x axis is plotted on a log scale, is generally described by a sigmoidal shaped curve, where the response percentage is increased incrementally at small doses and as the dose increases the proportion increases more sharply before a point is reached where increasing the dose will only increase the probability of observing the effect of interest slightly.

Each substance has a unique threshold dose, as well as unique efficacy and potency. The threshold dose is the dose at which the probability of a symptomatic response becomes non-zero. In some substances, such as water, the threshold dose can be very high, while other, more toxic substances would have a lower threshold dose. Efficacy is the maximum percentage response that can be achieved by a substance, meaning that it is essentially a measure of how effective a substance is at producing the toxic response. Potency describes how much of a given substance is required to elicit a response. These concepts can be seen more clearly in Figure 1.1, where Drug A and Drug B are two substances with similar efficacy, but where Drug A has a higher potency. The Drug C has a similar potency to Drug B, but a lower efficacy [4].

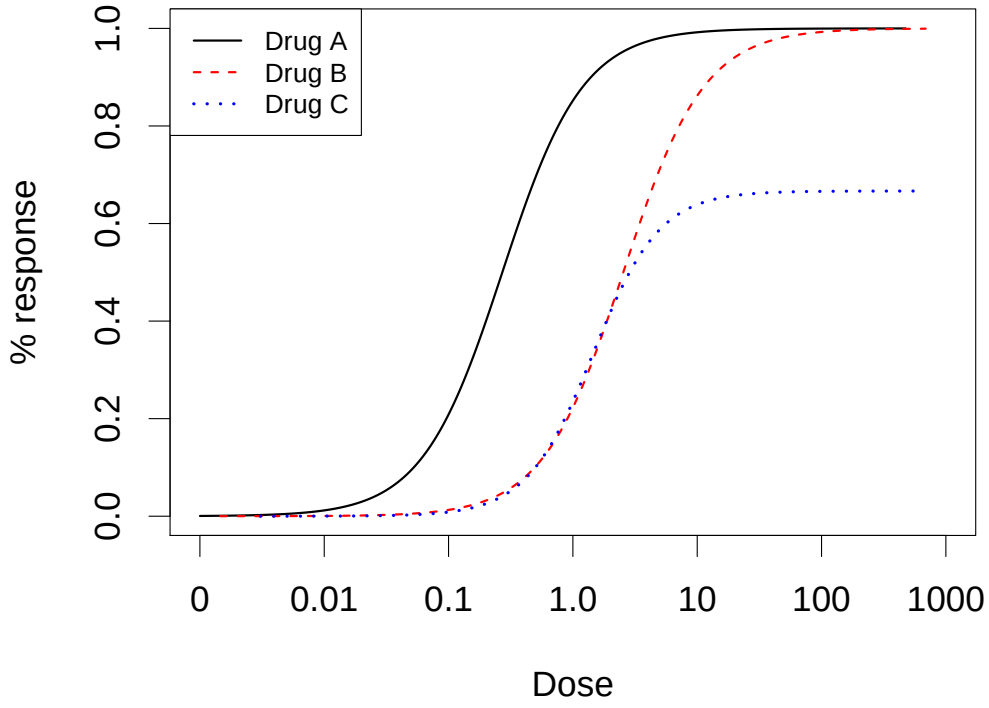


Figure 1.1: Plot showing the effect of efficacy and potency on the dose-response relationship

This relationship implies that an increase in dose will increase the toxicity. However the specific relationship described by the curves in Figure 1.1 assumes that the way in which people react to toxic substances is approximately normally distributed, which may not necessarily be the case. In some circumstances there can be a small subset of the population who are more vulnerable to a particular substance than can be described by a normal distribution, for example where individuals may have a rare trait which leads to excess vulnerability to certain substances.

1.1.4 Chemical Interactions

In addition to this dose-response relationship, there are some substances which interact with one another which may alter this relationship. These chemical interactions can have different effects and there are therefore several terms associated with this concept. The simplest of these is the additive effect, which occurs where the combined effect of two chemicals is equal to the sum of their individual effect, and this is most often what happens when two chemicals are given together. A synergistic effect, however, occurs where the combined effect of two chemicals is greater than the sum of their individual effects. Potentiation is a similar concept to synergy but potentiation means that one chemical does not have a toxic effect unless combined with another chemical.

The final concept to mention is antagonism. This is where the two chemicals administered interfere with each other, or one affects the action of the other. A specific instance of this, functional antagonism, is where two chemicals produce opposing reactions, leading to a balance of effects. This is a key concept in the study of poison antidotes [6].

1.2 Risk Factors for Poisoning

Incidences of poisoning have been found to be related with multiple risk factors. Studies will be presented which reflect the effects of gender, age, deprivation and rurality on incidences of poisoning.

1.2.1 Gender

Several studies have suggested that gender has an impact on the incidence of poisonings, particularly in the case of deliberate self-poisoning, which is thought to be more common in females than in males [7, 5]. One study examining the

demographics of poisoning admissions across three hospitals in Oxford, Leeds and Manchester indicated that in episodes of self-harm, females showed a greater rate of self-poisoning than did males [8]. The same study also looked at how the toxins associated with episodes of self-harm varied by gender. This indicated that the substances used varied by gender, with more females using paracetamol (including paracetamol combinations) and selective serotonin reuptake inhibitor (SSRI) and serotonin and norepinephrine reuptake inhibitors (SNRI) antidepressants than would be expected. The study also suggested that males were more likely than females to have consumed alcohol around the time of the self-harm incident.

A study carried out on cases of self-poisoning in Queens Medical Centre in Nottingham also found a relationship between gender and risk of poisoning. The ratio of females to males overall was found to be 1.45:1, with a higher ratio of almost 3:1 in the 16-20 age group [3].

Another study [9] showed similar findings. The relative risk of general poisoning in females compared to males was approximately 2, indicating that females are around twice as likely to be involved in an incidence of poisoning compared to males. This figure increased when considering cases specifically relating to self-poisoning, with females being more than 3 times more likely than males present with self-poisoning. In cases of unintentional poisoning it was found that females were on average 23% more likely than males to be involved. Further to this, the same study indicated that patterns in poisoning incidences by age differed between males and females, with female poisoning tending to peak at around 15 years of age compared to around the age of 20 in males.

1.2.2 Age

The impact of age on incidences of poisoning can also be seen across a wide range of studies, including some of those discussed previously, but age on its own can also contribute to the variability in poisonings.

One study indicated that deliberate self-poisoning was most prevalent among young adults between the ages of 15 and 35 [7]. The study carried out at Queens Medical Centre in Nottingham also found that the age distribution for self-poisoning patients was skewed towards younger patients. Patients under 30 were involved in 46.8% of cases, while only 3.3% of cases involved patients over 60 [3].

In addition to being an influencing factor in cases of self-poisoning in general, it has also been found that age may play a part in the likelihood of re-attendance for self-poisoning. One study suggested that individuals over 60 were more likely to reattend for self-poisoning than those in younger age groups [10].

There is some evidence that unintentional poisoning is most prevalent in children under 5, as children tend to explore their surroundings by putting things in their mouth. These are generally minor cases of poisoning, compared to episodes of deliberate self-poisoning. The rate of unintentional poisoning is seen to decrease in children over 5 as children become more informed about substances which could be harmful, meaning that accidental ingestion becomes less likely. The rate of poisoning then increases above the age of 10 as poisoning becomes more common in episodes of self-harm [7].

An English study focusing on hospital admissions due to unintentional poisoning in children younger than 5 indicated that admissions rates for poisonings have decreased overall between 2000 and 2011 [11]. The only products which proved an exception to this fell into the soaps and detergents category, for which admission rates had doubled over the study period.

This trend has not gone unnoticed. There was a reference to this in the 2014/15 National Poisons Information Service annual report where it was suggested that 96% of approximately 2,500 enquiries made regarding liquid laundry capsules concerned children. In addition to this, almost 93% of enquiries due to exposure to soluble film dishwashing tablets concerned children under 5. In both products the majority of cases were recorded as being minor, however a small proportion of these cases reported more severe indications of poisoning [1].

1.2.3 Deprivation

Deprivation is defined to be a damaging lack of material benefits considered to be basic necessities on society. This definition is relatively subjective meaning that it can be difficult to measure. There have been several proposed measures of deprivation, but the two most commonly used in the UK at the time of writing are the Scottish Index of Multiple Deprivation (SIMD) [12] and the Index of Multiple deprivation (IMD) [13], which have replaced earlier measures such as the Townsend Score [14] and the Carstairs Deprivation Index [15]. These scores attempt to evaluate deprivation across Scotland and England respectively.

These two measures are computed in similar but slightly different ways. Both scores take into account several factors across seven broad categories: employment, income, crime, housing, health, education and access to services. However the titles and contents of each of these categories differ between the two measures (Table 1.1 & Table 1.2). Although there are differences in how they are constructed, both measures place the most weight on factors in the Income and Employment categories.

Table 1.1: Table showing the indicators in each category of the SIMD

SIMD						
Employment	**Income**	**Crime**	**Housing**	**Health**	**Education**	**Access**
Unemployment Claimant Count averaged over 12 months	Income Support and Income-based Employment	Domestic House Breaking	Persons in households which are overcrowded	Standardised Mortality Ratio	School pupil attendance	Drive time to GP, to retail centre, to petrol station, to primary and secondary schools, to post office
Working age Incapacity Benefit or Employment Support Allowance recipients	Support Allowance claimants (16-59)	Drug Offences	Persons in households without central heating	Hospital stays related to alcohol misuse	School pupil performance	Public transport time to GP, to retail centre, to post office
Working Age Severe Disablement Allowance resipients	Job Seekers Allowance and Guaranteed Pension Credit Claimants (All ages)	Common Assault		Hospital stays related to drug misuse	Working age people with no qualifications	
	Universal Credit claimants with no employment marker.	Crimes of Violence		Comparative Illness Factor	17-21 year olds enrolling into full time higher education	
	Number of children in JSA, IS or ESA households	Vandalism		Emergency stays in hospital	School leavers aged 16-19 not in education, employment or training	
	Number of Adults and children dependent on adults in receipt of tax credits.	Sexual Offences		Proportion of population being prescribed drugs for anxiety, depression or psychosis		
				Proportion of live singleton births of low birth weight		

Table 1.2: Table showing the indicators in each category of the IMD

IMD						
Employment	**Income**	**Crime**	**Living Environment**	**Health**	**Education, Skills & Training**	**Barriers to Housing & services**
Claimants of Jobseeker's Allowance	Adults & children in Income Support families	Violence	Housing in poor condition	Years of potential life lost	Key stage 2 attainment	Road distance to post office; primary school; general store or supermarket; GP surgery
Claimants of Employment and Support Allowance	Adults & children in Income-based Jobseeker's Allowance families	Burglary	Houses without central heating	Comparative illness and disability ratio	Key stage 4 attainment	Household overcrowding
Claimants of Incapacity Benefit	Adults & children in Income-based Employment and Support Allowance families	Theft	Air quality	Acute morbidity	Secondary school absence	Homelessness
Claimants of Severe Disablement Allowance	Adults & children in Pension Credit (Guarantee) families	Criminal damage	Road traffic accidents	Mood and anxiety disorders	Staying on in education	Housing affordability
Claimants of Carer's Allowance	Adults & children in Child Tax Credit and Working Tax Credit families not already counted				Entry to higher education	
	Asylum Seekers in England in receipt of subsistence support, accommodation support or both				Adults with no or low qualifications	

There is some evidence to link the rate of poisoning with deprivation scores. In Scotland, a study of admissions to the Royal Infirmary of Edinburgh between 1981 and 2001 [16] indicated that the rate of admissions after a case of poisoning was higher for patients living in more deprived areas as measured by the Carstairs Deprivation Index. An additional study on poisoning presentations at the emergency department of Ninewells Hospital in Dundee indicated that patients seen tended to live in more deprived areas [17].

In contrast, a study carried out on poisoning admissions to Aberdeen Royal Infirmary found no indication that socio-economic status (based on SIMD) had an impact on poisonings [18]. It is, however, noted in this publication that this does not reflect trends seen in other centres and that the results may have been affected by the number of temporary residents in the study whose deprivation level could not be assessed.

Elsewhere in the UK, evidence of the link between deprivation and poisoning risk can be found in a study carried out by Tyrrell et al [9], where data from hospitals in England also indicated that there was an overall effect of socio-economic status on the relative risk of poisoning in adolescents. Specifically, the risk for intentional poisoning is approximately doubled in the most compared to the least deprived area and in unintentional poisonings the risk is approximately 50% higher for most compared to least deprived areas.

Two studies which investigated poisonings in children under five also found that the rate of admissions due to unintentional poisonings was higher for those children living in deprived areas [11, 19]. Specifically, the study found evidence that in 2011, children in the most deprived quintile were around 1.5 times more likely to present with a case of poisoning [11]. This represents a reduction on

the relative risk found by a study carried out in 2006, where children in the least deprived quintile were almost 2.5 times as likely to present with a case of poisoning [19].

The earlier of these studies found, not only that the overall risk of poisoning in young children varies by deprivation, but there is variability in the effect of deprivation on risk depending on the substance used in the poisoning incident [19]. The results from this study indicated that children in the most deprived third of the population are almost 6 times more likely than those in the least deprived third to present with poisoning by Benzodiazepines, a type of sedative. This is in contrast to poisonings where the substances involved were caused by one of non-opioid analgesics, antipyretics and antirheumatics, non-steroidal anti-inflammatories and salicylates. Children in the most deprived third of the population were found to be less than twice as likely to present with poisoning by one of these substances compared to those in the least deprived third.

Due to the complex nature of deprivation scores and the way in which they are calculated, there may be some uncertainty about whether these links with poisoning are centred around one particular area of deprivation as opposed to deprivation as a whole. As an example, it has been found that unemployment and previous criminal record, both of which are related to deprivation scores, have an impact on the risk of deliberate self-poisoning [20]. However, it seems reasonable to investigate deprivation as a whole, as a way of accounting for as much variability in poisoning risk under with one value, as opposed to many.

1.2.4 Drug Misuse

Recreational drug users are often described as a hidden population, meaning that they are unlikely to inform others that they belong to this group. For this reason, it is difficult to quantify the relative risk of drug misuse on poisoning, in turn mak-

ing it difficult to obtain data on people from this demographic. As such poisoning cases related to recreational drug use are likely to be under-reported, especially those incidents which may be caused by drug misuse but can be attributed to another cause. However, by definition, one would expect that recreational drug users would exhibit more complications due to drug toxicity than the general population [21]. A report on drug related deaths in Scotland estimated that the rate of drug deaths among problem drug users was 9.8 (per 1,000 problem drug users), which is unsurprisingly much higher than the estimated rate of drug death of 0.13 (per 1,000 population) for the general population [22]. Further to this it has been found that there is a link between drug abuse and repeated deliberate self-poisoning [20].

The Office of National Statistics produces an annual report on drug deaths each year. The focus of the 2015 report was deaths due to drug misuse, which the ONS define to be a death where the underlying cause is drug abuse or drug dependence or a death where the underlying cause is drug poisoning and where any of the substances involved are controlled under the Misuse of Drugs Act (1971) [23]. This indicates that there has been an overall increase in deaths due to the misuse of drugs that are commonly abused since 1993. Over the past few years this increasing trend has been more obvious in males compared to females. In contrast to the general trend in cases of poisoning, the report indicates that deaths due to drug misuse are more common in males than in females. It is also suggested that drug misuse deaths are most commonly due to accidental overdose. However, the proportion of drug misuse deaths which are reported as suicide is seen to be higher in females than in males, which is consistent with the gender patterns described previously.

A report produced by the Information Services Division (ISD) of the NHS in Scotland on drug related hospital statistics examines hospital admissions rather than deaths. However, this report indicated similar risk factors for poisoning due to drug misuse [24]. The report suggested that cases involving drug misuse are higher in males than in females. Further, the results in the Scottish report correlate with the results in the report for England and Wales, with evidence supporting that the majority of drug misuse incidents involve opioids.

There was also evidence in the ISD report that, while admissions due to drug misuse had remained stable in younger age groups, admissions rates were increasing in the older population. The example given for this is that the rate of admission among 40-44 year olds has increased from 20 patients per 100,000 of the population in 1996/1997 to 291 patients per 100,000 of the population in 2015/16. Finally, the ISD report indicated that drug misuse was higher for those areas which are more deprived.

Cases of poisoning due to recreational drug use are potentially problematic from the perspective of record keeping. A study carried out in 12 clinical coding departments in England and Wales NHS Trusts examined how different people coded different scenarios [25]. The researchers sent 12 hypothetical discharge summaries, 9 of which had a toxicological presentation, and asked the participants to assign a diagnosis code. The study indicated that those cases for which the presentation was non-toxicological were found to have little variation in the recommended diagnosis codes. The toxicological presentations, however, produced much larger variation in the suggested diagnoses, with the exception of those concerning alcohol intoxication and toxic symptoms of heroin and ecstasy. The researchers suggest that this may result in poisoning due to recreational drug toxicity being under-recorded, which in turn means that it is difficult to get an accurate representation of the epidemiology of recreational drug toxicity.

1.3 Changing Epidemiology of Poisoning in the UK

Within the last century, there have been changes in the chemicals used in poisoning incidents. In the 1950s, for example, a large proportion of self-poisoning incidences involved inhalation of coal gas, which was widely available at the time [5]. Incidence of self-poisoning using coal gas decreased as the carbon monoxide content within these products was reduced, and eventually the product was withdrawn from the market in favour of gases containing lower levels of carbon monoxide [5]. This is an example of where limiting access to a specific chemical has reduced rates of poisoning by that specific substance. This method plays on the impulsive nature of suicide attempts, where the availability of a particular method of suicide is seen to be a key factor in influencing the method used in a suicide attempt [26].

More recently a large number of self-poisonings in the UK involve paracetamol ingestion. It is estimated that between 82,000 and 90,000 patients present at emergency departments with paracetamol poisoning across the UK each year [2]. This is likely a reflection of the availability of paracetamol in the average household.

Paracetamol products are widely available in many forms over the counter, although legislation was brought into place in 1998 in an attempt to reduce the number of paracetamol overdoses. This legislation restricted the pack size of some over the counter preparations, meaning that paracetamol could only be bought in packs up to 16×500mg in non-pharmacy stores and pack sizes up to 32×500mg in pharmacies [27]. Further guidelines were released in 2009 stating that no more than two packs of paracetamol should be sold in one transaction and that it is illegal to sell more than 100 paracetamol tablets in one transaction.

There is conflicting evidence on whether this legislation change had a significant impact on the number of paracetamol overdoses. One study by Hawton et al [28] used mortality data from England and Wales and data from a liver transplant unit in England to show a reduction in both mortality and liver transplantation caused by paracetamol toxicity. Meanwhile studies by Bateman et al [29] and Newsome et al [30] carried out in Scotland indicate that the change in legislation had limited impact on the number and severity of paracetamol overdoses.

New Psychoactive Substances (NPS), more commonly known as legal highs, have been the subject of discussion in recent years, due to their increase in popularity. The term legal high covers several different types of substances, including cathinones and piperazines, all of which vary in their effect on the central nervous system. It is this that leads to difficulty in analysing the dangers of these types of drug. New Zealand was the first country to introduce legislation which specifically targeted NPS, with The Psychoactive Substances Act (2013) being almost unanimously supported [31]. This initially provided interim licenses for the sale of these products, before reports of adverse effects led to a blanket ban on the sale of legal highs.

At around the same time, Temporary Class Drug Orders (TCDOs) were being introduced in the UK [32]. TCDOs allowed the Home Secretary to place temporary restriction on substances which were not controlled under the Misuse of Drugs Act. Additional stipulations to a TCDO were that the substance was being misused and that this misuse was likely to be associated with having harmful effects.

The introduction of a TCDO has been found to have an impact on the number of cases relating to specific substances. The first TCDO in the UK was made for methoxetamine in 2012 [32]. The impact of this can be seen in Hill et al, where the introduction of the temporary order resulted in a large reduction in enquiries to

the National Poisons Information Service [32]. The National Poisons Information Service 2012/13 annual report suggested that after cathinone derivatives were made illegal, there was a sharp reduction in enquiries concerning mephedrone [33].

Subsequently, on the 26th May 2016, the Psychoactive Substances Act (2016) [34] was brought into effect. This legislation bans the production, sale and consumption of NPS. The act defines NPS to be "any substance which is capable of producing a psychoactive effect in a person who consumes it and is not an exempted substance". The act then states that a substance produces a psychoactive effect "if, by stimulating or depressing the persons central nervous system, it affects the persons mental functioning or emotional state". The term "exempted substances" refers to any drug which is already controlled by the Misuse of Drugs Act (1971) plus alcohol, caffeine, medicinal products, nicotine and tobacco products and food.

1.4 Legislation

1.4.1 UK Specific Legislation

The two pieces of legislation relating to paracetamol and NPS are both specific to the UK, and represent only a small proportion of the legislation which impacts the purchase and ownership of potentially harmful substances. Another such piece of legislation is The Misuse of Drugs Act (1971) [35] which enforces penalties on the possession, production and supply of controlled drugs. The penalties vary based on the offence committed and the classification of the substance in question. The Psychoactive Substances Act (2016) mentioned previously, supplements the Misuse of Drugs Act (1971) by placing restrictions on previously excluded substances.

Historically, dangerous substances have been regulated via a series of pieces of legislation. The Pharmacy Act (1852) established the existence of the Pharmaceutical Society of Great Britain, giving the body power to examine and certify its members [36]. Subsequently, the Pharmacy Act (1868) was brought into effect, introducing a Poisons List of 15 entries, which could be added to by the Pharmaceutical Society. These substances, or preparations containing these substances, could only be sold by registered pharmacists [36].

The 1868 Act also regulated the way in which poisons could be sold, with tighter restriction applied to those poisons thought to be more dangerous. The list of controlled poisons was extended by the 1908 Poisons and Pharmacy Act, which also updated its predecessor by including agricultural substances in the list of restricted products. The Act also stipulated that agricultural substances included on the Poisons List could be sold by any licensed retailer[36].

A further update came with the Pharmacy and Poisons Act (1933) which established a disciplinary body, which governs corporate bodies and pharmacists convicted of offences under the Pharmacy Act [36]. The Pharmaceutical Society was then authorised to appoint inspectors, whose job would be to enforce the Act. The 1933 Act was repealed under the Medicines Act (1968) [37], which replaced all previous legislation relating to medicine. However, restriction to the sale of medicines was limited to those medicines which contained substances listed in the Poisons List. For that reason, the quality of medicines was somewhat controlled by the Food and Drugs Act (1955) [36].

The Poisons Act (1972) [38] came into effect in order to regulate the sale of non-medicinal poisons. A simplification of the restrictions outlined in the Poisons Act (1972) was brought into effect by The Deregulation Act (2015) [39]. This legislation defines the difference between restricted and reportable substances. The Act states that a license must be held in order to use, possess, acquire or

import a regulated substance and that suppliers will also be held accountable for supplying regulated substances without checking first whether the recipient is a licence holder. The legislation also makes it an offence if a supplier does not report any transaction involving either a reportable or regulated substance where there are grounds to believe that the substance is intended for illicit use.

1.4.2 EU Legislation

Further to the rules stipulated by UK legislation, the EU has additional regulations regarding hazardous chemicals. The labelling and sale of dangerous chemicals is, in the EU, governed by various laws and regulations. The Dangerous Substances Directive [40] came into force in 1967, and applies to products placed on the market in the European Union. Potentially dangerous substances were to be labelled and the symbols shown in Figure 1.2 were created for consistency. The appropriate symbols had to be shown on packaging by law.

The Dangerous Preparation Directive was written in 1988, and subsequently rewritten in 1999 as a complement to the Dangerous Substances Directive [41]. This was brought in to effect in order to extend the regulations applied in the Dangerous Substances Directive over preparations, in addition to pure substances. In addition to defining the labelling conditions of the packaging, both of these directives regulate the packaging of these products such that the substance or preparation contained within should not be able to escape.

These two pieces of legislation were amended when classification, labelling and packaging (CLP) regulations came into effect in 2008 [42], as a supplement to the previously mentioned directives. The CLP made alterations to the warning symbols outlined in the Dangerous Substances Directive, with the intention of making these clearer (Figure 1.2). The CLP regulation also mandates that each member government appoint a body to be responsible for receiving information

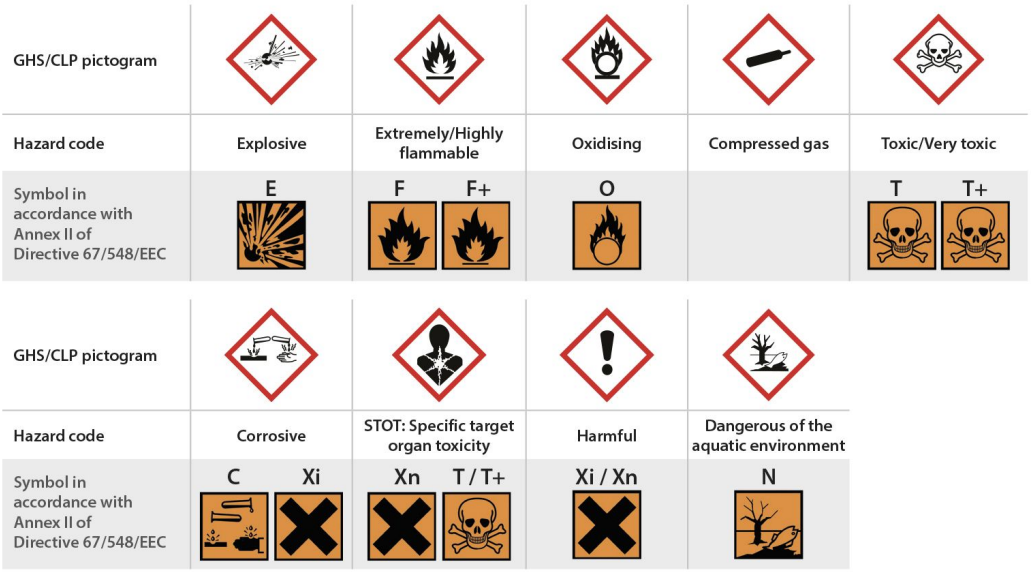


Figure 1.2: Symbols to be put on packaging under the Dangerous Substances Directive and the new symbols under the CLP regulations (Source: http://www.ghs-schuelke.com/ghs-en/Details.php)

about the ingredients of chemical products that are placed on the market. This is to be done to make this information accessible to health care professionals treating patients showing adverse symptoms following harmful exposure to these chemicals. In the UK this body is the National Poisons Information Service (NPIS) [43].

1.5 Poisons Information in the UK

The NPIS is comprised of 4 units: two in England (Newcastle and Birmingham), one in Wales (Cardiff) and one in Scotland (Edinburgh). The main aim of the NPIS is to provide year round accurate, up to date advice on the diagnosis, management and treatment of poisoning in the UK [1]. To that end, NPIS has produced and continues to maintain a database, TOXBASE, containing information on the treatment and management of cases of poisoning.

The provision of poisoning advice to healthcare professionals began in 1963, when the NPIS began a telephone enquiry service. This was, subsequently partly moved onto a Viewdata platform in 1983. As the Viewdata platform became obsolete, TOXBASE was moved onto an internet platform, to provide faster access to potentially vital information [44].

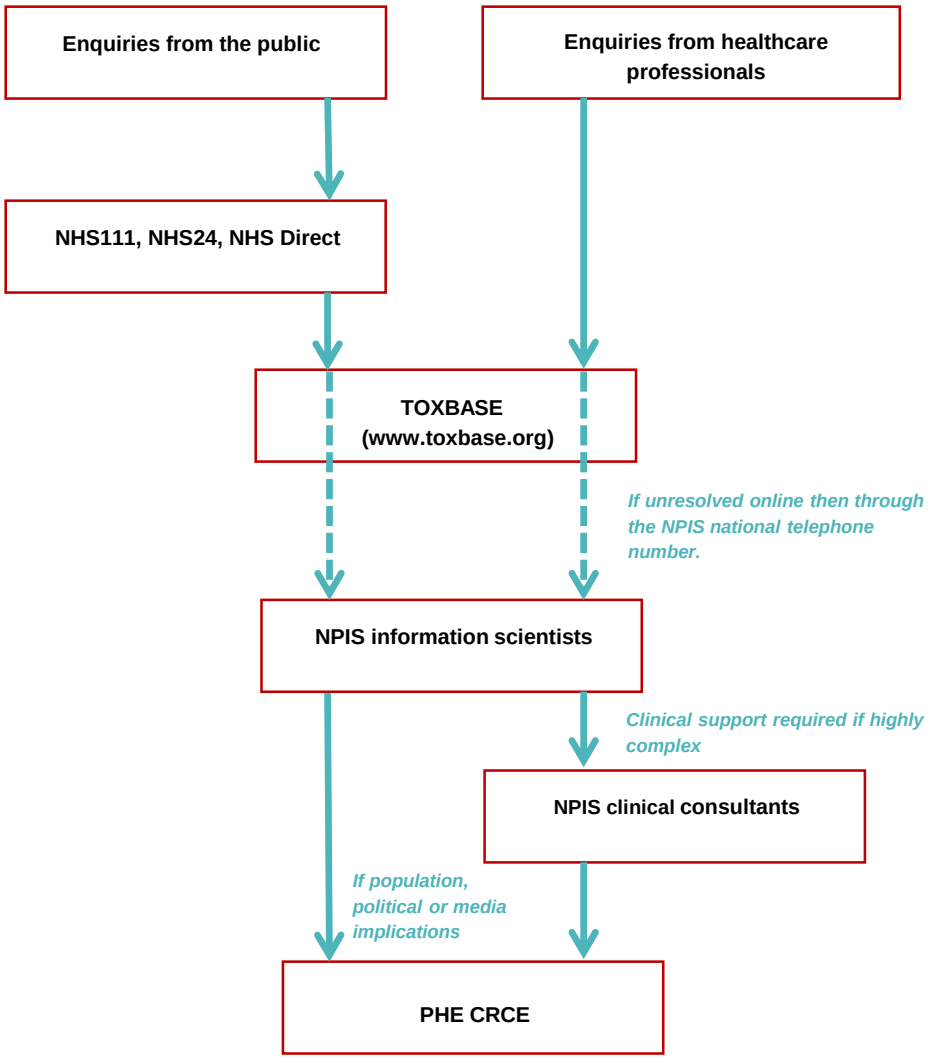


Figure 1.3: Diagram showing the path of enquiries through the NPIS. The first step is to look up TOXBASE, or contact the NPIS telephone service in complex cases. Any information that may have wide scale implications should be passed on to the Public Health England Centre for Radiation, Chemical and Environmental Hazards (PHE CRCE)

Three out of the four units (Birmingham, Cardiff and Newcastle) provide 24 hour telephone support, while the Edinburgh unit responds to telephone enquiries during the day in addition to maintaining the TOXBASE database. The flow of poisoning queries is shown in Figure 1.3. Healthcare professionals can access the TOXBASE database directly for information. However, members of the public do not have access to TOXBASE and would therefore likely contact one of the NHS telephone hotlines. If the query is not resolved by accessing TOXBASE, the query may be resolved via the NPIS telephone service, which is staffed by specialists in poisons information. Depending on how complex the case is, an NPIS clinical consultant may provide additional support. Information from enquiries may be passed on to the Public Health England Centre for Radiation, Chemical and Environmental Hazards, if this information is likely to have any population, political or media implications [1].

1.6 Use of the TOXBASE Database

The TOXBASE database is free to access for any medical professional within the NHS, with medical professionals outside of the NHS able to purchase access. Currently, the database has information on over 17,000 different toxins, with up to 5,000 product entries being written or revised each year, with 4,100 product entries edited in 2014/2015 [1]. TOXBASE provides information about diagnosis, treatment and management of patients suffering from exposure to various substances, spanning from pharmaceuticals to plants and animals [1].

The TOXBASE system makes a record of every access to the system, giving rise to a database containing information on the dates and times that accesses are made, which healthcare facilities are using the system and which pages are involved in user sessions. The National Institute for Health Care Excellence

(NICE), in its monograph on poisoning and overdose, indicates that TOXBASE should be accessed in cases of poisoning where there is any uncertainty in the treatment of a specific case of poisoning [45]. As such, examining the database of TOXBASE accesses should provide an insight into the prevalence of poisonings across the UK, although the true prevalence will be masked by variation in both individual and hospital level use. Since each access is logged, analysis can focus more specifically on individual product types or even type of healthcare facility. These analyses can involve specific chemical types, or may concern the general usage of the system.

1.6.1 Previous Studies on TOXBASE Usage

Each of the NPIS annual reports outlines TOXBASE usage for that financial year. There have been two studies in addition to these which have examined TOXBASE use as a standalone entity [46, 44]. The first of these [46], examined how usage of the TOXBASE database has impacted the usage of the telephone hotline over the first five years after TOXBASE went online. This indicates that as use of the online database increased, the volume of poisoning enquiries to the telephone service decreased.

The same study examined which chemicals, according to TOXBASE, were most commonly involved in poisonings. The findings from this indicates that the majority of accesses to the database concern pharmaceutical products, with 9 out of the top 10 products accessed by emergency departments at the time of study identified as medicines. The other product commonly accessed was the drug of abuse ecstasy [46].

This study also noted that TOXBASE use differs by user type. The number of accesses from each type of user was calculated as a rate per user and it was found that there were two particularly prolific user types: hospital emergency departments and the NHS public advice hotlines NHS 24 and NHS 111 [46].

The second study [44] was focussed on TOXBASE use in Scotland. This study noted the same negative relationship between TOXBASE use and usage of the NPIS telephone service. It was also of note that in the year 2000, emergency departments and minor injuries units were by far the most prolific users of TOXBASE, accounting for 23,061 agents and 14,713 sessions out of a total 27,712 agents accessed and 18,142 user sessions. This study also indicated that pharmaceuticals, was the most frequently accessed category in TOXBASE, accounting for around 73% of agents accessed.

This study also examined the rate of TOXBASE per 100,000 of the population for each of the Scottish health boards. This indicated differences in TOXBASE accesses by region, with NHS Grampian having the most sessions per 100,000 population (516) and NHS Western Isles having the lowest (179). This analysis excluded NHS Lothian, as the Royal Infirmary of Edinburgh is home to the Scottish NPIS unit, and NHS Shetland, which made no accesses over the year.

1.7 Poisoning in the UK Healthcare System

Cases of poisoning interact with various parts of the UK healthcare system, from General Practitioners (GPs) to hospital departments. All of these varied sections, however, encounter different aspects of poisonings; with GPs tending to deal with people who are contemplating self-harm [47, 48], and emergency departments tending to deal with more acute cases of poisoning [49].

GPs are well placed to discuss self-harm, of which poisoning represents a large proportion, with individuals as they may have a wider knowledge of other circumstances, for example an illness in the family, which may act as a trigger [48]. However, a study on GPs feelings on communicating with patients, in particular young patients, about self-harm found that they felt unsure on how to broach the subject and what language they should use when discussing self-harm [47].

Previously, GPs had shown a preference for using the NPIS telephone hotline when seeking advice about poisoning. The 2014/15 annual report indicated that GP usage of TOXBASE has been on the rise, indicating that they are becoming more familiar with the online system [1].

Although General Practitioners are vital in discussing and possibly preventing self-poisoning, they are less likely to be involved in incidents of unintentional poisoning, where there are no indications of the event prior to its occurrence. In this case, depending on the severity of the symptoms, these cases are likely to be handled by NHS telephone services or emergency departments.

These two services are the most prolific users of TOXBASE; in 2014/15 65% of all user sessions were from hospital departments, with 85% of those coming from emergency departments. Around 13% of TOXBASE accesses came from NHS 24 or NHS Direct. In comparison, GP surgeries accounted for approximately 6% of accesses and around 22.5% of telephone enquiries [1]. It is for this reason that this project specifically focuses on the NHS telephone services and emergency departments and their encounters with poisoning. These two services are discussed in more detail below.

1.7.1 NHS Public Telephone Service

The NHS public telephone service comprises two parts: NHS 111 (formerly NHS Direct) services England and Wales, while NHS 24 services Scotland, with both having the objective of reducing the demand on other sectors of the NHS, by providing a consistent and accessible out of hours service to the whole of the UK [50, 51]. NHS Direct was implemented in stages between March 1998 and November 2000. NHS Direct became NHS 111 in February 2014, in order to provide users with a more memorable contact number. NHS 24 launched in August 2002, providing an out of hours telephone service across Scotland.

Since 2014, NHS 24 has used the same 111 number as NHS 111. Calls are answered by trained call handlers, supported by experienced nurses and paramedics, who will ask questions relating to location and symptoms in order to best direct the call. Calls can be transferred directly to the relevant people, or the NHS call handlers can book an ambulance if needed. The NHS telephone services are for use when medical advice is needed quickly, but the symptoms do not require treatment at an emergency department. In cases of poisoning, it is likely that the NHS hotline would be the first step in getting treatment, particularly where there are no severe symptoms.

In this setting, TOXBASE can be used in order to help the operators to identify which symptoms in a case of poisoning by a specific substance would be considered mild or non-life threatening and which symptoms might be considered particularly dangerous, where medical advice should be sought immediately. The NHS telephone hotlines made 15,941 telephone enquiries to NPIS in 2014/15, an increase compared to previous years. It was found that hospital admission and GP referral was avoided in 55.7% of these cases. Preventing these unnecessary admissions and referrals represents a significant saving to the NHS [1].

1.7.2 Emergency Departments

Emergency departments (EDs) are at the front line of the provision of emergent health care. In 2014/2015, emergency departments in England saw 22.4 million attendances (as reported by A&E situation reports) [52], with a further 1.34 million attendances in Scotland [53] and 981,000 in Wales [54]. Approximately 1% of attendances in England had a primary diagnosis of poisoning [52].

Within an ED, the patient will be seen by a doctor or specialist nurse, who will assess the symptoms and attempt to determine which substances the patient has been exposed to. Where possible the patient will be treated and discharged immediately, however in cases where the poisoning is severe, the patient may be admitted for further treatment or observation. TOXBASE helps in patient assessment by providing a structured management plan, including best practices in the treatment of cases of poisoning by specific substances.

1.8 Thesis Outline

This thesis contains a variety of different analyses, all of which combine to form a multi-faceted overview of trends in TOXBASE usage and poisoning. Prior to this piece of work, analysis of TOXBASE use has comprised of simple statistical analyses, over a fairly limited period of time. This thesis will outline common spatial and temporal trends across a variety of NHS services. The thesis will also outline the linkage of NHS attendance and admissions data with the TOXBASE database, before moving on to an analysis of the relationship between TOXBASE use and admissions due to poisoning, which is something that has not been examined before.

The first step of this project was to determine whether there were any consistent trends in poisoning, and the exploration of these trends is reflected in Chapters 2, 3 and 5.

Chapter 2 utilises generalised additive modelling alongside spatio-temporal methods to assess both spatial and temporal patterns in NHS 24 calls with respect to poisoning. Chapter 3 then documents a similar assessment of the trends in accesses to TOXBASE made by emergency departments across the UK. The work in Chapter 3 was published in Human and Experimental Toxicology in 2018 [55].

Chapter 4 will focus on the development of a new method for parameter estimation in the presence of data that are missing not at random, with specific focus on data sets with suppressed values, a feature commonly found in data obtained from healthcare organisations. Chapter 5 will then apply these methods to data on UK wide emergency admissions due to poisoning, in order to assess the trends in poisoning admission as a proportion of poisoning attendances.

Data on TOXBASE usage at a hospital level is then linked with the data on emergency admissions. Chapter 6 outlines commonalities in hospitals which are unusual in terms of admissions or TOXBASE use. Chapter 7 then discusses an analysis of the impact of TOXBASE use on emergency admissions. This chapter is currently being developed into a manuscript for publication.

A final discussion of the analysis is presented in Chapter 8.

Chapter 2

Trends in Poisoning Calls to NHS24

This chapter documents the first set of analyses that was carried out as part of this project. Its focus is on modelling changes in NHS 24 calls which may concern poisoning or overdose. The goal of this analysis is to examine these data for consistent temporal trends, and also to assess relationships between poisoning and specific demographic variables selected based on the literature review in Chapter 1. This data set also provided an opportunity to assess a variety of spatial and temporal modelling techniques, including introductory spatio-temporal modelling described in Section 2.2.

The NHS 24 call data will be outlined in Section 2.1, along with a description of the demographic variables which were used as part of the analysis. Initially, spatial and temporal analyses were conducted independently by aggregating the data. This provided a basis for the results of the combined spatio-temporal analysis.

2.1 Description of data

NHS 24 records calls and call reasons as standard practice. The goal of this analysis was to examine temporal trends and covariate effects in order to gain an understanding of what factors may influence poisoning. Data were obtained from Health Protection Scotland, and contained the number of calls made to NHS 24 on a certain date in a certain postcode district. A postcode district is an area that is formed by combining all streets which share the same first half of the postcode. Since poisoning is a relatively rare occurrence, these postcode districts were combined to form 49 regions, as described by Barry Lynch, who supplied these data, in his thesis [56]. Further details of each region can be seen in Appendix A.

The regions obtained were a result of the data collection procedure, and other region structures may have been preferable were this possible. For example, having the data at data zone or intermediate zone level would have made finding possible covariates more straightforward.

A count of the number of poisoning related calls for each day was extracted by identifying specific key words or phrases, with three overlapping categories: "All", "Possible" and "Minimum". These categories refer to how likely each is to contain calls which are not poisoning related. As can be seen from Table 2.1, the "All" category contains the most key words and phrases and is therefore the most likely to have counts which contain calls unrelated to poisoning, the "Possible" category sits somewhere in the middle and the "Minimum" category is the least likely to count calls which are unrelated to poisoning.

These codes were chosen by consultants in public health in collaboration with colleagues from the National Poisons Information Service as part of an initial examination of temporal trends in poisoning related calls to NHS 24. These

data were therefore obtained before the beginning of this project, and while these codes are likely to cover the majority of poisoning calls, there are some calls which may not have been captured by these categories, such as toxic inhalation or toxic substances in the eye.

Table 2.1: Key words or phrases included in each of the "All", "Possible" and "Minimum" categories

	All	**Possible**	**Minimum**
Foreign Bodies	✓		
Foreign Body in Ear Adult	✓		
Foreign Body in Ear Child (5-16 years)	✓		
Foreign Body in Ear Toddler (1-4 years)	✓		
Foreign Body in Nose Adult	✓		
Foreign Body in Nose Child (5-16 years)	✓		
Foreign Body in Nose Infant (0-1 year)	✓		
Foreign Body in Nose Toddler (1-4 years)	✓		
Ingestion	✓	✓	✓
Ingestion - Baby	✓	✓	
Ingestion Foreign Body	✓	✓	
Ingestion Toxic Substance	✓	✓	✓
Poisoning	✓	✓	✓
Swallowing	✓	✓	

The data were initially aggregated up to the 49 regions, the first and last 3 rows can be seen in Table 2.2. These data contained daily counts of poisoning calls in each of the categories outlined in Tables 2.1 from 1st January 2006 until the 17th September 2013, for each of the 49 regions. However, poisoning is not a particularly common reason for calling NHS 24 and these data were sparse. In

fact, out of 134,399 observations, 87,618 (65%) were zero in the "All" category, 89,978 (67%) were zero in the "Possible" category and 99,146 (74%) were zero in the "Minimum" category. This led to the decision to aggregate the data up to a monthly level.

Table 2.2: Data after aggregation to the 49 regions used in this analysis

Date	Area	All	Possible	Minimum	Total
2006-01-01	Aberdeen Inner	0	0	0	38
2006-01-02	Aberdeen Inner	0	0	0	51
2006-01-03	Aberdeen Inner	0	0	0	46
...	...	...	...	...	...
2013-09-15	West Moray	0	0	0	15
2013-09-16	West Moray	0	0	0	8
2013-09-17	West Moray	0	0	0	7

After aggregating the data there were 4,557 observations. Of these there were 269 (5.9%) which were zero in the "All" category, 284 (6.2%) in the "Possible" category and 361 (7.9%) in the "Minimum" category.

In addition to the NHS 24 call data, covariates were collected based on the discussion in Chapter 1. The Scottish Neighbourhood Statistics DataMart [57] was queried for appropriate variables. The choices were limited by the spatial granularity in the data set, caused by the fact that the data were aggregated up from postcode district level into 49 made up regions. The variables obtained related to rurality, deprivation and age distribution.

Rurality was assessed using the Scottish Government 6-fold Urban/Rural classification [58]. These data were available at postcode level, meaning that the postcode district could be extracted by removing the last three characters of the postcode. The urban/rural classifications were available alongside a postcode level population, as measured in 2011. These populations were then aggregated up to a regional level using the links shown in Appendix A in order to get a total population for each region. The populations for only those postcodes with

a rural score of 4, 5 or 6 were aggregated in a similar way, to obtain the number of people in each region who were living in a rural or remote area. This allowed for the calculation of the percentage of the population living in a rural or remote area to be calculated. This produced the data set shown in Table 2.3.

Table 2.3: Table showing an excerpt from the rural data

Region	Prop
Aberdeen Inner	0.0000
Aberdeen Outer	10.0826
Aberdeenshire North	62.8643
Aberdeenshire South	38.8907
...	...
Stirling	34.2419
West Dunbartonshire	4.1815
West Lothian	10.7822
West Moray	54.0831

"Prop" is the percentage of the population of the region who live in a rural or remote area. Aberdeen Inner is a region in the centre of the city of Aberdeen, meaning that none of its population live in a rural area, whereas around 54% of the population in West Moray live in a remote or rural area.

Figure 2.1 shows that these data have a large number of regions with a small percentage living in a rural or remote area, with less having a moderate percentage. In fact, a large number of regions have less than 10% of the population living in a rural or remote area. Six of those regions, in fact, had less than 1% living in a rural or remote area: Aberdeen Inner, Edinburgh East, Glasgow East, Glasgow South, Glasgow West and South Lanarkshire West.

There appears to be a small spike in the number of regions with a high percentage living in a rural or remote area. There were ten regions which had more than 80% of their population living in a rural or remote area. Nine of these were regions where 100% of the population were described as living in a rural or remote area: Argyll & Bute Islands, Argyll & Bute Mainland, Arran & Cumbrae,

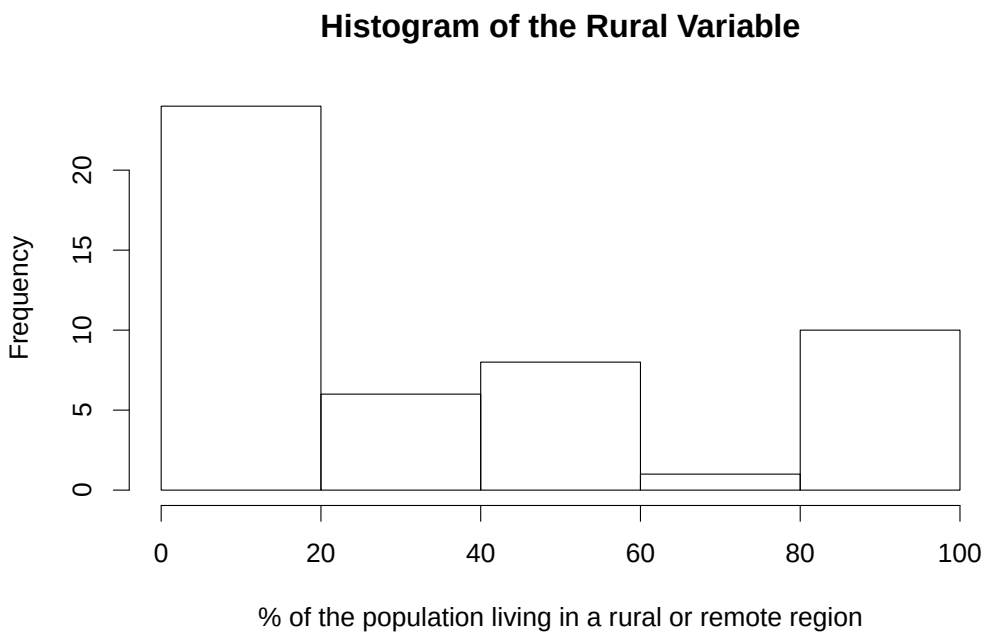


Figure 2.1: Histogram showing the distribution of the "Rural" variable

Badenoch & Strathspey, Caithness & Sutherland, Eilean Siar, Orkney Islands, Shetland Islands and Skye & Lochalsh. The other region, Ross & Cromarty, had 98.3% of its population living in a rural or remote area.

The measure used for deprivation was the Scottish Index of Multiple Deprivation (SIMD), which is a measure constructed at data zone level. The postcodes were mapped to data zones, in order to carry out a similar aggregation to that used in rurality. As before, the postcode level populations were aggregated up to the regional level, as were the populations for those living at a postcode located in one of the most deprived 20% of areas in Scotland, i.e. in SIMD quintile 1. This led to a calculation of the percentage of the population in each region who live in a deprived area. The resulting data set is shown in Table 2.4:

Table 2.4: Table showing an excerpt from the deprivation data

Region	Prop
Aberdeen Inner	18.0898
Aberdeen Outer	8.7026
Aberdeenshire North	3.5936
Aberdeenshire South	0.6853
...	...
Stirling	9.1293
West Dunbartonshire	35.3015
West Lothian	14.6570
West Moray	2.5752

In these data, "Prop" is the percentage of the population living in an area which is in the 20% most deprived areas in Scotland. It can be seen that West Dunbartonshire had a relatively high proportion of the population living in a deprived area at 35%, and Aberdeenshire South had a relatively low proportion at 0.69%.

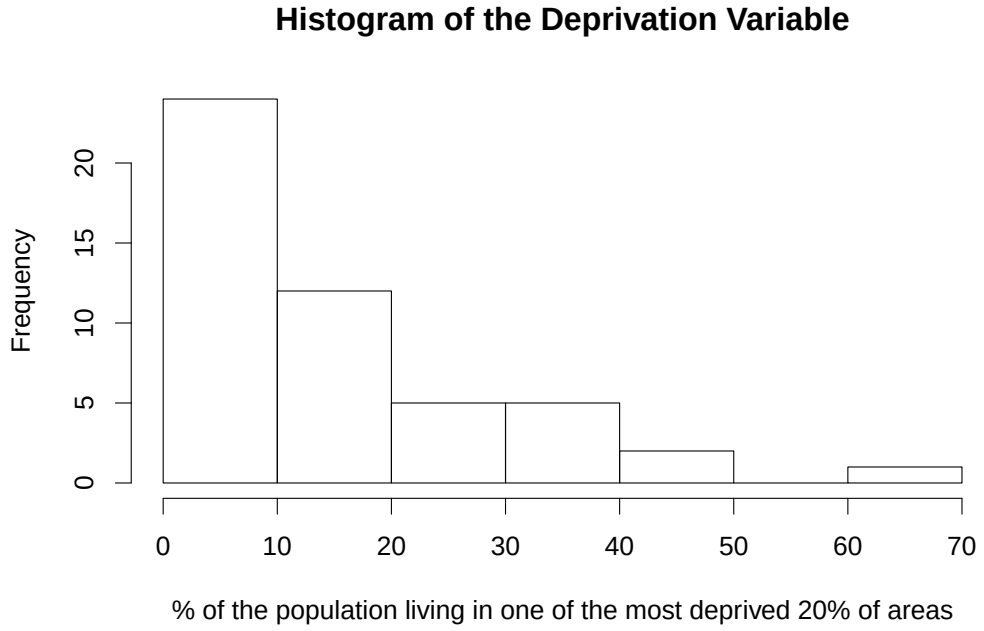


Figure 2.2: Histogram showing the distribution of the "Deprivation" variable

Figure 2.2 shows the distribution of these data. There were a large number of regions which had less than 10% of the population living in a deprived area. The region with the highest proportion of the population living in a deprived area was Glasgow East, with 69.3%. There were several regions which had no members of the population living in one of the 20% most deprived areas in Scotland: Argyll & Bute Islands, Arran & Cumbrae, Badenoch & Strathspey, Eilean Siar, Orkney Islands, Shetland Islands and Skye & Lochalsh.

The age data set had the population size separated by gender and by age category for each postcode sector in Scotland. The age categories were in groups of 5 years, starting for 0-4, with the final category being 90+. These age groups were separated into three variables. These were "Toddler", the percentage of children aged 0-4, "Child", the percentage of children aged 5-15 and "Pensionable", the percentage of the population above 65, which was the pensionable age in Scotland at the time of writing. The other age groups were excluded in order to reduce the risk of severe multicollinearity.

As with the deprivation and urban/rural variables, the population in each postcode sector was aggregated up to a regional level using the table in Appendix A, which allowed for the calculation of the percentage of the population in each of the three age categories described. This resulted in the data set shown in Table 2.5.

The top left plot in Figure 2.3 shows the distribution of the percentage of children aged 0 to 4 years. It would appear that the population in most regions comprise around 5.5-6% toddlers. The minimum percentage of toddlers in any region was 4.33% in Arran & Cumbrae, closely followed by 4.52% in the Aberdeen Inner region. West Lothian and North Lanarkshire West had the two largest percentages, with 6.46% and 6.32% of the population being aged between 0 and 4 respectively.

Table 2.5: Table showing an excerpt from the age distribution data

Region	Toddler	Child	Pensionable
Aberdeen Inner	4.5622	8.6405	14.5032
Aberdeen Outer	5.7712	13.9832	15.5445
Aberdeenshire North	5.7065	13.1849	14.4436
Aberdeenshire South	5.9623	14.0627	13.0292
...	...	...	...
Stirling	5.6261	12.2756	15.1560
West Dunbartonshire	5.8597	13.4824	15.8893
West Lothian	6.4556	13.5721	10.9475
West Moray	6.0753	13.6355	14.6222

The top right plot in Figure 2.3 displays a histogram of the percentage of children aged 5 to 15 in each region. The majority of regions have around 13-14% of their population aged 5 to 15. The Shetland Islands had the largest percentage of children at 14.97%, with the smallest proportion of children being seen in Aberdeen Inner at 8.64%

The bottom plot shows the distribution of the percentage of the population in each region who are of pensionable age. This indicates that the majority of regions have around 15% of their population over 65. This is particularly high, at 25.6%, in Arran & Cumbrae. The region with the smallest percentage was West Lothian, where around 10.95% of the population were of pensionable age.

These covariates were obtained for the year 2011 and were treated as fixed through time. This is reasonable in the case of rurality, which does not change over time. Deprivation does vary over time, although does not tend to change by a large amount. There are generally small shifts in the rankings, however these data were treated as the proportion of the population living in one of the 20% most deprived area, which makes large changes in these data very unlikely. The age distribution may change more through time, however the treatment of these data as proportions will mitigate these changes, and the regional variation in age distribution will be much greater than the changes through time.

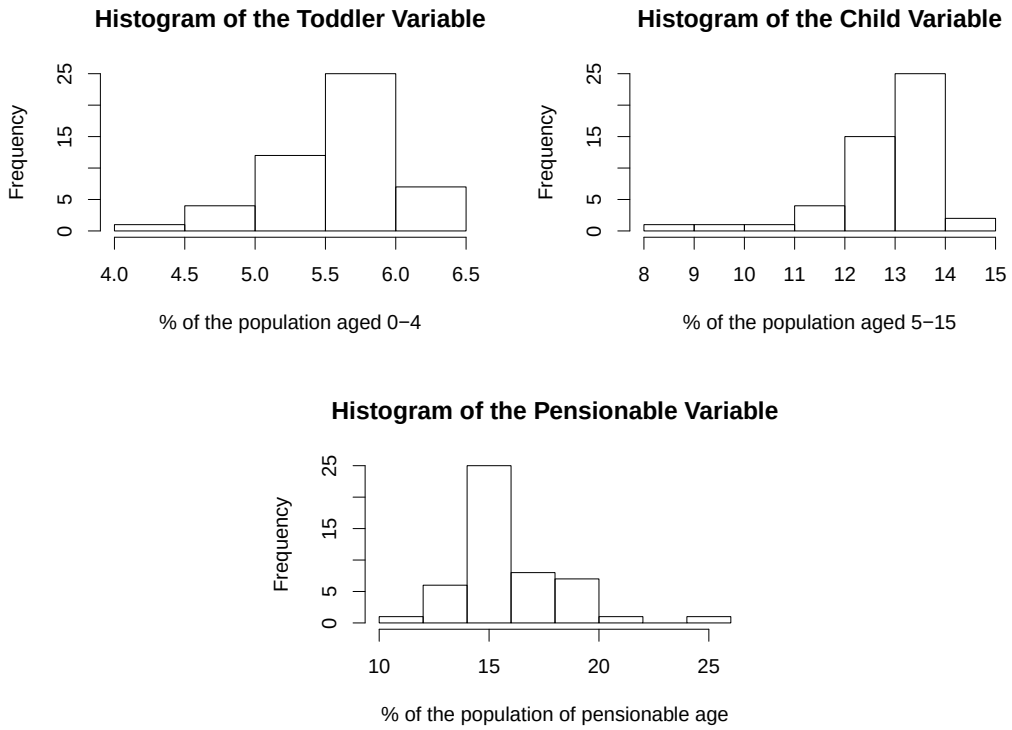


Figure 2.3: Histogram showing the distribution of the age related variables

2.2 Methods

The data were assessed from three different perspectives. First the data were aggregated across the regions, to get the total number of poisoning calls across Scotland for a given month. A temporal analysis was conducted on these as outlined in Section 2.2.1. The next step was to aggregate the data across the entire time period, in order to get the total number of poisoning calls in each region between January 2006 and September 2013. A spatial analysis was carried out on these using the demographic variables described in Section 2.1, using the methods outlined in Section 2.2.2. Finally, these two analyses were combined and a spatio-temporal analysis was conducted on the monthly level calls data, using the fixed time demographic variables, as outlined in Section 2.2.3.

2.2.1 Temporal Modelling

As part of this analysis, it was decided that the temporal patterns should be treated as non-linear. This is largely to reflect the inherent non-linearity in within year trends, which should start at the same level at which it finishes in order to provide continuity to the fitted function. This means that simple regression techniques are not appropriate. As such, in order to assess the temporal trends in the data, Generalised Additive Models (GAMs) [59] were used as implemented in the `mgcv` package [60, 61, 62] in `R` [63].

Generalised additive models are a non-linear regression technique, which allow complex functions to be modelled using splines. The interest of this part of the analysis is in seasonal and long term patterns in poisoning calls to NHS 24, which have been extracted using a model with fixed effect component of the form described in Equation 2.1.

$$E(\log(y_i)) = \log(c_i) + \beta_0 + f(m_i) + g(t_i) \tag{2.1}$$

The response, denoted y_i is the number of poisoning calls made at time point i. In order to account for differences in overall call volume the offset, $\log(c_i)$, was included. This represents the logged number of total calls made to NHS 24 at time i. The inclusion of this offset terms means that any inference drawn from this model is about the proportion of poisoning calls, rather than the number of poisoning calls. The within-year seasonality is represented by the term $f(m_i)$, where m_i is a month indicator. The value t_i represents a time indicator so that $g(t_i)$ will show the overall trend in calls. These functions were estimated using cubic regression splines.

The model above makes use of splines in order to fit the seasonality term $f(m_i)$ and the long-term trend $g(t_i)$. Splines use piecewise polynomial functions, which are smoothly joined together at predefined locations called knots. The type of polynomial function which is fit in the intervals is determined by the type of spline used. In this analysis, the long-term trend was fit using a cubic regression spline. This type of spline fits a polynomial function of degree 3, a cubic function, within each subdivision of that data.

If we have k knots, and we fit a function between each pair of knots, then there are $k-1$ polynomial functions. In order to fit a smooth curve, these functions have to conform to specific conditions. To demonstrate these conditions suppose that we are fitting a curve, where we have three knots: one at the beginning of the data, one in the middle of the data and one at the end of the data, as shown in grey in Figure 2.4.

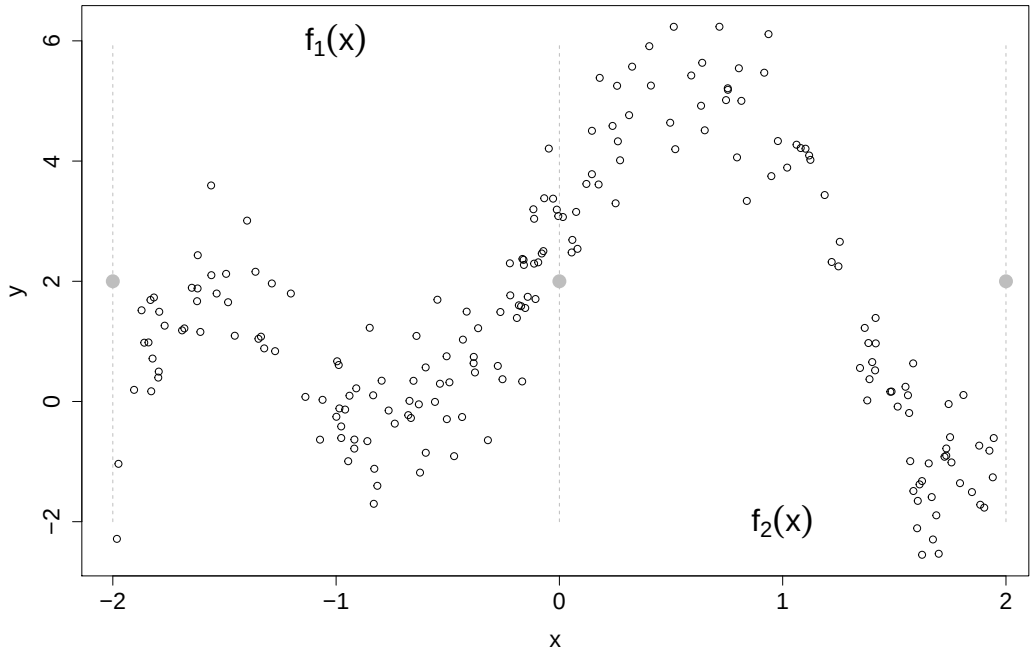


Figure 2.4: Diagram describing the setup for the example. The knots are shown in grey. The function $f_1(x)$ is fit on the first half of the data (x values between -2 and 0) and $f_2(x)$ is fit on the second half of the data (x values between 0 and 2)

Then we fit two curves, one between the first and second knots and one between the second and third knots. These will be referred to as $f_1(x)$ and $f_2(x)$ respectively. Then these curves have to satisfy:

1. That $f_1(x) = f_2(x)$ at their joining knot, that is the middle knot in this example. This results in C^0 continuity.

2. The slope of $f_1(x)$ must be equal to the slope of $f_2(x)$ at their joining knot. This is equivalent to saying that the first derivative of the functions evaluated at the knot must be equal: $f_1'(x) = f_2'(x)$. This results in C^1 continuity.

3. The curvature of $f_1(x)$ must be equal to the curvature of $f_2(x)$ at their joining knot. This is equivalent to saying that the second derivative of the functions evaluated at the knot must be equal: $f_1''(x) = f_2''(x)$. This results in C^2 continuity.

These three conditions form a system of equations, which can be solved to obtain estimates of the cubic polynomial coefficients. When being fit to data, these functions have to be fit using a numerical optimization, where each function, $f(x)$ is fit such that $\sum_{i=1}^{n}(y_i - f(x_i))^2 + \lambda \int_a^b [f''(x)]^2 dx$ is minimised, this is penalised least squares. In the term $\lambda \int_a^b [f''(x)]^2 dx$, a is some constant which is smaller than the minimum covariate value and b is some constant which is larger than the maximum covariate value. This term is used to penalise terms which are too rough.

In situations where multiple functions are being fit, as in this analysis, y_i would denote the partial residuals. The partial residuals for a given term represent the variation in the response which are not explained by the other terms in the model. For example when fitting the long term trend, the partial residuals would

be $\log(y_i) - \log(c_i) - \beta_0 - f(m_i)$. Each of the fitted function will be presented on the scale of the partial residuals, which will be referred to throughout as the "centred response".

The seasonal trend was fit using a cyclical cubic regression spline which additionally constrains the fitted curve to adhere to the conditions outlined above at the end knots as well as those within the data. The implication for this analysis, is that the seasonal trend takes the same value in January as it did in December.

There are two ways in which the smoothness of the fitted functions can be manipulated. The first is by setting the smoothing parameter, which is denoted by λ in the penalised least squares function above. When this parameter is large, there is a larger penalty on rough functions, resulting in a smoother estimated function. When this parameter is small the function is more likely to be rough.

The second way to manipulate the smoothness of the functions is to set the number of knots. The fewer knots used in the function fitting, the smoother the fitted function will be. Conversely, if there are a large number of knots, then the fitted function will be rough. For example, if the number of knots is equal to the number of observations, the estimated curve can be an interpolation of the data.

Within the `mgcv` package, the smoothing parameter is estimated by cross validation, unless it is specified by the user. This procedure can result in functions which overfit the data, however the process does allow for some flexibility. Therefore, the decision was made to manipulate the smoothness of the fit by changing the number of knots, and allowing the smoothing parameter to be estimated. The removal of knots was done by observation, bearing in mind that smoother functions are less likely to overfit the data, and therefore more likely to capture the underlying real world process.

Since the response variable is made up of counts, the natural assumption is that these data follow a Poisson distribution. However, use of the Poisson distribution makes the strong assumption that $E(y_i) = \text{var}(y_i)$, that is that the mean is equal to the variance. This assumption is commonly not met, however, and was assessed in this case by examining the residual deviance divided by the residual degrees of freedom, also known as the dispersion factor. This uses the idea that, when data are overdispersed, the variance σ^2 can be written as:

$$\sigma^2 = \mu\phi$$

where μ is the expected value and ϕ is a constant multiplier. This can be rearranged so that $\phi = \sigma^2/\mu$. Then where the data are correctly dispersed, $\phi = 1$. To assess this, the model deviance can be used as an approximation for σ^2 and the residual degrees of freedom can be used as an approximation for the mean. This means that if the ratio of these two is close to 1, correct dispersion can be assumed. If this value is much larger than 1, then there is evidence for overdispersion, and if the value is much less than 1 then there is evidence of underdispersion. Where the data exhibited overdispersion, a Negative Binomial distribution was used instead of the Poisson distribution.

In any regression, one of the underlying assumptions is that the errors from the fitted model are independent, which is often not the case when modelling time series. A breach of this assumption means that variances are likely to be under-estimated and therefore confidence intervals too narrow. In order to assess this assumption, the autocorrelation function (ACF) and partial autocorrelation function (PACF) were examined.

The autocorrelation function measures the correlation between values at different lags. These values are presented on a plot as shown in Figure 2.5, where the dashed lines represent approximate 95% confidence limits.

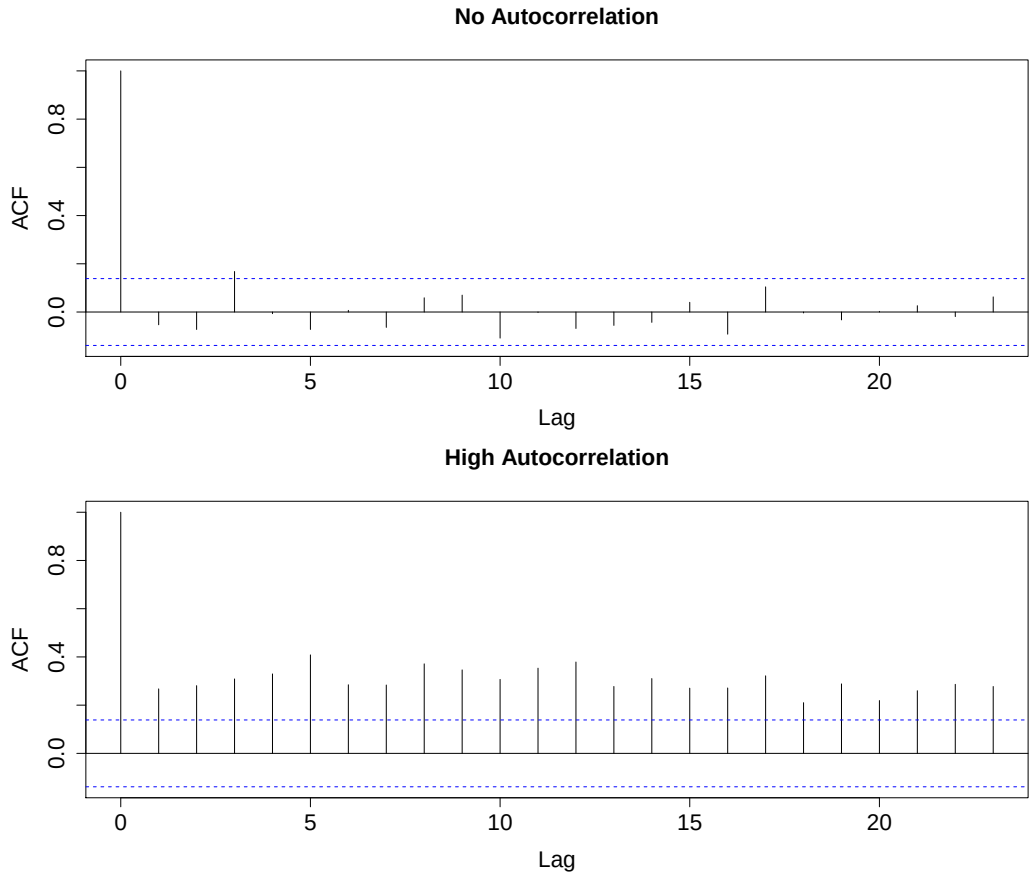


Figure 2.5: Example of the autocorrelation function when there is no correlation (top) and high correlation (bottom)

In these plots the first value in the plot corresponds to lag 0. That is the measure of each value with themselves, and therefore always has the value 1. Any bars which represent small lags which lie outside of the dashed blue lines are an indication of temporal autocorrelation. Note that there are still some lags outside of the confidence limits in the ACF of the data where there is no autocorrelation; this is fine, since the plot uses a 95% interval one would expect

that some values would lie outside of the interval by chance. In combination with the partial autocorrelation function, this type of plot allows for the assessment of the number of lags to be included in a moving average (MA) process.

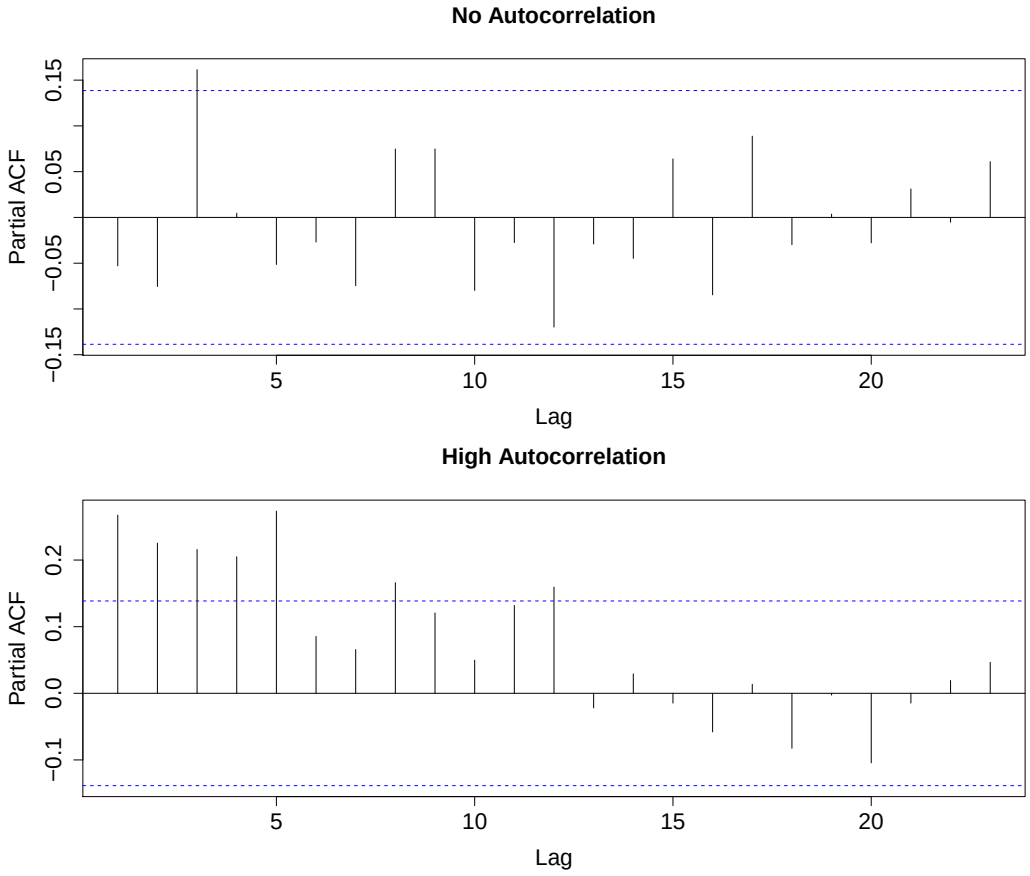


Figure 2.6: Example of the partial autocorrelation function when there is no correlation (top) and high correlation (bottom)

The partial autocorrelation function is similar to the autocorrelation function, except that it is a measure of the autocorrelation between values at different lags, after accounting for the intermediate lags. That is, it is the autocorrelation between observations at time i and $i + t$ after accounting for the relationship between observations at time i and times $i + 1, i + 2, ..., i + t - 1$. These are plotted and interpreted in a similar way to the autocorrelation function (Figure 2.6), and allow for the assessment of the number of lags to be included in an autoregressive (AR) process.

The two examples in Figure 2.6 use the same data as those in Figure 2.5. Focussing on the bottom plot in each of these, it can be seen that the ACF consistently lies outside of the approximate 95% interval, while the PACF is outside of the interval for the first three lags. This is an indication that an AR(3) process would be appropriate for these data. A similar assessment was conducted on the residuals from the temporal model, and ARMA processes were incorporated into the model fit using a generalised additive mixed model (`gamm`) where necessary to account for autocorrelation.

2.2.2 Spatial Modelling

The spatial modelling was done on the total number of calls made to NHS 24 between January 2006 and September 2013, where the model in Equation 2.2 was used as a starting point in the modelling procedure. The response, denoted y_j is the number of poisoning related calls made to NHS 24 between January 2006 and September 2013 in region j, while c_j is the total number of calls made to NHS 24 in region j. The inclusion of c_j as on offset term means that interpretation is made with respect to the proportion of poisoning calls, as in the temporal model.

The covariates were fit as linear terms. The urban/rural covariate is denoted in Equation 2.2 by U_j, which represents the percentage of the population in region j who live in a rural or remote area. The percentage of the population in region j living in a deprived area is denoted D_j. Then the percentage of the population aged between 0 and 4 was denoted T_j, K_j represents the percentage of the population aged between 5 and 15 and P_j denotes the percentage of the population aged 65 or over.

$$E(\log(y_j)) = \log(c_j) + \beta_0 + \beta_1 U_j + \beta_2 D_j + +\beta_3 T_j + \beta_5 K_j + \beta_6 P_j \qquad (2.2)$$

The final fixed effects component of the model was determined using backwards selection using significance testing.

When modelling observations across spatial locations, it is common to find that the model errors are spatially autocorrelated, due to near regions being more similar than distant regions. This is particularly true where some spatially correlated covariate has not been included in the model. The residuals of the model were therefore assessed for spatial association using Moran's I.

In order to calculate Moran's I, it is necessary to have some description of the spatial structure. In this case the data are observed in regions, or areal units, meaning that it is appropriate to construct a spatial adjacency matrix. This was constructed on the basis of rook adjacency, such that regions were defined as neighbours if they share a border, neighbours are denoted by a 1 in the adjacency matrix, while non-neighbouring regions are denoted by 0.

In order to see this more clearly, consider the straightforward example in Figure 2.7.

There are four regions numbered 1 to 4, focussing on region 1, we can see that it shares a border with regions 2 and 4. This means that the first row, and first column since the adjacency matrix must be symmetric, will be 0, 1, 0, 1. This can be repeated for all regions to obtain the following adjacency matrix:

$$\begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}$$

The data set under consideration in this chapter, is slightly more complicated, in that it contains information on the island regions of Scotland. These regions would, by the previous definition, have no neighbours. When regions have no

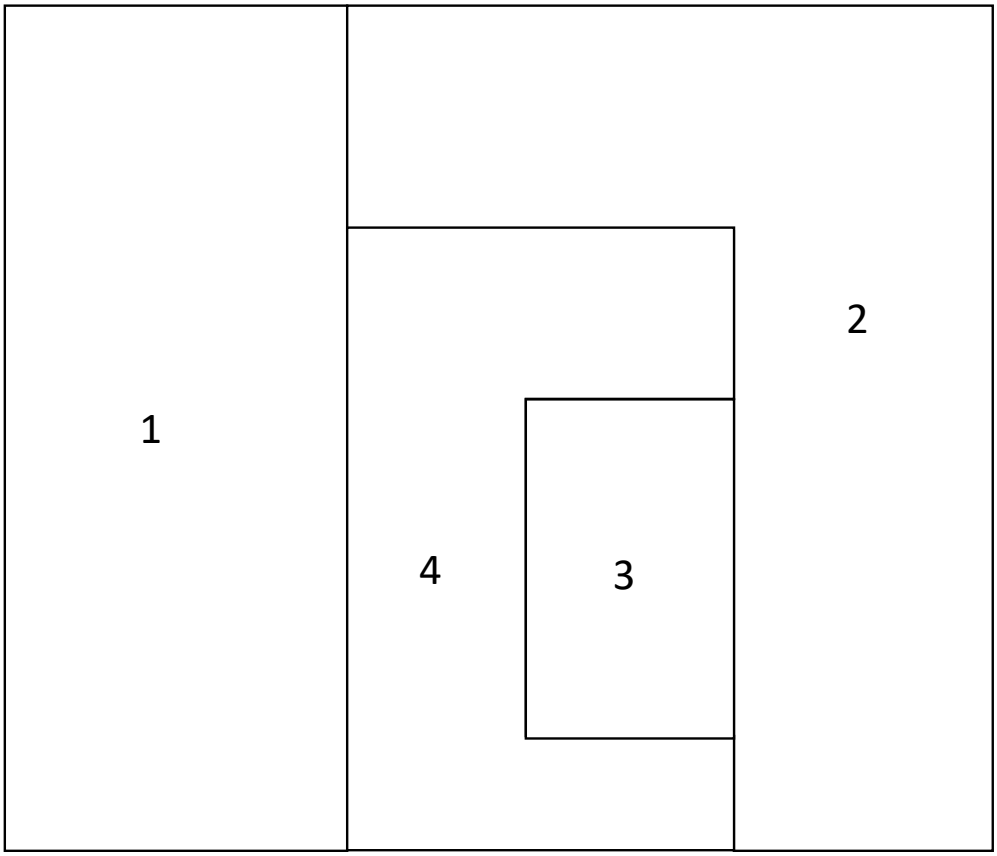


Figure 2.7: Image showing the regions used in the example of constructing a spatial adjacency matrix

neighbours, all elements in that row of the adjacency matrix would be zero, this leads to problems in the estimation of spatial random effects and for correlation estimates. To account for this, each island region was assigned a neighbour based on its closest neighbour, as has been done in a previous study involving Scottish islands [64].

Using the method outlined above, Argyll & Bute Islands was assigned to Argyll & Bute Mainland, Arran and Cumbrae was assigned to North Ayrshire, Eilean Siar was assigned to Skye & Lochalsh, the Orkney Islands were assigned to Caithness & Sutherland and the Shetland Islands were assigned to the Orkney

Islands. In order to preserve symmetry, these links were also implemented in reverse, so that Argyll & Bute Mainland was assigned as a neighbour to Argyll & Bute Islands, for example.

Residual spatial dependence was assessed using Moran's I, the general form of which can be seen in Equation 2.3 [65]. The value w_{ij} is the neighbour indicator for regions i and j and is the i, jth entry in the spatial adjacency matrix. This means that only those pairs of regions which are neighbours have an effect on the calculation of Moran's I. The term $w_{..}$ is the sum over all of the weights in the spatial adjacency matrix. In general z_i is the observed value for region i and $\bar{z}$ represents the mean of the observations. In this analysis, Moran's I will be used to calculate spatial autocorrelation within the model residuals.

$$I = \frac{n \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}(z_i - \bar{z})(z_j - \bar{z})}{w_{..} \sum_{i=1}^{n}(z_i - \bar{z})^2} \tag{2.3}$$

Moran's I was assessed in two ways. First the statistic was calculated for all data, Moran's I was then recalculated for the data excluding the island regions in order to ensure that any residual spatial autocorrelation was not masked by including the island regions. A permutation test was used in order to test for significant spatial association within the residuals. The null hypothesis of this test is that there is no spatial association, and the alternative hypothesis is that there is spatial association.

The test is performed by using the fact that, once observations are permuted to different locations, they lose their spatial trend. The data are therefore randomly distributed among the regions some number of times M, and I is calculated for each permutation. The number of permutations must be chosen, but should be large enough to obtain a distribution of Moran's I under spatial randomness. For

this analysis, 1,000 permutations was chosen, as it is sufficient to get an idea of the distribution, but small enough that it would not take too long to run. The test statistic was calculated as:

$$\frac{2}{M+1}\sum_{k=1}^{M} I(I_k > |I_{obs}|)$$

where I_k is Moran's I for permutation k and I_{obs} is Moran's I for the original data. The numerator 2 appears in order to make this a two-sided test.

Any residual spatial autocorrelation was accounted for using Markov random fields (MRFs). A Markov Random Field is a process such that the joint distribution can be written as a product of conditional distributions, which in a spatial setting are based on the neighbouring regions. For n regions this means that:

$$f(Z_1, ..., Z_n) = \prod_{i=1}^{n} f(Z_i | Z_j : j \in N_{Z_i})$$

This is the foundation of conditional autoregressive models, which are commonly used in spatial modelling. These models take the form

$$Z_j \sim N\left(\frac{\sum_{i=1}^{n} w_{ij} Z_i}{\sum_{i=1}^{n} w_{ij}}, \frac{\sigma^2}{\sum_{i=1}^{n} w_{ij}}\right)$$

so that the distribution of each observation has a mean which is equal to the mean of all of its neighbouring observations, with variance equal to some constant which is scaled by the number of neighbouring regions.

This structure was incorporated into the models where required using functionality in the `mgcv` package. This is done within a smooth term as follows:

```
s(region, bs='mrf', xt=list(nb=neighbours))
```

The first argument specifies the variable which contains the region indicators. The `bs` argument in general specifies which type of smooth to fit, in this case `'mrf'` indicates that a Markov Random Field should be estimated. Then the neighbourhood information is specified through the `xt` argument, and requires that the adjacency matrix be passed as a list, where each list entry corresponds to a specific region and contains the index or name of all neighbours to that region.

2.2.3 Spatio-Temporal Modelling

The spatio-temporal modelling involved combining the fixed effects from the two previous models as shown in Equation 2.4.

$$E(\log(y_{ij})) = \log(c_{ij}) + \beta_0 + \beta_1 U_j + \beta_2 D_j + \beta_3 T_j + \beta_4 K_j + \beta_5 P_j + f(m_i) + g(t_i) \quad (2.4)$$

As before, U_j represents the percentage of the population in region j living in a rural or remote area, D_j represents the percentage of the populations in region j living in an area which is in the top 20% most deprived in Scotland. The terms T_j, K_j and P_j are the percentage of the population in region j aged 0 to 4, 5 to 15 and 65+ respectively. $f(m_i)$ is the smooth term for within year seasonality, where m_i is the month indicator for time i and $g(t_i)$ is the smooth term for the long term trend, where t_i is a time indicator. The response y_{ij} is the number of poisoning calls made to NHS 24 at time i in region j, and c_{ij} is the total number of calls made to NHS 24 at time i in region j. The inclusion of the term $\log(c_{ij})$ means that inference from this model is done in relation to the proportion of poisoning calls.

A similar method to that in the spatial analysis was used in the spatio-temporal model. This was extended, so that the Markov Random Field covered the spatio-temporal process, rather than just a spatial process as previously [66, Chapter 11]. The idea behind this is that the spatio-temporal process can be modelled as

$$Y_{st} = \mu_{st} + \omega_{st} + \varepsilon_{st}$$

where μ_{st} is the mean of the process, as described in Equation 2.4 and ε_{st} represents a set of independent errors. The term ω_{st} represents the residual spatio-temporal structure, and can be decomposed as follows

$$\omega_{st} = \nu_s + \gamma_t + \kappa_{st}$$

where ν_s represents the underlying spatial structure, or spatial random effects as modelled in Section 2.2.2. The term γ_t represents the underlying temporal structure, or temporal random effects as discussed in Section 2.2.1 and κ_{st} represents the interaction of spatial and temporal structures, which makes allowance for different temporal structures in different locations and different spatial structure at different time points.

This structure was accounted for in the model using the MRF capabilities in `mgcv` as before. The ν_s component was incorporated using the code described in Section 2.2.2:

```
s(region, bs='mrf', xt=list(nb=neighbours))
```

The γ_t component was incorporated using the `correlation` argument in the `gamm` function as in Section 2.2.1. Since it is difficult to visually assess the overall temporal autocorrelation, the parameters in the ARMA process were selected using AIC. Finally, the κ_{st} component was incorporated via the following code

```
ti(region, year, bs=c("mrf", "cr"),xt=list(nb=neighbours))
```

where `ti` denotes a smooth interaction between two continuous variables. The `'mrf'` basis corresponds to region, while `'cr'` corresponds to the long term trend variable. The `xt` argument takes the spatial information to feed into the computation of the spatial markov random field.

Moran's I was used to assess spatio-temporal association as in the spatial analysis, meaning that a new adjacency structure is required. The construction of this spatio-temporal adjacency matrix required the spatial adjacency matrix as described previously, and a temporal adjacency matrix. The temporal adjacency matrix was constructed using the same rules as the spatial adjacency matrix, meaning that positions where the time points were next to each other contained a 1. To illustrate this, consider a time series with 4 time points. Then time 1 neighbours time 2, time 2 neighbours time 3 and time 3 neighbours time 4. This would be expressed as the following adjacency matrix

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

The spatio-temporal adjacency matrix was constructed by taking the kronecker product of the spatial adjacency matrix as described in the previous section and a temporal adjacency matrix similar to that above, but spanning 93 time points.

2.3 Results

The analysis was carried out for all three of the call categories. The number of calls in each category are highly correlated, as can be seen in Table 2.6. Due to this high correlation, the results of the analyses were almost identical across the three groups. For this reason, only the results for the "All" category are presented here, as these data consist of larger call numbers than the two other categories.

Table 2.6: Pearson correlations between the number of calls in the "All", "Possible" and "Minimum" categories

	All	Possible	Minimum
All	1.000	0.996	0.985
Possible	0.996	1.000	0.992
Minimum	0.985	0.992	1.000

2.3.1 Temporal Analysis

Initially, the data were aggregated across Scotland in order to get an overall picture of the temporal trends present in poisoning calls to NHS 24. The first and last four rows of the data are shown in Table 2.7 below.

Table 2.7: Table showing the structure of the data used in the temporal analysis

Date	All	Possible	Minimum	Total	Month	Year
2006-01	387	370	290	48733	1	0.0833
2006-02	314	304	234	38563	2	0.1667
2006-03	386	362	284	42556	3	0.2500
2006-04	362	343	280	48918	4	0.3333
...	...	...	...	...	...	...
2013-06	762	706	497	103771	6	7.5000
2013-07	770	706	500	99461	7	7.5833
2013-08	820	756	517	96616	8	7.6667
2013-09	440	419	272	54484	9	7.7500

The data used were aggregated up to month level, where the "Month" column contains a numeric value for month of the year. The "Year" column combines information on the year and the month, for use in fitting the generalised additive model. This takes the value year $-$ 2006 $+$ month/12.

The distribution of calls made to NHS 24 can be seen in Figure 2.8. This indicates that the number of poisoning calls in a given month ranges from 300 to 1,000, with a most months having above 500 calls. The distribution is altered when the proportion is taken, where, in the majority of months, around 0.8% of the call volume were related to the "All" category. In both counts and proportions, the maximum value is around three times the minimum value.

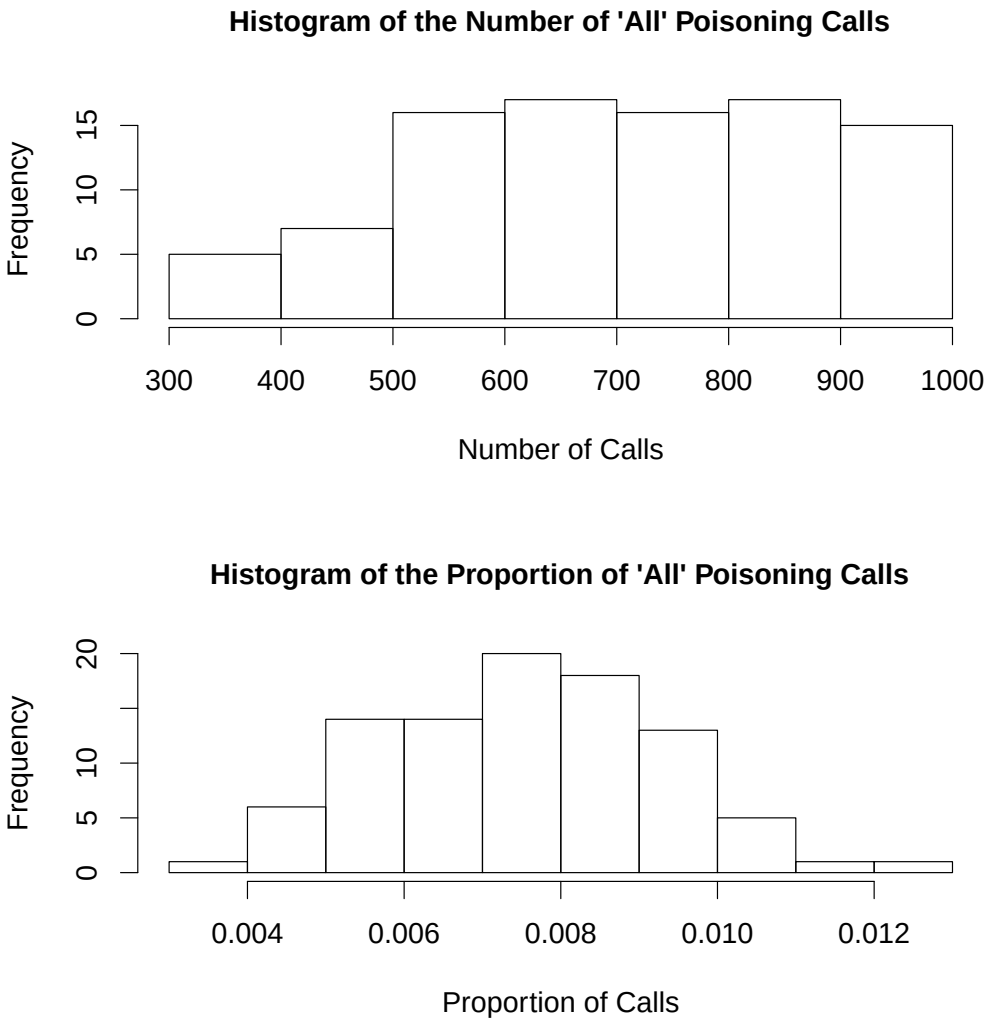


Figure 2.8: Histogram of the number (top) and proportion (bottom) of poisoning calls to NHS 24 in the "All" category per month

Figure 2.9 shows these data in three ways: (a) poisoning calls as a proportion of all NHS 24 calls, (b) the number of poisoning calls and (c) the total number of calls. The latter are provided to assess whether the trends present are a result of poisoning calls or if the trends are an artefact of temporal trends in the number of NHS 24 calls overall.

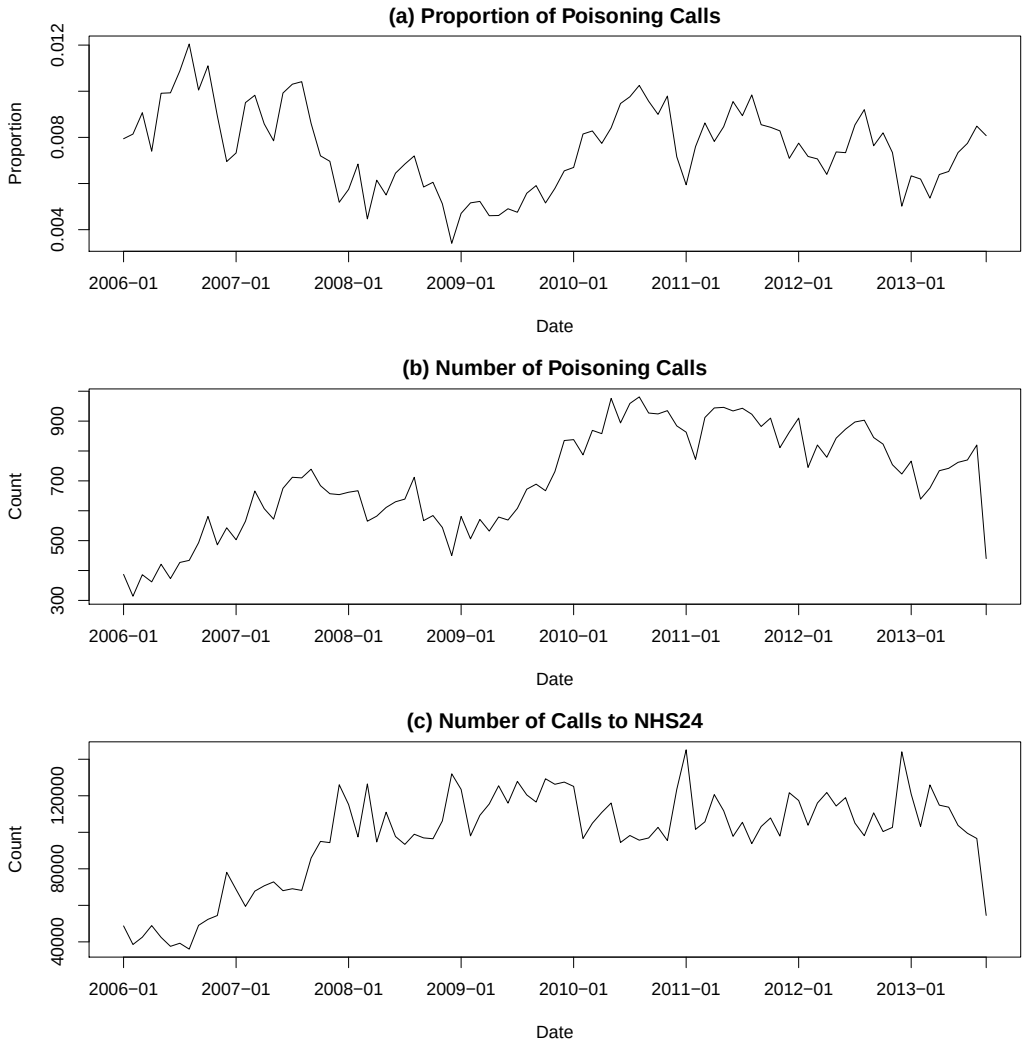


Figure 2.9: Plots showing the proportion of poisoning calls to NHS24 (a), the number of poisoning calls to NHS24 (b) and the total number of calls to NHS24 (c)

Figure 2.9 suggests that the proportion of calls concerning poisoning has a fairly strong seasonal trend, which appears to peak in August of each year, with the exception of 2009. This trend appears to be present in the overall number of poisoning calls, although it is much less clear until after 2011. There is also some evidence of an opposing trend in the overall number of NHS 24 calls, although this is not as obvious as that shown in the proportion of calls.

This figure also shows a gradual increase in call numbers between the inception of NHS 24 in 2006 and 2008, where the call numbers have gradually settled into a steady state. Another feature of note in these plots is the sharp drop off in calls in September of 2013. This is because the final date obtained in these data was the 17th September 2013, meaning that there is only a half a month's worth of calls recorded. This does not present an issue in these analyses as the focus is on modelling the proportion of calls, rather than the absolute number of calls.

From Figure 2.9, it seems that poisoning calls accounted for a slightly higher proportion of NHS 24 calls in 2006, the proportion then decreased until around 2009. Over the same period, the number of poisoning calls and NHS 24 usage overall were increasing, with NHS 24 use overall increasing at a greater rate than poisoning calls. The number of calls to NHS 24 overall levelled off around 2008, while the number of poisoning related calls increased between 2009 and mid 2010 before levelling off, resulting in an overall increase in the proportion of poisoning calls to NHS 24 over the same period. From 2010 onwards, the proportion of calls concerning poisoning has approximately levelled off, as have the number of poisoning calls and the overall number of calls to NHS 24.

In order to formally assess the trends discussed above, a generalised additive model was fit to these data of the form described in Equation 2.1. This initially assumed a poisson distribution for the number of poisoning calls. The ratio of the deviance to the residual degrees of freedom was found to be 7.44, which is clearly much larger than 1. This indicates that there is overdispersion in the data, hence the model was refit assuming that the data follow a negative binomial distribution. For this model, the dispersion factor for this model was found to be 0.999, indicating that the residuals are correctly dispersed.

The residuals were also assessed for autocorrelation at this stage. The autocorrelation and partial autocorrelation functions were plotted and are shown in Figure 2.10. This shows that there is significant correlation at lag 1 in both plots, though both bars only just lie above the confidence limit.

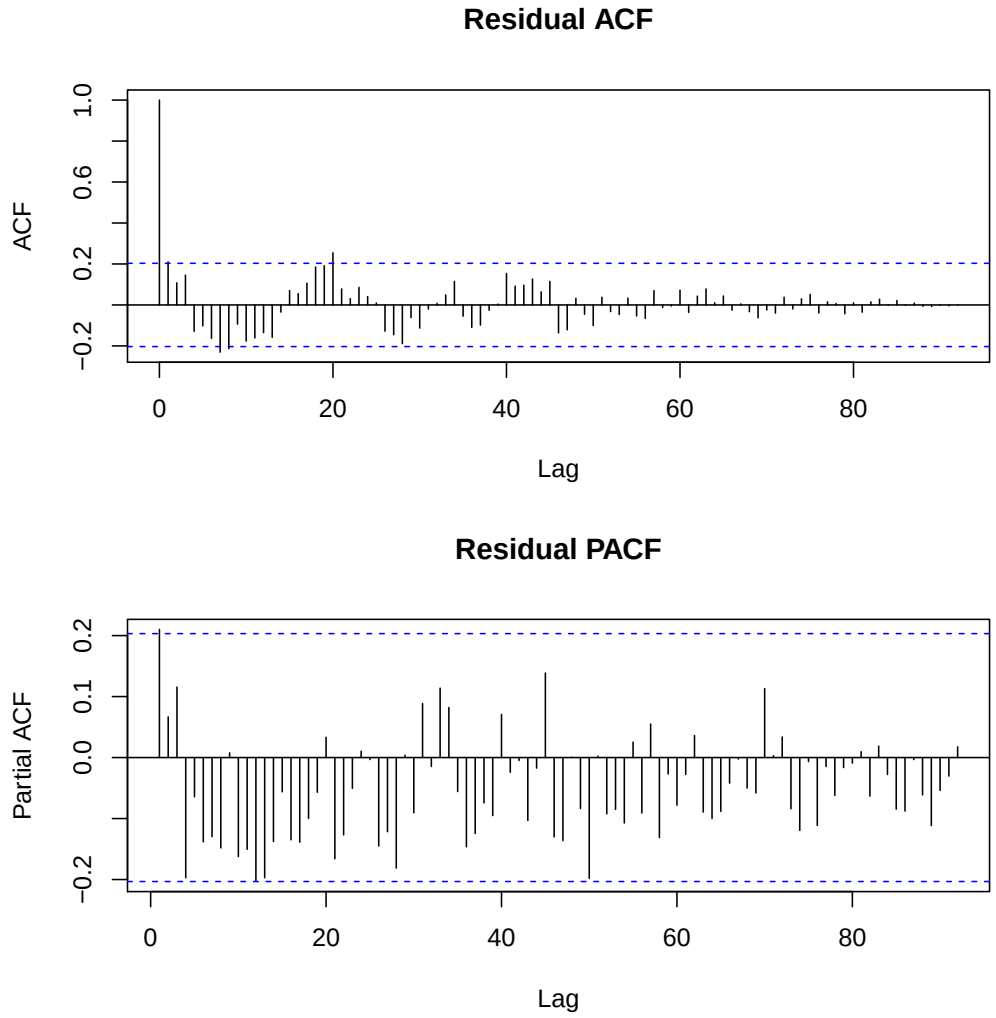


Figure 2.10: Plot of autocorrelation function plot (Top) and partial autocorrelation plot (Bottom)

In order to be conservative, the decision was made to incorporate an AR(1) correlation structure. Resulting in uncorrelated residuals as shown in Figure 2.11.

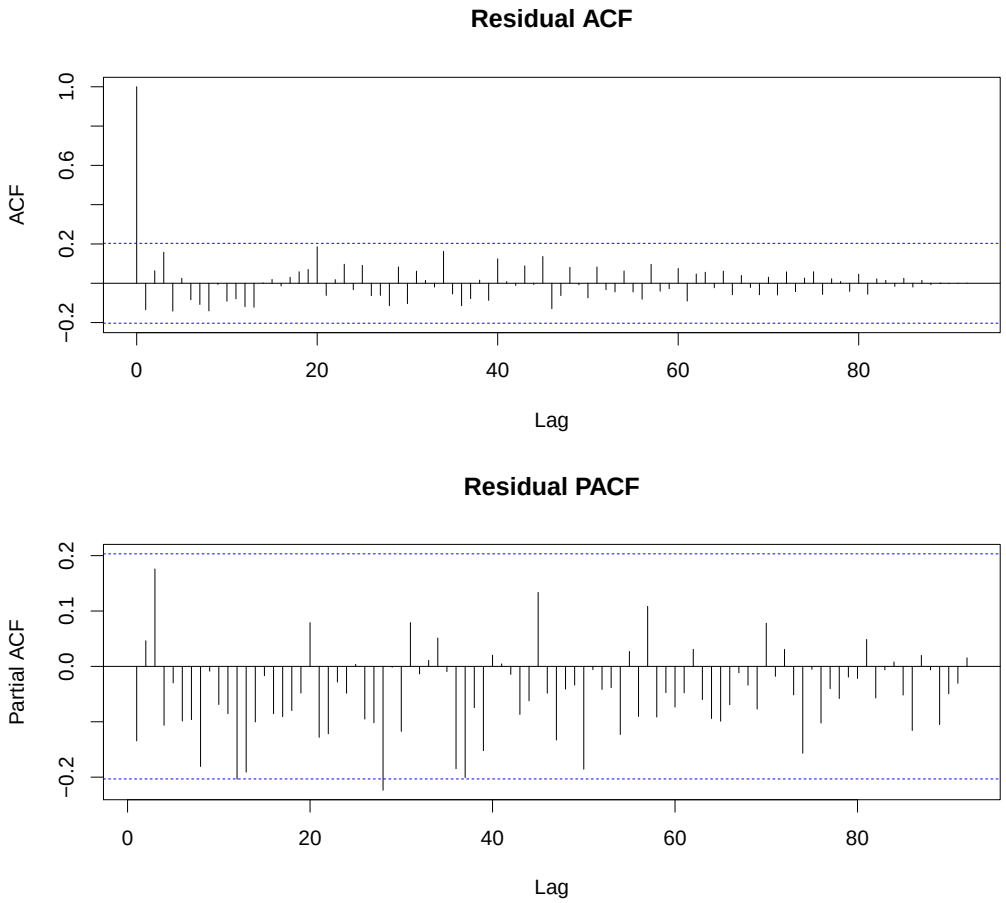


Figure 2.11: Plot of autocorrelation function plot (Top) and partial autocorrelation plot (Bottom)

The final model assumes a negative binomial distribution for the data and incorporates and AR(1) correlation structure in order to model the residual temporal autocorrelation. The fitted trends from this model are shown in Figure 2.12. The seasonal trend, on the left, shows that there is indeed annual variation in the proportion of calls to NHS 24. There is a small peak in February, which was not detected in Figure 2.9, which corresponds to an increase in call proportions of 12.5% on average (95% CI: 3.5%, 22.4%) from January. The seasonal trend peaks in August, before decreasing by 38.8% (95% CI: 27.4%, 51.1%) before December.

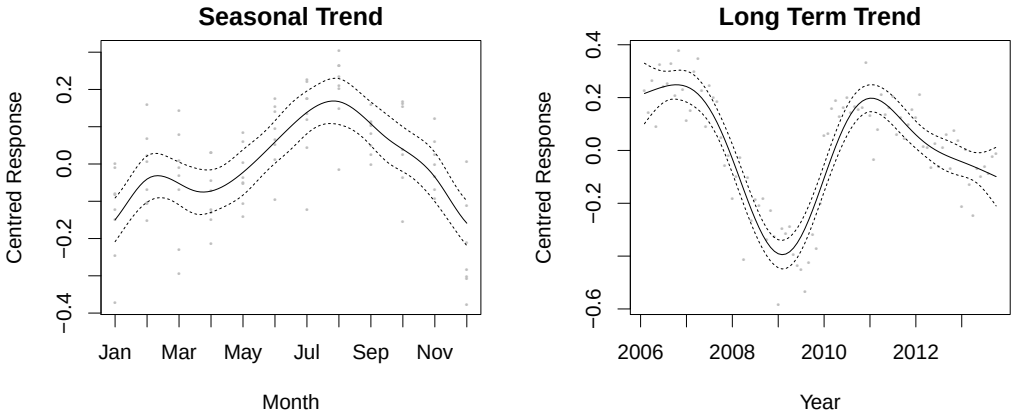


Figure 2.12: Plots showing the seasonal (left) and long term (right) trends estimated using a GAM

The long term trend, on the right of Figure 2.12, suggests that the proportion of poisoning calls is approximately level through 2006, before the call proportion decreases between 2007 and 2009 by 87.7% (95% CI: 73.2%, 103%). Call proportions then increase by 80.8% (95% CI: 67.8%, 94.8%) between 2009 and 2011. The call proportions then decrease by 34.6% (95% CI: 19.1%, 52.1%) from 2011 until 2013.

2.3.2 Spatial Analysis

The second step in this analysis was to get an impression of the spatial variation in NHS24 poisoning calls. To that end, the calls were aggregated across time, so that each region examined was related to a single, total number of poisoning calls, and total number of NHS 24 calls overall. These data were merged with the covariate data described previously, to provide the data shown in Table 2.8 on the next page.

Table 2.8: Table showing the structure of the data used in the spatial analysis

Region	All	Possible	Minimum	Total	Toddler	Child	Pensionable	Urban	Deprivation
Aberdeen Inner	1122	1053	781	190532	4.562	8.641	14.503	0.000	18.090
Aberdeen Outer	1286	1206	883	195227	5.771	13.983	15.545	10.083	8.103
Aberdeenshire North	1359	1276	913	180651	5.706	13.185	14.444	62.864	3.594
Aberdeenshire South	1481	1382	1009	177510	5.962	14.063	13.029	38.891	0.685
...	...	...	...	...	...	...	...	...	...
Stirling	1176	1098	774	143533	5.626	12.276	15.156	34.242	9.129
West Dunbartonshire	1065	980	783	143373	5.860	13.482	15.889	4.182	35.301
West Lothian	2589	2394	1763	340324	6.456	13.572	10.947	10.782	14.657
West Moray	230	216	176	26034	6.075	13.635	14.622	54.083	2.575

Initially, the distribution of poisoning calls was examined using histograms as can be seen in Figure 2.13. The number of calls in a region varied from a minimum of 8 in Arran & Cumbrae and 3,003 in Edinburgh East. The proportion of poisoning calls ranged from 0.005 in Arran & Cumbrae to 0.0139 in the Argyll & Bute Islands region. The majority of regions had a call proportion of around 0.007.

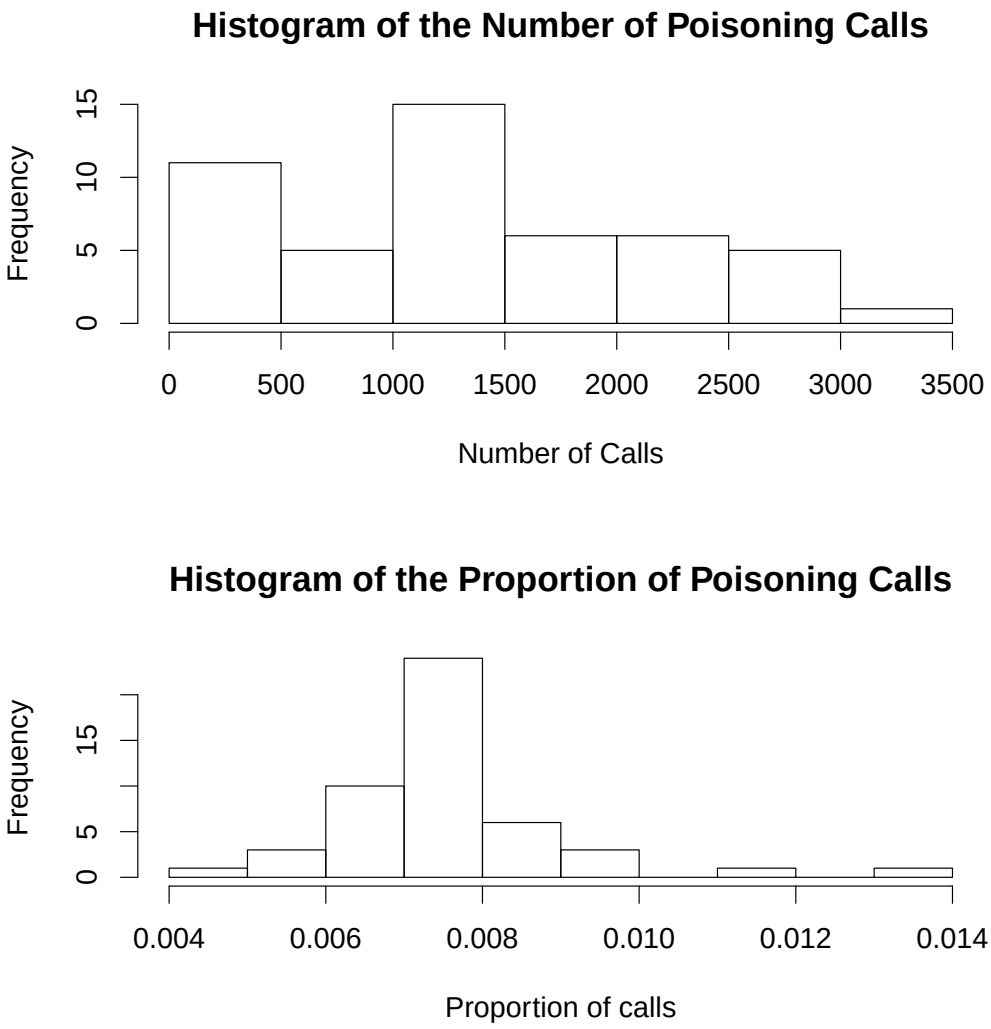


Figure 2.13: Histograms showing the distribution of the number (top) and proportion (bottom) of poisoning calls made to NHS 24 by region

The Argyll & Bute Islands region can be seen as the pink region in Figure 2.14. This map shows how the proportion of poisoning calls varies across Scotland, with dark green indicating a low proportion, yellow indicating a moderate proportion and pink indicating a high proportion relative to the other regions.

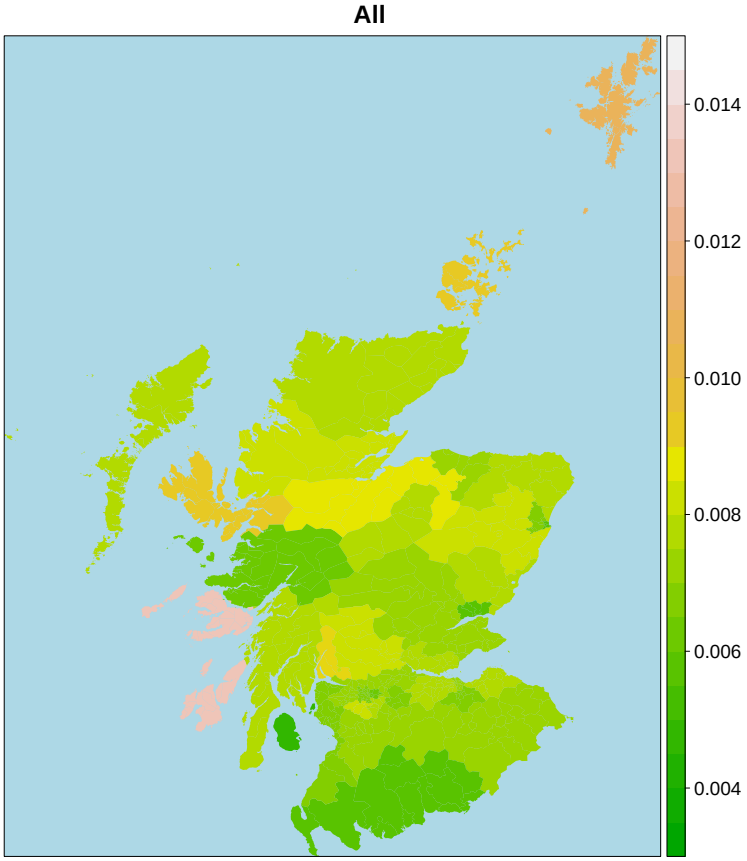


Figure 2.14: Map showing how the rate of poisoning calls in the "All" category per call to NHS24 for any reason vary across Scotland

This map shows that call proportions tend to be slightly higher in the north east of Scotland and seem to decrease towards the south west. The island regions tend to have higher call rates in general. In fact, the regions with the four highest percentages are island regions: Argyll & Bute Islands, Shetland Islands, Skye & Lochalsh and the Orkney Islands. The exception to this is Arran & Cumbrae, which has the lowest proportion of poisoning calls, as described previously.

The relationship between the covariates and the proportion of poisoning calls was examined, as shown in Figure 2.15. From this, there appears to be a positive relationship between the proportion of poisoning calls and the percentage of children and toddlers in the region, as well as the percentage of the population living in a rural area. The proportion of poisoning calls appears to decrease with an increasing proportion of adults of pensionable age and an increasing proportion of the population living in deprived areas. This set of plot also demonstrates that it is reasonable to assume a linear relationship between each of the covariates and the proportion of poisoning calls.

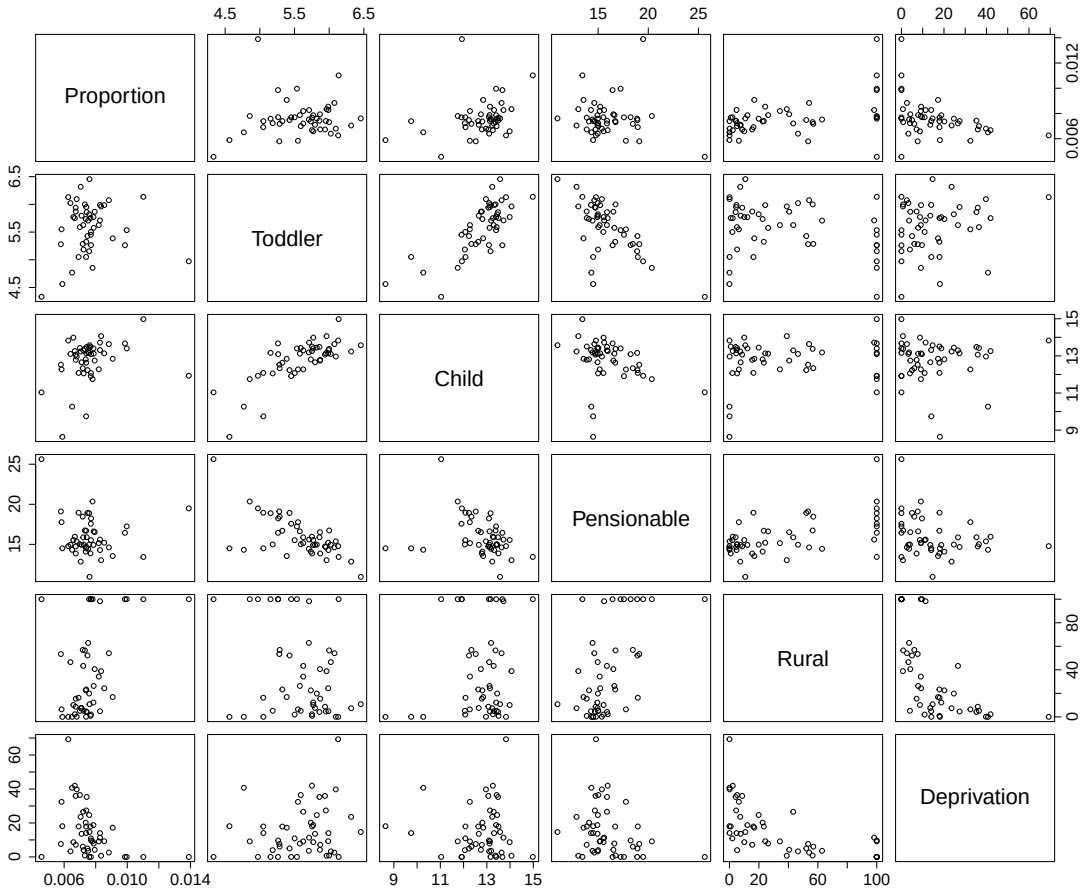


Figure 2.15: Plot showing the relationships between the proportion of poisoning calls and the covariates. Each point within each panel describes one of the 49 regions used in the analysis

Figure 2.15 also shows the relationship between each of the independent variables. The correlations do not appear to be particularly strong, with the exception of the "Toddler" variable with both the "Child" and "Pensionable" variables. These correlations were 0.751 and -0.663 respectively.

The variables described in the methods section were used as covariates in a generalised linear model. As in the temporal analysis, the data were initially assumed to follow a poisson distribution. For the poisson model incorporating all covariate effects, the dispersion parameter was large at 8.269, indicating overdispersion. A negative binomial model was implemented, and the model with all of the covariates in it had a dispersion parameter of 1.220.

The initial model fit with all covariates had parameter estimates as shown in Table 2.9, alongside their associated p-values.

Table 2.9: Table showing coefficient estimates on the model scale for the full model

	Estimate	**Standard Error**	**P-value**
Child	0.0491	0.0223	0.0278
Toddler	-0.0908	0.0648	0.1616
Pensionable	-0.0305	0.0095	0.0014
Urban/Rural	0.0015	0.0006	0.0136
Deprivation	-0.0026	0.0011	0.0221

The only term removed from the model was the "Toddler" variable, which had a p-value of 0.1616. In the model excluding this term, the "Child" variable had a p-value of 0.0715, which, although not significant at the 5% level, is still very small, and the decision was made to keep this term in this model.

Table 2.10 shows the effect of each independent variable in the final model on the call proportion. The table presents the exponent of the coefficients, which reflect the percentage change in call proportions for a unit increase in each of the covariates.

Table 2.10: Table showing coefficient estimates on the scale of the data and residual Moran's I results

	Estimate	**95% Confidence Interval**
Child	2.42%	-0.02%, 5.13%
Pensionable	-2.41%	-4.01%, -0.78%
Urban/Rural	0.18%	0.06%, 0.30%
Deprivation	-0.26%	-0.48%, -0.03%
	Moran's I	**P-value**
Overall	0.090	0.260
Minus Islands	0.067	0.412

The two age variables have effects of similar size, although in opposite directions. An increase of 1% in the percentage of children living in a region increases the proportion of poisoning calls by 2.42% on average, while an increase in 1% in the percentage of adults of pensionable age living in a region decreases the proportion of poisoning calls by 2.41% on average. A greater percentage of the population living in a rural or remote area was linked to an increase of 0.18% in the proportion of poisoning calls for every increase of 1%. The deprivation variable was associated with a decrease in the proportion of poisoning calls, such that for every 1% increase in the percentage of the population living in a deprived area, the proportion of poisoning calls decreased on average by 0.26%.

Moran's I, shown in Table 2.10, was used to assess residual spatial autocorrelation. The non-significant permutation test indicates that the observed value of I for the residuals is consistent with the assumption of spatial independence.

The model predictions and residuals are shown in the maps in Figure 2.16, with the model predictions in the left panel and the model residuals in the right panel. The model appears to capture some of the north east to south west trends described previously, with higher predicted values in the north east and slightly lower predicted values in the south west. However, the model overestimates

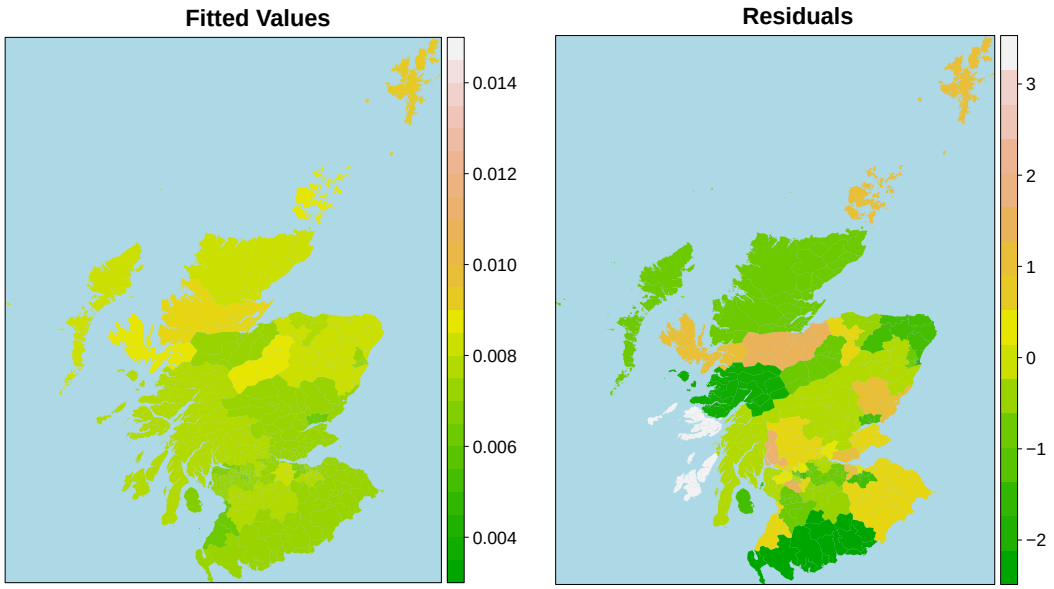


Figure 2.16: Model fitted values (left) and model residuals (right)

call proportions in Dumfries and Galloway, while greatly underestimating call proportion in the Argyll and Bute Islands region (shown in white in the residual map).

The residual map in the right panel shows regions with negative residuals, those which have been overestimated, in dark green and those with positive residuals, those which have been underestimated, in orange, pink and white. This shows a random scatter of colours across the map, reflecting the result given by the permutation tests in Table 2.10 that there is no residual spatial autocorrelation.

2.3.3 Spatio-temporal Analysis

The final stage in this piece of work was to carry out a full spatio-temporal analysis on the NHS 24 call data. This involved fitting a model of the form described in Equation 2.4. The model selection process was done as in the spatial analysis. The coefficients (where appropriate) and p-values for the full model are shown in Table 2.11.

Table 2.11: Table showing coefficient estimates on the model scale, alongside the model p-values

	Estimate	**Standard Error**	**P-value**
Toddler	-0.053	0.0234	0.024
Child	0.036	0.0079	4.63×10^{-6}
Pensionable	-0.025	0.0035	4.41×10^{-13}
Urban/Rural	0.001	0.0003	4.53×10^{-4}
Deprivation	-0.003	0.0004	7.54×10^{-15}
s(Month)	—	—	$< 2 \times 10^{-16}$
s(Year)	—	—	$< 2 \times 10^{-16}$

Based on Table 2.11, there were no non-significant terms in the model. The residual autocorrelation for this model was assessed using Moran's I, incorporating the spatio-temporal adjacency matrix as described in Section 2.2.3. This model exhibited residual autocorrelation, with $I = 0.049$ ($P = 1.84 \times 10^{-9}$), indicating significant, though small, residual autocorrelation.

The next model incorporated a temporally smooth spatial structure, allowing a different spatial structure for each observed month of data. The residual correlation was still found to be significant, with $I = 0.039$ ($P = 1.27 \times 10^{-6}$). Having accounted for the spatial dependency structure, it was then likely that this residual autocorrelation was largely due to temporal, rather than spatial structure, hence a model which additionally incorporated temporal structure was necessary. In this setting, it is difficult to visualise the appropriate parameters

for the temporal ARMA process. Hence, the fixed effect component of the model was held fixed and Akaike's Information Criterion (AIC) was minimised in order to select an appropriate temporal structure, as can be seen in Table 2.12.

Table 2.12: Table showing AIC values for different temporal structures

Temporal Structure	AIC
AR(1)	4502.121
AR(2)	4495.433
AR(3)	4499.666
MA(1)	4502.284
MA(2)	4493.496
MA(3)	4496.790
ARMA(1,1)	4512.883
ARMA(1,2)	4496.044
ARMA(1,3)	4498.657
ARMA(2,1)	4502.529
ARMA(2,2)	4505.764
ARMA(2,3)	4506.391
ARMA(3,1)	4507.015

From this, there were two models which performed similarly: AR(2) and MA(2) had the two lowest AIC values. Since the AIC values for each of these was so similar, the AR(2) model was carried forward in order to remain consistent with the temporal analysis, and the fixed effect part of the model was re-evaluated.

The coefficient estimates and their associated p-values for the full model with spatio-temporal structure are shown in Table 2.13

Table 2.13: Table showing coefficient estimates and residual Moran's I results

	Estimate	Standard Error	P-Value
Toddler	-0.0707	0.0673	0.2939
Child	0.0394	0.0229	0.0857
Pensionable	-0.0259	0.0094	0.0060
Urban/Rural	0.0009	0.0007	0.2223
Deprivation	-0.0032	0.0011	0.0050

This shows that some of the model coefficients which were highly significant in the model ignoring spatial and temporal autocorrelation, are now non-significant. The non-significant independent variables were removed from the model, starting with the "Toddler" variable which had the largest P-value of 0.294. The model was refit and it was found that the "Urban/Rural" variable remained non-significant, resulting in its removal from the model with a p-value of 0.224. The model was refit for a third time and, while not significant at the 5% level, the "Child" coefficient had a relatively small p-value of 0.097, and, in order to be conservative, the decision was made to retain this term in the model. The spatial covariate estimates from the final model are shown in Table 2.14, along with Moran's I. The temporal trends are presented in Figure 2.17.

Table 2.14: Table showing coefficient estimates on the data scale and residual Moran's I results

	Estimate	**95% Confidence Interval**
Child	2.26%	-0.41%, 5.00%
Pensionable	-1.52%	-2.85%, -0.16%
Deprivation	-0.39%	-0.60%, -0.18%
	Moran's I	**P-value**
Overall	0.0131	0.1037
Minus Islands	0.0180	0.0279

In this model, the coefficient estimate for the "Child" covariate was found to suggest that a 1% increase in the percentage of children in a region was associated with an average increase in poisoning calls of 2.26%, a result which was not significant at the 5% level. The pensionable covariate was statistically significant, such that an increase of 1% in the percentage of pensionable age adults decreases the call proportion by 1.52% on average. This model also indicates a significant relationship between deprivation and poisoning calls, such that an increase of 1% in the percentage of the population living in a deprived area indicates an average decrease of 0.39% in poisoning calls to NHS 24.

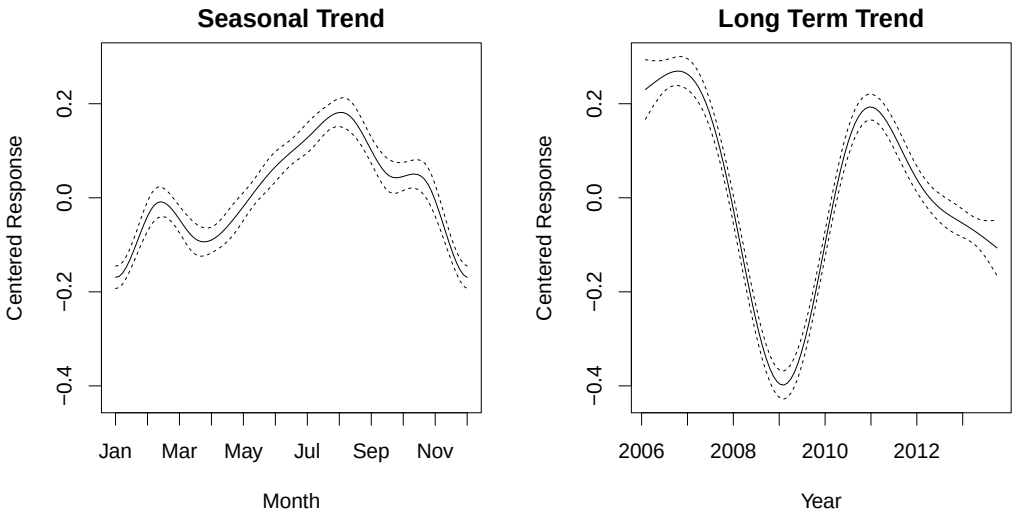


Figure 2.17: Estimated seasonal (left) and long term (right) trends from the spatio-temporal model

The temporal trends seen in Figure 2.17 are similar to those seen in the temporal analysis. The within year trend shows a small peak around February, with the overall peak in poisoning calls occurring in August. In addition to these, there appears to be another small peak in October, which was not present in the temporal analysis.

The long term trend shows a slight increase in the proportion of poisoning calls throughout 2006, there was then a sharp decrease in the proportion of calls between 2007 and 2009. Call proportions then increased once more between 2009 and 2011, before decreasing again until 2013.

2.4 Discussion

This analysis sought to assess whether there are any consistent trends in poisoning, by making use of routinely collected information on calls to NHS 24. Initially, looking at the call data temporally indicated that there is some seasonality in the proportion of calls to NHS 24. This trend could be driven by two things: season-

ality in the actual poisoning calls themselves, or seasonality in the use of the NHS 24 phone information system as a whole. By examining plots of the time series of these two factors there is a suggestion that it may, in fact, be a combination of the two which drives the seasonality.

American poison centres handle calls made by the public which relate to cases of poisoning. In their 2015 annual report, the American Association of Poison Control Centres presented data on the number of "exposure calls" (i.e. those calls which relate to a human exposure to a poisonous substance), which showed a clear seasonal trend in these calls, which peaked in summer [67]. This peak in cases of poisoning over the summer may be explained by school holidays. One study found that emergency attendances were generally higher over the summer for ages 1-14 [68]. The summer peak may also be partly due to the recorded presence of drugs at festivals [69]. However, without more specific details on the calls these hypotheses cannot be tested.

Poisoning calls overall appeared to be higher in the summer, however it is likely that the seasonality in the proportion of calls was emphasised by the fact that winter illnesses mean that NHS 24 calls are higher over winter [70]. The fact that calls overall are higher in winter, in addition to the fact that poisoning calls are lower in winter, means that the effect of the reduction in the proportion of poisoning calls is increased.

The relationship found with the demographic variables in the spatio-temporal analysis was surprising, and these demographics only accounted for small levels of variability in the proportion of poisoning calls. In Chapter 1 several studies were identified which indicated higher risk for poisoning in the most deprived areas [9, 11, 19], however an increasing proportion of the population living in deprived

areas was found to indicate a small reduction in the proportion of poisoning calls, which may indicate that these individuals may tend to attend hospital rather than contact NHS 24.

In the marginal spatial analysis the urban/rural covariate was significant, and indicated an increase in the proportion of calls as the proportion of the population living in rural or remote areas increased. This may be due to the link between rurality and deprivation, as deprivation takes into account factors such as overcrowding and crime, which are much more common in urban areas than in rural areas, meaning that those which have a greater proportion in rural areas are more likely to have a lower proportion in deprived areas.

In both the marginal spatial and the spatio-temporal model fits, an increase in the proportion of the population of pensionable age was found to indicate a decrease in the proportion of poisoning calls to NHS 24. This is likely because it is generally younger adults who are at the highest risk of poisoning, with adults aged 15-35 being at most risk from deliberate self poisoning and children under five being at most risk from unintentional self-poisoning [7].

AR(p) correlation structures were used to account for residual temporal autocorrelation that was found within the data. This type of autocorrelation indicates that values are related to the preceding p observations, and in both the spatial and spatio-temporal models, the parameter p was estimated to be 1 and 2 respectively. This is an indication that some systematic month to month variability was not being captured by these models, which may be related to lacking precise information of population dynamics in these regions each month.

This piece of analysis has provided an interesting basis for the following chapter outlining trends in TOXBASE use in emergency departments across the UK. Comparing this analysis with the TOXBASE analysis will allow for some insight into whether the trends discussed here are true trends in poisoning or whether these trends are caused by other factors specific to NHS 24 use.

Chapter 3

Trends in accesses to the TOXBASE Database

This chapter focusses on examining the temporal trends that are present in accesses to the TOXBASE database. Gaining an understanding of how TOXBASE use varies is an integral part of examining the impact of TOXBASE on patient management and will help to provide insight into the analysis presented in Chapters 6 and 7. This piece of work has been published in the journal *Human and Experimental Toxicology* [55].

The main aim of the analysis outlined in this chapter is to examine the underlying long-term and seasonal trends in accesses to TOXBASE. Due to changes in the database system in 2008, the analysis is limited to accesses made between January 2008 and December 2015. Due to the large proportion of accesses to TOXBASE from hospital emergency departments [1], it was of interest to extract data from TOXBASE pertaining to these key TOXBASE users in Great Britain. To that end this piece of work details trends in accesses to TOXBASE from emergency departments across Great Britain.

Figure 3.1: Map showing locations of the English government office regions

The analysis initially concentrated on assessing accesses from emergency departments across Britain as a whole. The data were aggregated up, in order to create one complete series of daily accesses. Following this, analysis was carried out at a regional level. Scotland and Wales were taken as two regions, while England was separated into its Government Office Regions: North East, North West, Yorkshire & the Humber, East Midlands, West Midlands, East of England, South East, South West and London. The relative locations of these regions can be seen in Figure 3.1.

The initial step in carrying out this analysis involved making sure that the data were complete and correct. This process is outlined in section 3.1, and was undertaken in conjunction with the database manager based in the National

Poisons Information Service, Edinburgh. The methodology used will be outlined in Section 3.2. These methods are similar to those in Chapter 2, therefore the focus will be on the specifics of the models fit. The results of the analysis are then provided for both the overall analysis and for the regional level analysis. The chapter concludes with a more detailed discussion of the work.

3.1 Data Cleaning

3.1.1 Description of the Database

The data held on TOXBASE use was available from 1998, when TOXBASE was first moved online, up to the end of 2015. This came in the form of a Microsoft Access database with access details stored in separate tables by year, with slight changes in the structre of these table after 2008, when the TOXBASE information database was moved to a new server. A sample of the product access tables from 2006 and 2008, named ProdAcc2006 and ProdAcc 2008 respectively, is shown in Figure 3.2.

ProdAcc2006

ProductAccessI	UserID	AccessDateTime	ProductID
1962499	NHSD18	01/01/2006 00:00:36	381
1947963	gp101	01/01/2006 00:01:31	870
1947957	gp101	01/01/2006 00:03:34	870
1957385	NHSD9	01/01/2006 00:04:36	13596
1947965	gp101	01/01/2006 00:04:41	870
1952119	H204	01/01/2006 00:05:10	13486
1952122	H204	01/01/2006 00:05:30	13486
1946775	H1805	01/01/2006 00:05:58	15518
1962500	NHSD18	01/01/2006 00:07:14	381
1952129	H204	01/01/2006 00:07:45	13486
1961546	H1385	01/01/2006 00:09:10	1051

ProdAcc2008

ProductAccessID	UserID	AccessDateTime	ProductID	EPiServerPage
518387	NHSD10	23/12/2008 13:41:05	86575	87915
526969	NHSD1	25/12/2008 10:05:07	86579	87909
526970	NHSD1	25/12/2008 10:05:07	86579	87909
529341	NPISCARD	27/12/2008 09:57:33	86579	87909
529340	AMB258	27/12/2008 09:58:11	86579	87909
534576	H272	29/12/2008 00:19:50	86579	87909
534577	H272	29/12/2008 00:25:38	86579	87909
510636	H152	20/12/2008 03:43:52	86566	87896
510638	H739	20/12/2008 14:42:36	86566	87896
510637	H739	20/12/2008 14:42:36	86566	87896
513011	GP895	21/12/2008 05:54:43	86566	87896

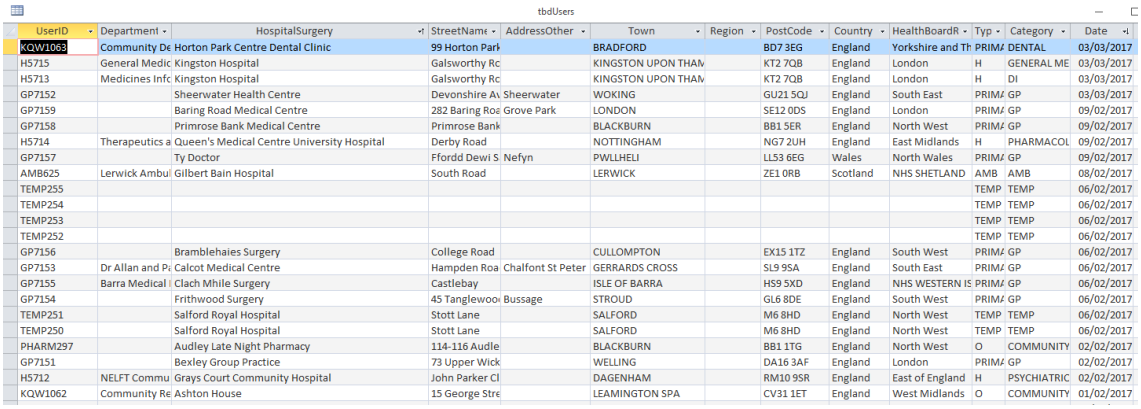
Figure 3.2: Image showing a sample of observations from the product access tables for 2006 and 2008 (ProdAcc2006 and ProdAcc2008)

All of the product access tables have 4 columns. The first is 'ProductAccessID', which is a unique identifier for each access to a page in TOXBASE. The second column is 'UserID' which acts as a link to another table which contains more detail on each registered TOXBASE user. The third column, 'AccessDateTime'

contains a timestamp for each individual page access, and the fourth contains the 'ProductID' which links to more detailed product information. Each product access table from 2008 onwards contains an additional column EPiServerPage, which is an identifier relevant to the storage of the database.

Each row of these tables is called an access, which is when a user looks at a specific page within TOXBASE. A group of accesses made by one user pertaining to one search is called a session, for example in ProdAcc2006, there were two accesses made by user H204 20 seconds apart, which would be considered a session. However, it is difficult to distinguish the start and end points of these sessions. There was another access to H204 two minutes after the initial two accesses, which could be part of the same user session. However, user IDs correspond to a location or team, rather than an individual, so if the hospital is busy, there may be multiple individuals using the same user ID to look up information for different patients, so this access may not be part of the same user session. Due to the difficulties, and subjectivity, in extracting sessions, it was decided that for this project it would be more appropriate to look at accesses.

The information on individual registered TOXBASE users is stored in a table called tbdUsers. The UserID column in the product access tables links to the primary key of tbdUsers, also called UserID. Figure 3.3 shows a sample of rows from the user information table. The first column contains the 'UserID'. The second column contains information on the department that the ID corresponds to and third column 'HospitalSurgery' contains information on the location that the ID corresponds to. The next six columns contain address information, including postcode and country. This is followed by 'HealthBoardRegion' information, which has been used in this chapter. The next two columns provide information on the type of user: 'TypeOfUser' and 'Category'. The final column, 'Date', provides information on when the user registered with TOXBASE.

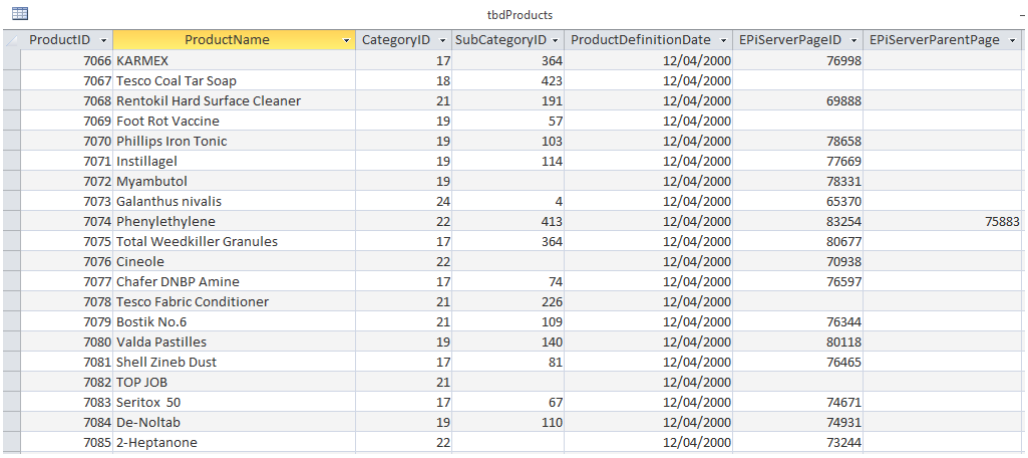


tbdUsers

UserID	Department	HospitalSurgery	StreetName	AddressOther	Town	Region	PostCode	Country	HealthBoardR	Typ	Category	Date
KQW1063	Community De	Horton Park Centre Dental Clinic	99 Horton Park		BRADFORD		BD7 3EG	England	Yorkshire and Th	PRIMA	DENTAL	03/03/2017
H5715	General Medic	Kingston Hospital	Galsworthy Rc		KINGSTON UPON THAM		KT2 7QB	England	London	H	GENERAL ME	03/03/2017
H5713	Medicines Infc	Kingston Hospital	Galsworthy Rc		KINGSTON UPON THAM		KT2 7QB	England	London	H	DI	03/03/2017
GP7152		Sheerwater Health Centre	Devonshire Av	Sheerwater	WOKING		GU21 5QJ	England	South East	PRIMA	GP	03/03/2017
GP7159		Baring Road Medical Centre	282 Baring Roa	Grove Park	LONDON		SE12 0DS	England	London	PRIMA	GP	09/02/2017
GP7158		Primrose Bank Medical Centre	Primrose Bank		BLACKBURN		BB1 5ER	England	North West	PRIMA	GP	09/02/2017
H5714	Therapeutics a	Queen's Medical Centre University Hospital	Derby Road		NOTTINGHAM		NG7 2UH	England	East Midlands	H	PHARMACOL	09/02/2017
GP7157		Ty Doctor	Ffordd Dewi S	Nefyn	PWLLHELI		LL53 6EG	Wales	North Wales	PRIMA	GP	09/02/2017
AMB625	Lerwick Ambul	Gilbert Bain Hospital	South Road		LERWICK		ZE1 0RB	Scotland	NHS SHETLAND	AMB	AMB	08/02/2017
TEMP255										TEMP	TEMP	06/02/2017
TEMP254										TEMP	TEMP	06/02/2017
TEMP253										TEMP	TEMP	06/02/2017
TEMP252										TEMP	TEMP	06/02/2017
GP7156		Bramblehaies Surgery	College Road		CULLOMPTON		EX15 1TZ	England	South West	PRIMA	GP	06/02/2017
GP7153	Dr Allan and Pa	Calcot Medical Centre	Hampden Roa	Chalfont St Peter	GERRARDS CROSS		SL9 9SA	England	South East	PRIMA	GP	06/02/2017
GP7155	Barra Medical	Clach Mhile Surgery	Castlebay		ISLE OF BARRA		HS9 5XD	England	NHS WESTERN IS	PRIMA	GP	06/02/2017
GP7154		Frithwood Surgery	45 Tanglewoo	Bussage	STROUD		GL6 8DE	England	South West	PRIMA	GP	06/02/2017
TEMP251		Salford Royal Hospital	Stott Lane		SALFORD		M6 8HD	England	North West	TEMP	TEMP	06/02/2017
TEMP250		Salford Royal Hospital	Stott Lane		SALFORD		M6 8HD	England	North West	TEMP	TEMP	06/02/2017
PHARM297		Audley Late Night Pharmacy	114-116 Audle		BLACKBURN		BB1 1TG	England	North West	O	COMMUNITY	02/02/2017
GP7151		Bexley Group Practice	73 Upper Wick		WELLING		DA16 3AF	England	London	PRIMA	GP	02/02/2017
H5712	NELFT Commu	Grays Court Community Hospital	John Parker Cl		DAGENHAM		RM10 9SR	England	East of England	H	PSYCHIATRIC	02/02/2017
KQW1062	Community Re	Ashton House	15 George Stre		LEAMINGTON SPA		CV31 1ET	England	West Midlands	O	COMMUNITY	01/02/2017

Figure 3.3: Image showing a sample of users from the user information table (tbdUsers)

Information on the products being accessed is contained within a table called tbdProducts, which is linked to the product access tables via its primary key 'ProductID'. Each row in the product information table corresponds to a specific page within TOXBASE, a sample of which can be seen in Figure 3.4.



tbdProducts

ProductID	ProductName	CategoryID	SubCategoryID	ProductDefinitionDate	EPiServerPageID	EPiServerParentPage
7066	KARMEX	17	364	12/04/2000	76998	
7067	Tesco Coal Tar Soap	18	423	12/04/2000		
7068	Rentokil Hard Surface Cleaner	21	191	12/04/2000	69888	
7069	Foot Rot Vaccine	19	57	12/04/2000		
7070	Phillips Iron Tonic	19	103	12/04/2000	78658	
7071	Instillagel	19	114	12/04/2000	77669	
7072	Myambutol	19		12/04/2000	78331	
7073	Galanthus nivalis	24	4	12/04/2000	65370	
7074	Phenylethylene	22	413	12/04/2000	83254	75883
7075	Total Weedkiller Granules	17	364	12/04/2000	80677	
7076	Cineole	22		12/04/2000	70938	
7077	Chafer DNBP Amine	17	74	12/04/2000	76597	
7078	Tesco Fabric Conditioner	21	226	12/04/2000		
7079	Bostik No.6	21	109	12/04/2000	76344	
7080	Valda Pastilles	19	140	12/04/2000	80118	
7081	Shell Zineb Dust	17	81	12/04/2000	76465	
7082	TOP JOB	21		12/04/2000		
7083	Seritox 50	17	67	12/04/2000	74671	
7084	De-Noltab	19	110	12/04/2000	74931	
7085	2-Heptanone	22		12/04/2000	73244	

Figure 3.4: Image showing a sample of rows from the product information table (tbdProducts)

The first column is the primary key 'ProductID', while the second column, 'ProductName' provides the name of the product. The next two columns, 'CategoryID' and 'SubCategoryID' provide information that allows for the analysis

of similar products. The Category information table, tbdProductCategory, was used in Chapter 7 and is shown in Figure 3.5. The subcategory information allows for finer scale categorisation of products than the categories do; this information was also used in Chapter 7 and is shown in Figure 3.6.

The fifth column in the product information table is 'ProductDefinitionDate' which defines the date at which the page entered TOXBASE. The final two columns contain identifiers which correspond to the storage TOXBASE. If the 'EPiServerPageID' is empty, then this indicates that the product was introduced into, and then removed from the database prior to 2008, though these entries are still required in the product information table to link to earlier years. EPiServerParentPage is an indicator that the particular search result that was clicked on takes you to another page, which is essentially an umbrella for a variety of substances. For example Dettol disinfectants have different "pages" for different scents, but the information for all of these is contained within one page for Dettol disinfectants.

tbdProductCategory

CategoryID	CategoryName
7	Miscellaneous
12	Veterinary products
16	Biologicals
17	A - agrochemicals
18	C - cosmetics
19	D - pharmaceuticals
20	F - fungi
21	H - household
22	I - chemicals
23	M - micro-organisms
24	P - plants
26	Z - animals

Figure 3.5: Image showing the twelve product categories as stored in tbdProductCategory

The broadest categories within TOXBASE are shown in 3.5. There are 12 categories: miscellaneous, veterinary products, biologicals, agrochemicals, cosmetics, pharmaceuticals, fungi, household, chemicals, micro-organisms, plants and animals. These can be further split into 401 sub-categories, a subset of which are shown in Figure 3.6, where the first column contains a sub-category ID, the second column contains the sub-category name and the third column contains the category ID of the relevant category.

tbdProductSubCategory

SubCategoryID	SubCategoryName	CategoryID
9	D 9.5.3 Fluoride	19
17	D Herbal	19
18	Aquarium product	7
19	SCRA Drugs of abuse	19
20	Amfetamines and similar stimulants Drugs of abu:	19
21	Piperazines Drugs of abuse	19
22	Synthetic Opioids Drugs of abuse	19
23	Tryptamines Drugs of abuse	19
24	Benzodiazepines Drugs of abuse	19
25	Synthetic Cocaines Drugs of abuse	19
26	Unknown drug of abuse	19
27	Dissociative drugs of abuse	19
28	D 5.3 Antiviral drugs	19
29	Lysergic acid diethylamide and related agents	19
30	Quinazoline derivatives	19
56	Petroleum distillate	21
57	Veterinary products - Drugs	19
58	Veterinary products - Pesticides	17
59	Veterinary products - Household	21
60	D 4.3.1 Tricyclics	19

Figure 3.6: Image showing a subset of the 401 product categories as stored in tbdProductSubCategory

3.1.2 Data Cleaning

In order to ensure that the tables within the database were correct, the full data set was read into `R`, via a connection to the TOXBASE access database which was set up using the `RODBC` package in `R`. The data set was then read in and merged using SQL as follows:

```
SELECT *
FROM ProdAcc2008 p, tbdProducts t, tbdUsers u
WHERE p.ProductID = t.ProductID AND
                UCASE(p.UserID) = UCASE(u.UserID)
```

Upon reading in the data, it became apparent that some of the data were being excluded. The issue was that, as time had passed and the storage of the database had evolved, some past users no longer have their information stored on the system and similarly, older products have disappeared from the current version of the database. This meant that, as the tables were being merged through the SQL procedure, accesses which had no match in tbdProducts or tbdUsers were being excluded.

In order to determine which items were being lost, the product access tables were read in individually as follows (taking the 2008 accesses as an example):

```
SELECT *
FROM ProdAcc2008
```

A copy of these which was merged with the user information was read in:

```
SELECT *
FROM ProdAcc2008 p, tbdProducts t
WHERE p.ProductID = t.ProductID
```

and a copy merged with the product information was read in:

```
SELECT *
FROM ProdAcc2008 p, tbdUsers u
WHERE UCASE(p.UserID)=UCASE(u.UserID)
```

The `UCASE()` function was used as username is not case sensitive, and is sometimes lowercase in the product access tables.

By assessing which product IDs were in the product access tables, but not in the versions that had been merged with the product information table, it was relatively straightforward to find the IDs of the products missing from the product table. A similar procedure was followed to find the user IDs which were not present in the user information table.

The linkage between the old and new databases was carried out by the database manager at NPIS. The missing product information was found using the EPi Server Page ID where possible. However, as this information was not available prior to 2008 when the new system was implemented, there were a large number of observations which could not be resolved. At this stage, the decision was made to use the TOXBASE data from 2008 onward only.

3.1.3 Data Used for Analysis

Following the data cleaning process, the TOXBASE access data pertaining to emergency department users were read in to `R` using SQL. This meant restricting the values being read from 'tbdUsers' to those where the 'Category' field was 'A&E'. Since the primary focus of this analysis was hospitals in Great Britain, the data were also restricted such that the 'Country' field in tbdUsers was equal to one of Scotland, England or Wales. The SQL code to implement this for the year 2008 was as follows:

```
SELECT ProductAccessID, p.UserID, AccessDateTime, t.CategoryID
FROM ProdAcc2008 p,  tbdUsers u, tbdProducts t
WHERE p.ProductID = t.ProductID AND
      UCASE(u.UserID) = UCASE(p.UserID) AND
      u.Category='A&E' AND
      u.Country IN ('Scotland','England','Wales')
```

The columns retained were restricted to those which were likely to be used in the following analyses.

Specific product types were also assessed for comparison with the overall seasonal trends. The first group chosen was antidepressants, which are the most commonly accessed type of drug. Agrochemicals were also selected as they are likely to have a strong seasonal component. The final category examined was drugs of abuse which represents one of the most commonly accessed categories.

These were each extracted in different way. The agrochemicals accesses were the most straightforward, as these could be obtained directly from the data set produced above, by extracting all observations for which 'CategoryID' was 17. The SQL code for this is as follows

```
SELECT ProductAccessID, p.UserID, AccessDateTime, t.CategoryID
FROM ProdAcc2008 p,  tbdUsers u, tbdProducts t
WHERE p.ProductID = t.ProductID AND
      UCASE(u.UserID) = UCASE(p.UserID) AND
      u.Category='A&E' AND
      u.Country IN ('Scotland','England','Wales') AND
      t.CategoryID = '17'
```

The other two categories had to be extracted from the database by referencing the product sub-category. The sub-category IDs which correspond to antidepressants are 93 (Monoamine Oxidase Inhibitors - MAOIs), 94 (Selective serotonin reuptake inhibitors - SSRIs) and 102 (Antidepressants - other). Since drugs of abuse contains a wide range of drugs, the sub-category IDs for these are shown in Figure 3.7.

SubCategoryID	SubCategoryName	CategoryID
19	SCRA Drugs of abuse	19
20	Amfetamines and similar stimulants Drugs of abu	19
21	Piperazines Drugs of abuse	19
22	Synthetic Opioids Drugs of abuse	19
23	Tryptamines Drugs of abuse	19
24	Benzodiazepines Drugs of abuse	19
25	Synthetic Cocaines Drugs of abuse	19
26	Unknown drug of abuse	19
27	Dissociative drugs of abuse	19
28	Cannabis (and synonyms)	19
29	Lysergic acid diethylamide and related agents	19
30	Quinazoline derivatives	19
31	Amfetamine (and synonyms)	19
32	Methamfetamine	19
33	MDMA (and synonyms)	19
34	GHB and analogues	19
35	Organic nitrites (and synonyms)	19
36	Heroin (and synonyms)	19
37	Cocaine (and synonyms)	19
38	LSD (and synonyms)	19
39	Alpha-adrenoceptor blocking drugs Drugs of abus	19
40	Butylone (and synonyms)	19
41	Mephedrone (and synonyms)	19
42	Methedrone (and synonyms)	19
43	Methylone (and synonyms)	19
44	Methcathinone (and synonyms)	19
45	PMA (and synonyms)	19
46	AMT (and synonyms)	19
47	Khat (and synonyms)	19
91	Misc Drugs of abuse	19
92	Substances of abuse	7

Figure 3.7: Image showing the sub-categories included in the Drugs of Abuse analysis

The code to extract antidepressants and drugs of abuse related accesses was similar to that used to extract the agrochemicals related accesses. The only change was that the line `t.CategoryID='17'` became `t.SubCategoryID IN IDvec`, where `IDvec` is used to denote a vector containing the sub-category IDs for antidepressants or drugs of abuse as relevant.

The SQL code presented corresponds to one year only. These queries were repeated for each year included in the analysis (2008-2015) and the results from each of these were combined to form the full data set.

3.2 Methods

The models used in this chapter are Generalised Additive Models [59], as in Chapter 2, which are implemented in the `mgcv` [60, 61, 62, 59] package within the `R` statistical software package [63]. The goal with this piece of work was to examine the underlying trends in accesses, therefore only temporal components were examined.

The data were initially aggregated to give the overall number of accesses on a given day across Britain. A model was fit to this series which would extract a day of the week component of trend alongside seasonal and long term trend components. This model can be seen in Equation 3.1

$$E(\log(y(t))) = \log(n(t)) + \beta_1 + \beta_{day(t)} + f(d(t)) + g(t) \tag{3.1}$$

In this equation, $y(t)$ is the number of TOXBASE accesses on day t, while $n(t)$ is the number of registered TOXBASE users on day t, meaning that the output from this model corresponds to the number of accesses made per user. The term β_1 represents the average number of accesses made on a Monday. The subscript $day(t)$ represents a given day of the week, which means that $\beta_{day(t)}$ gives the difference between the average number of accesses on Monday and the average number of accesses on that day of the week. The term $f(d(t))$ represents the seasonal trend in the data, where $d(t)$ is a given day within a year. In the equation, t is an overall time indicator, so that $g(t)$ represents the long term trend. The data were assessed for overdispersion as in Chapter 2. Where appropriate, a negative binomial model was assumed using the log link as shown in Equation 3.1.

As in the previous chapter, residual temporal autocorrelation was assessed using the autocorrelation function and the partial autocorrelation function. ARMA correlation structures were then incorporated into the model as necessary.

Regional level models were then constructed. It was recognised that there were various ways in which the trends could vary by region. Therefore, multiple models were fit, each reflecting a different way in which TOXBASE accesses may vary by location. The simplest scenario is that all regions access TOXBASE in the same way, with no regional effect, other than accounting for the population in each region, as seen in Equation 3.2.

$$E(\log(y_l(t))) = \log(x_l) + \beta_1 + \beta_{day(t)} + f(d(t)) + g(t) \tag{3.2}$$

All of the terms here are as before, the only addition is that of $\log(x_l)$ which is the log of the population in region l. This acts as an offset and has been scaled such that the model predictions are rates of access per 1,000 members of the population based in the 2015 mid-year population estimates [71]. The most complex model in the selection process indicated that all of the trends varied by region, as shown in Equation 3.3.

$$E(\log(y_l(t))) = \log(x_l(t)) + \beta_1 + \beta_{day(t)} + \gamma_l + (\beta\gamma)_{day(t)l} + f_l(d(t)) + g_l(t) \tag{3.3}$$

Here the interaction of $\beta_{day(t)}$ and γ_l indicates a within week effect which varies by region and the subscript l on each of the smooth terms indicates that these differ by region. Spatio-temporal autocorrelation was accounted for by incorporating regional random effects in a Markov Random Field within the additive model as in the spatio-temporal analysis outlined in Chapter 2, with a temporal autocorrelation structure incorporated where necessary.

This chapter additionally assesses the location of any stationary points in the smooth functions. This allows for more precise estimates of where the peaks in seasonality lie, and also means that levelling off in the long term trend can be assessed.

For a known functional form, this would be done by examining the first derivative, which is the gradient of the function. The smooth terms from generalised additive models, however, do not have a known functional form, meaning that obtaining the first derivative is not straightforward. In order to obtain estimates of the derivatives, the finite differences method was used. This uses the fact that the first derivative of a function $f(a)$ can be expressed as

$$f'(a) = \lim_{h \to 0} \frac{f(a+h) - f(a)}{h}$$

and evaluates the fraction in the above, for some small fixed value of h across a variety of points a on the curve. Each of these evaluations is then approximately equal to the first derivative of the smooth function at point a. In this case, the smallest h possible was a difference of 1 day, due to the presence of the factor for day of the week within the model.

Within `mgcv`, these derivatives were found by taking predictions from the model, using the `lpmatrix` structure. This produces a matrix $\boldsymbol{X}$ such that:

$$\hat{\boldsymbol{y}} = \boldsymbol{X}\boldsymbol{p}$$

where $\boldsymbol{p}$ is the vector of parameters for the smooths and $\hat{\boldsymbol{y}}$ is the vector of fitted values. If $\boldsymbol{X_1}$ and $\boldsymbol{X_0}$ are matrices of this type evaluated at $a + h$ and a respectively, then a matrix $\boldsymbol{X_p}$ can be found such that

$$\boldsymbol{X_p} = \frac{\boldsymbol{X_1} - \boldsymbol{X_0}}{h}$$

The derivative for a specific smooth can then be found by post-multiplying the matrix made up of columns of $\boldsymbol{X_p}$ corresponding to that smooth by a vector of the coefficients corresponding to that smooth.

Getting a measure of uncertainty for the derivatives was also of interest, as this would allow for the estimation of a range of plausible values for the stationary points of the smooths. This was done by constructing a matrix $\boldsymbol{X_i}$ which is equal to $\boldsymbol{X_p}$ in the columns corresponding to the smooth of interest and zero otherwise. The standard errors are then the row sums of $\boldsymbol{X_i V_p} \odot \boldsymbol{X_i}$, where $\odot$ denotes componentwise matrix multiplication and $\boldsymbol{V_p}$ is the $p \times p$ covariance matrix for the model parameters. These standard errors can then be used to construct confidence intervals [72, 73].

3.3 Results

3.3.1 Overall Analysis

The data used for this stage of the analysis were aggregated across Great Britain, to produce the data shown in Table 3.1.

Table 3.1: Table showing the structure of the data used in the analysis over Great Britain

Accesses	Day	Month	Year	Hospitals
1414	Tuesday	1.032	2008.003	371
1035	Wednesday	1.065	2008.005	371
1033	Thursday	1.097	2008.008	371
1072	Friday	1.129	2008.011	371
...	...	...	...	...
1789	Monday	12.903	2015.992	395
1864	Tuesday	12.935	2015.995	395
1896	Wednesday	12.968	2015.997	395
1839	Thursday	13.000	2016.000	395

Since January 1st 2008 was a Tuesday, the day of the week variable was constructed by incrementing the day of the week from then. The month of year was obtained by taking the month component from the date variable and adding the day component divided by the number of days in that month. As an example, the month values for January were 1 + Day/31. The year variable was defined in a similar way, where the year was taken to be the observed year plus the number of days into that year divided by the number of days in that year. For 2008, this would be 2008 + DayOfYear/366, since 2008 was a leap year. 'Hospitals' contains information on the number of registered TOXBASE users on that date, in order to ensure that any changes in the long term were not due to a large increase in the number of users.

There were a total of 436,060 TOXBASE accesses made by emergency departments in Great Britain in 2008, which had increased substantially to 733,363 accesses in 2015. In 2008 there were 371 registered users of TOXBASE who were categorised as emergency departments, and there were an average of 3.21 accesses per user per day. The number of emergency department users registered to TOXBASE had increased to 395 by 2015, with an increase in average use to 5.09 accesses per registered user per day.

The raw time series of the number of accesses per registered user per day is shown in Figure 3.8. There is a clear increasing trend between 2008 and mid-2010, there is then a slight decrease. This is followed by an increase again from 2011 up until 2014, with the series seeming to level off over the final two years of data. It is also worth noting that there is a dip in accesses each year which, with a couple of exceptions, occurs in mid to late December, around Christmas.

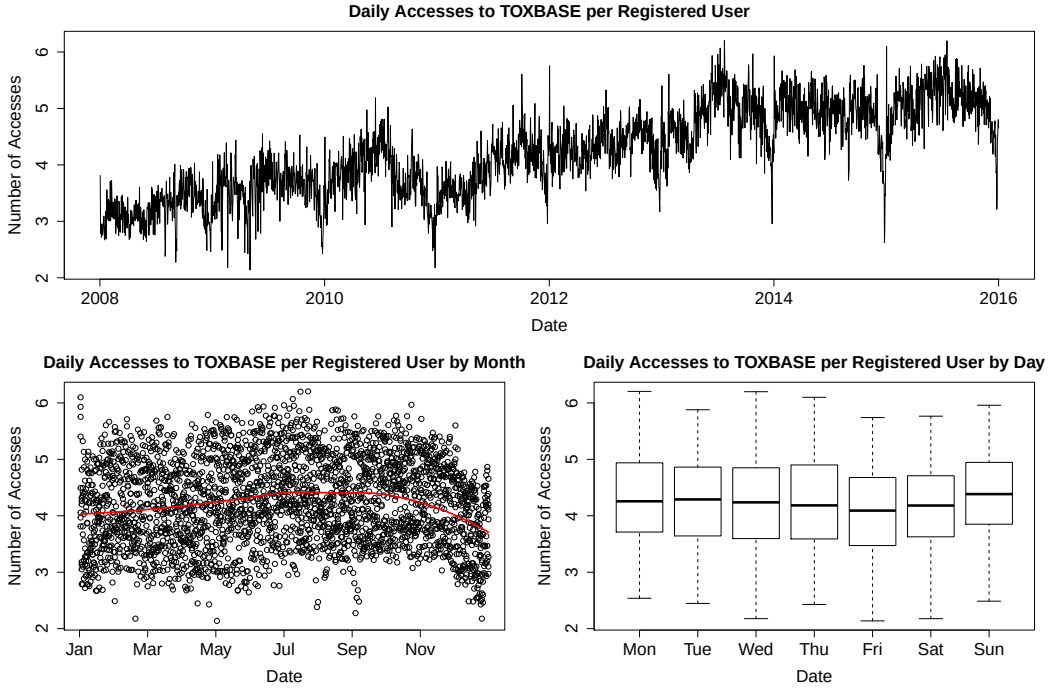


Figure 3.8: Daily number of accesses per registered user of TOXBASE plotted against year (top), month (bottom left) and Day (bottom right)

A LOESS curve has been added to the plot of accesses by month shown in the bottom left plot in Figure 3.8 in order to aid interpretation. From this, there appears to be a slight seasonal trend, which peaks in late summer. The dip in accesses in December can also be seen here.

The day of the week trend (bottom right Figure 3.8) does not appear to be particularly strong, though there is a small tendency towards higher accesses on Sundays in comparison to the rest of the week, as well as a tendency for accesses to be lower on a Friday compared to the rest of the week.

The model was initially fit assuming that the data come from a poisson distribution. This initial model had a dispersion factor of 17981.61, indicating that overdispersion was a substantial problem with these data. A negative binomial model was then implemented, resulting in a dispersion factor of 1.006, indicating correct dispersion.

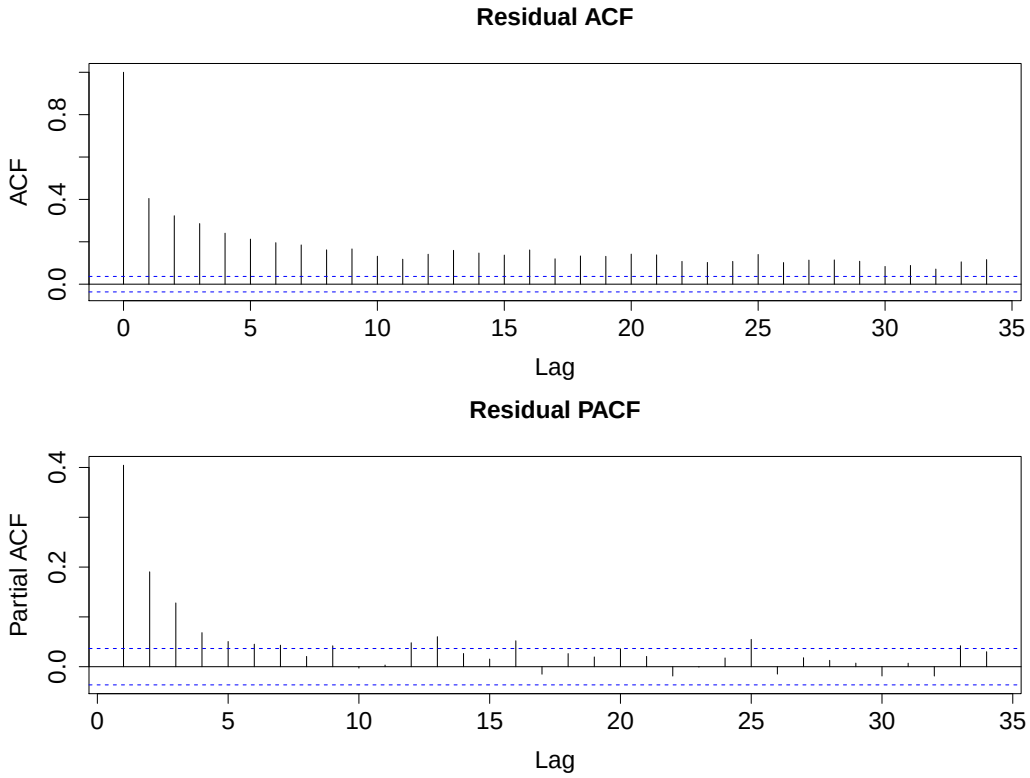


Figure 3.9: The autocorrelation function (top) and the partial autocorrelation function (bottom) for the initial temporal model fit to TOXBASE accesses

The ACF and PACF from the negative binomial model are shown in Figure 3.9. This shows that there is significant autocorrelation up to a high number of lags in the ACF, with seven significant lags in the PACF. This indicates that an AR(7) process may be most appropriate for this data, which seems rather high. Since lags 5, 6 and 7 show autocorrelation that is only slightly outside of the lines, an AR(3) process may also be plausible.

Correlation structures ranging from AR(1) up to AR(7) were fit and compared using AIC, in order to determine which would be most appropriate. In order to select the most appropriate correlation model, AIC and BIC values were calculated for each model (Table 3.2). In order to balance fit and complexity, the point at which these metrics are no longer decreasing by a large amount was found, and the corresponding correlation structure used. This shows that the AIC

begins to level off at around AR(4), and BIC begins to level off around AR(3). This suggests that there is limited benefit to incorporating any lag greater than 4. The results from the model containing an AR(4) autocorrelation structure will be presented here.

Table 3.2: AIC for each of the correlation structures examined

Model	AIC	BIC
AR(1)	-6438.32	-6372.61
AR(2)	-6568.35	-6496.76
AR(3)	-6623.85	-6546.29
AR(4)	-6642.32	-6558.80
AR(5)	-6650.91	-6561.42
AR(6)	-6657.38	-6561.92
AR(7)	-6662.96	-6561.54

Figure 3.10 shows the autocorrelation and partial autocorrelation function for the residuals of the model incorporating an AR(4) autocorrelation structure. This shows that the exponential decrease in the ACF is no longer present and the first few lags of the PACF are non-significant, which indicates that the temporal correlation structure used was sufficient in capturing the residual autocorrelation structure.

The model estimated trends are depicted in Figure 3.11. The long term trend (right) shows that accesses increased fairly consistently between 2008 and 2015. The increasing trend slowed in 2010, before increasing more rapidly from 2012 until 2014. In total it was estimated that accesses had increased by 66.7% (95% CI: 56.1%, 78.0%) over the 8 year period. The first derivative of this smooth function was examined to determine whether the long term trend does, in fact, level off. This statistic indicated that, while access to TOXBASE appears to have been increasing, it is plausible that the trend in accesses started to level off in October 2013, which is where zero first appears in the confidence interval for the first derivative.

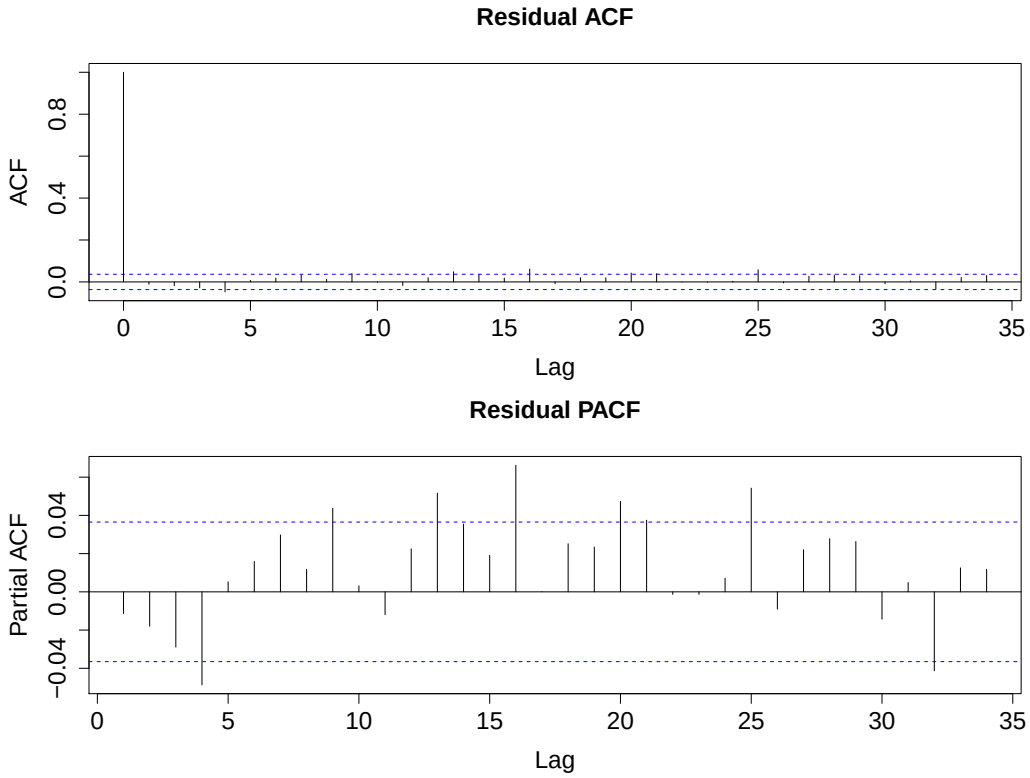


Figure 3.10: The autocorrelation function (top) and the partial autocorrelation function (bottom) for the temporal model fit incorporating an AR(4) correlation structure

The seasonal trend (left) shows three distinct peaks in accesses throughout the year. The first peak represents an increase of 13.8% (95% CI: 9.7%, 18.2%) between January and mid to late February. From this peak, there is a 6% (95% CI: 2.2%, 9.9%) decrease to a minimum in late March or early April, which is followed by an increase of 9.2% (95% CI: 5.5%, 12.9%) to a peak in mid-July. There was then a decrease of 5.3% (95% CI: 1.6%, 9.0%). There was then a small but not statistically significant increase of 3.4% (95% CI: -0.4%, 7.4%) between August and October. There is a large decrease from the October peak to the minimum around the Christmas period of 15.9% (95% CI: 11.7%, 20.2%). It is also worth noting that this trend is similar to that seen in the NHS24 calls data in Chapter 2.

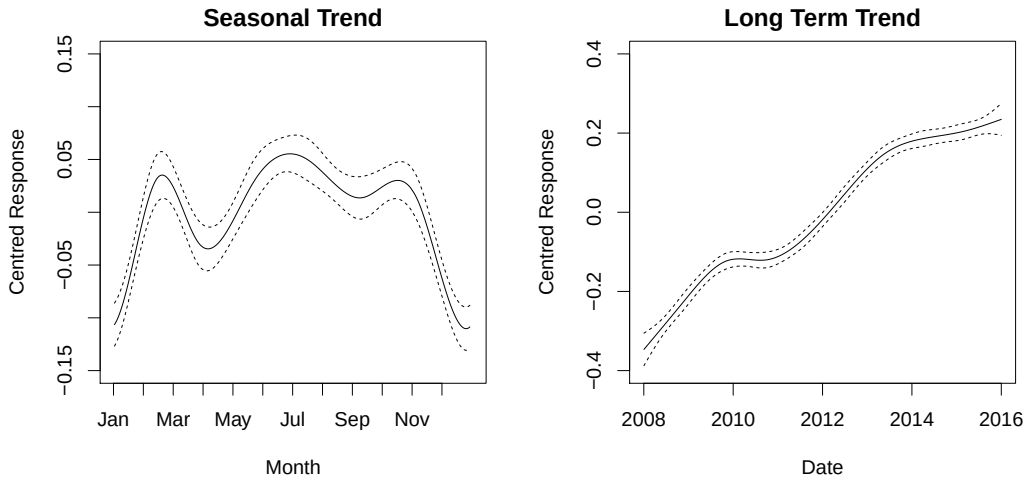


Figure 3.11: Seasonal(left) and long term (right) trends in access to the TOXBASE database, as estimated from a generalised additive model (dashed lines represent a 95% confidence interval

The seasonal pattern was investigated further by considering the seasonality in accesses to specific toxins. The seasonal trends in accesses to antidepressants, agrochemicals and drugs of abuse pages were examined (Figure 3.12), as recommended by colleagues at NPIS. Antidepressants are the most commonly accessed drug group, accounting for around 11% of all accesses, while drugs of abuse and agrochemicals represent around 6% and 1.5% of accesses respectively.

It can be seen that accesses to both agrochemicals and, less prominently, drugs of abuse pages peak around July, which is approximately the same time as the second peak in the overall seasonal trend. What is more interesting, however, is that accesses to antidepressants pages show a very similar seasonal pattern to the trend in overall accesses, although with a greater peak in February.

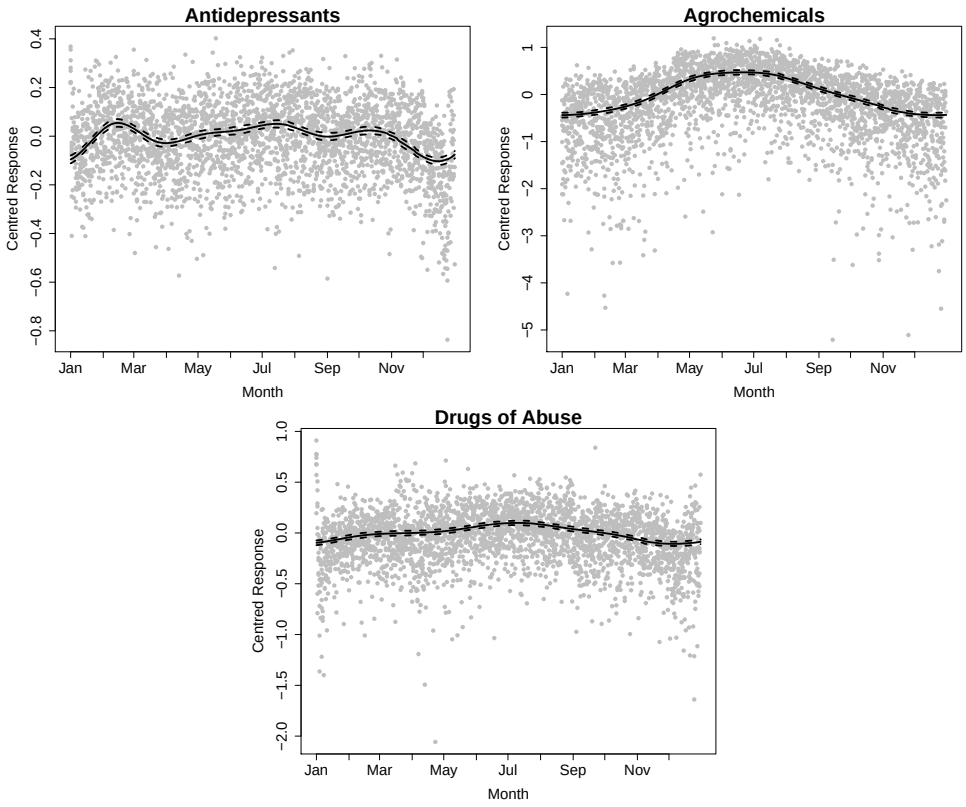


Figure 3.12: Seasonality in the number of accesses made to pages related to antidepressants (top left), agrochemicals (top right) and drugs of abuse (bottom)

The average number of accesses per registered user on each day of the week is shown in Figure 3.13. Here "day" runs from midnight to midnight, meaning that accesses made due to an overdose on Friday night may actually contribute to Saturday's accesses. From this, on average, accesses are highest on Sundays, with 4.35 (95% CI: 4.20, 4.31). Accesses then decrease during the working week by 4.6% (95% CI: 4.1%, 6.1%) on average. Accesses then increase between Friday and Sunday by 7.5% (95% CI: 6.1%, 8.9%).

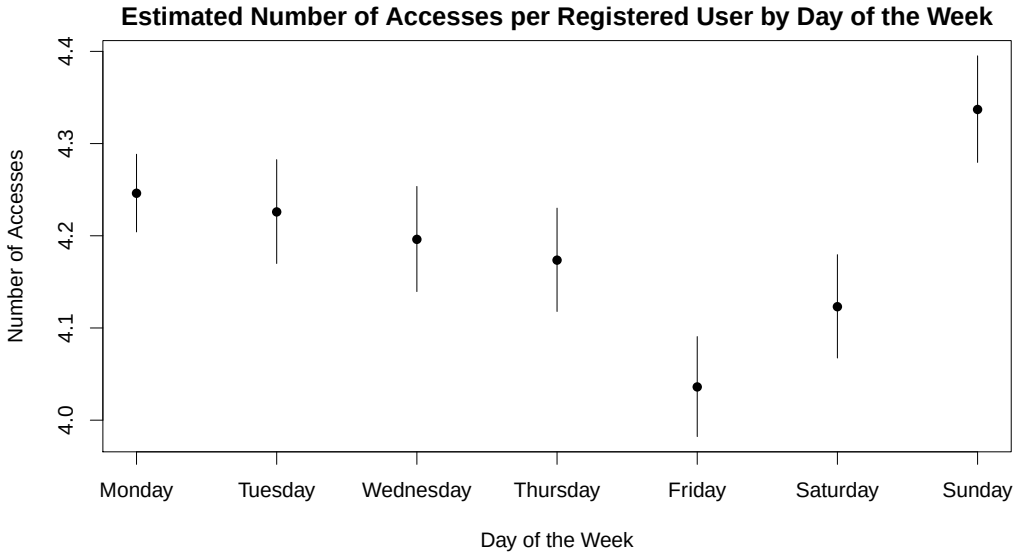


Figure 3.13: Average number of accesses for each day of the week (vertical lines are 95% confidence intervals)

3.3.2 Regional Analysis

The data used in this part of the analysis were separated into 11 regions: East Midlands, East of England, London, North East, North West, Scotland, South East, South West, Wales, West Midlands and Yorkshire & the Humber. The data set contained daily data for each region, with a total of 31,713 observations. The first and last few rows are shown in Table 3.3.

Table 3.3: Table showing the structure of the data used in the analysis over Great Britain

Region	Date	Accesses	Population
east midlands	2008-01-01	121	4677038
east of england	2008-01-01	94	6076451
london	2008-01-01	132	8673713
north east	2008-01-01	77	2624621
...	...	...	...
south west	2015-12-31	128	5471180
wales	2015-12-31	111	3099086
west midlands	2015-12-31	227	5751000
yorkshire and the humber	2015-12-31	197	5390576

Small changes in the population are going to have a miniscule impact on the rate of accesses due to the relative size of accesses compared to the population. This meant that the populations data could be treated as fixed, and were taken to be the mid-2015 population estimates obtained from the Office for National Statistics website (`www.ons.gov.uk`).

The time series of TOXBASE accesses for each region are shown in Figure 3.14. This indicates that the overall trends in each region are fairly similar. There are some slight differences however. For example there appears to be a more pronounced dip in accesses in 2011 in some regions, such as the North West, in comparison to other. There are also some trends which look more consistent. For example accesses in both Wales and Scotland appear to have been increasing linearly with time.

In order to be consistent, models were fit using a negative binomial model, rather than a poisson distribution to account for overdispersion. The AIC and BIC of each of the models under consideration is shown in Table 3.4.

Table 3.4: Table showing the AIC and BIC of each of the models fit

	AIC	**BIC**
All regions are the same	320,262	320,470
Seasonal trend differs by region	320,309	321,143
Long term trends differs by region	319,861	320,660
Seasonal & long term trends differ by region	319,912	321,338
Rate of accesses...		
...differs by region	303,138	303,430
...& seasonal trend differ by region	303,089	304,026
...& long term trend differ by region	302,282	303,228
...&seasonal & long term trend differ by region	302,233	303,829
Rate of accesses & day of the week effect...		
...differ by region	303,014	303,808
...& seasonal trend differ by region	302,963	304,402
...& long term trend differ by region	302,152	303,600
All regions are different	302,102	304,200

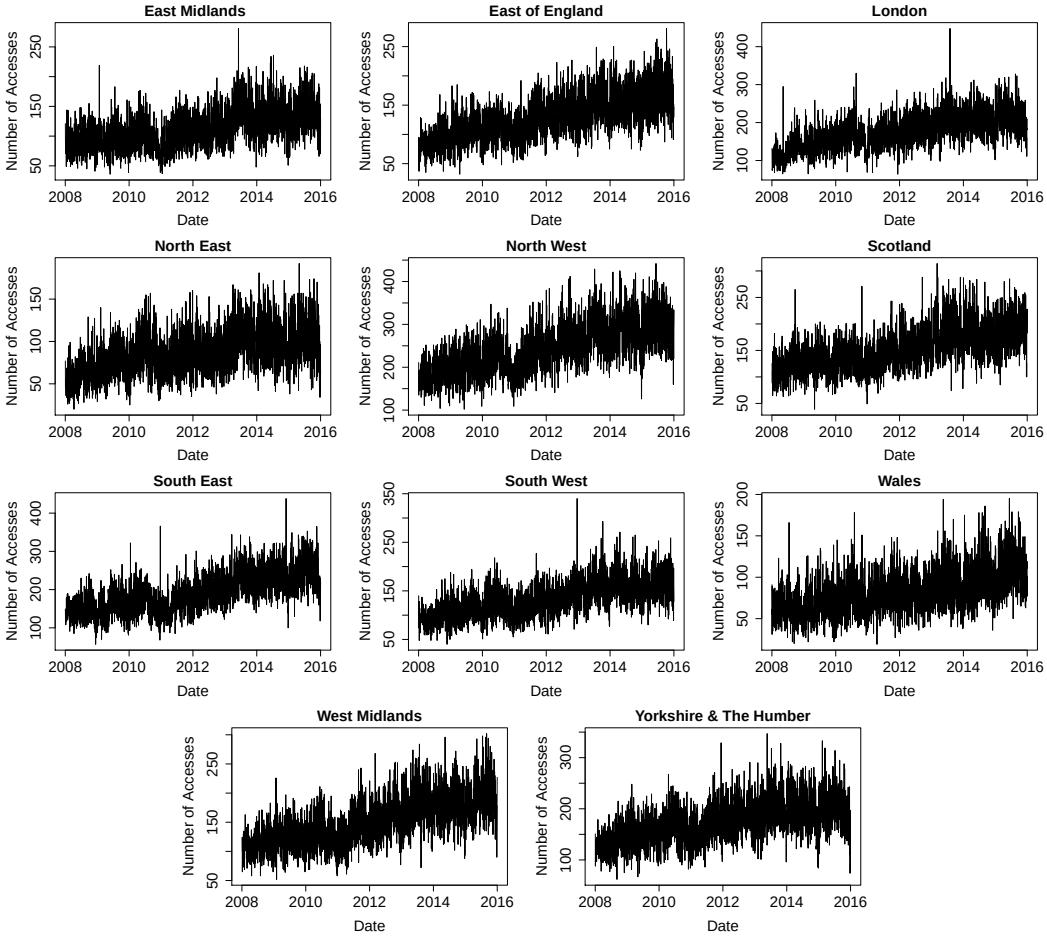


Figure 3.14: Plots showing the temporal trends in each of the regions

There were two models which appeared to be similar under these criteria: the model with differences in the rate of accesses by region only (AIC: 303,138 & BIC: 303,430), and the model which differed in terms of the rate of access and in the long term trend (AIC: 302,282 & BIC: 303,829). The long term trends from the more complex of these models were plotted in order to see which regions differed, and by how much (Figure 3.15).

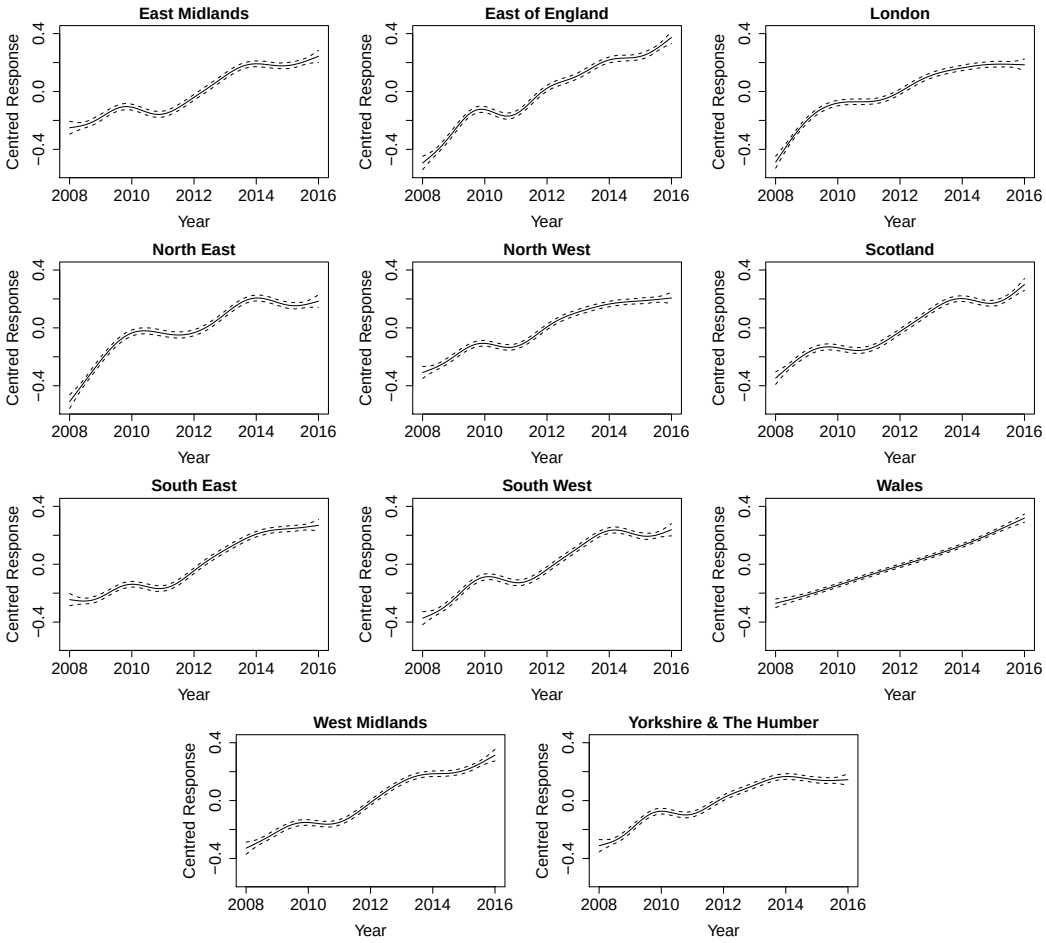


Figure 3.15: Estimated long term trends for each region

From Figure 3.15, it can be seen that the majority of the trends shown have similar features: An increase from 2008 until around 2010, where accesses levelled off, followed by another period of increase and another period of levelling off. The only region which does not seem to follow this trend is Wales, where it appears that access to TOXBASE has been consistently increasing since 2008.

The model was not significantly improved by accounting for this difference in Wales, a likelihood ratio test comparing the two models returned a p-value of 0.199. Therefore results will be presented for the simpler model, the form of which is shown below.

$$E(log(y_l(t))) = log(x_l(t)) + \beta_1 + \beta_{day(t)} + \gamma_l + f(d(t)) + g(t) \qquad (3.4)$$

In this equation, γ_l facilitates each region having different rates of access, there is no interaction between any of the other terms and region.

For these data, the spatio-temporal adjacency matrix would be 31,713 $\times$ 31,713, which is too large for `R` to store and use, making it impossible to measure the spatio-temporal autocorrelation in the data. In order to best assess whether correlation structure was appropriate, a spatio-temporal structure was incorporated, and the AIC and BIC values of this model were compared to that of the model without spatio-temporal structure. For this model the AIC and BIC values were 302,731 and 303,150, which are lower than those for the model carried forward (303,138 and 303,430). Interpretation was then made using the model which incorporates the spatio-temporal structure in the data.

The estimated trends were similar to those described previously, with the rate of access in each region increasing as seen in Figure 3.11. In this instance, however, the first derivative indicated no evidence of levelling off, with the 95% confidence interval only slightly above zero for across the final two years of data.

The seasonal trend was also similar to that in Figure 3.11, once again showing three peaks in accesses to TOXBASE in February, July and October. In addition, the day of the week effect shows a peak in accesses on Sundays, with accesses generally decreasing during the week and increasing over the weekend, as in Figure 3.13.

Table 3.5 shows how the rates varied across Great Britain, with reference to the West Midlands, which had the median rate of accesses over the whole study period. This corresponds to γ_l in Equation 3.4.

Table 3.5: Relative rate of accesses made to TOXBASE for each region compared to the West Midlands region. These values also account for the number of hospitals in each region at a given time point

Region	Estimate
North East	1.3281 (1.2478, 1.4136)
Yorkshire & The Humber	1.2227 (1.1590, 1.2898)
North West	1.1050 (0.9824, 1.2428)
East Midlands	1.0216 (0.9375, 1.1133)
Wales	1.0076 (0.9547, 1.0635)
West Midlands	1.0000
Scotland	0.8807 (0.7592, 1.0215)
East of England	0.8429 (0.7944, 0.8943)
South West	0.8292 (0.7555, 0.9101)
South East	0.6863 (0.6075, 0.7753)
London	0.6631 (0.5919, 0.7429)

This shows that regions in the north of the country seem to be the most prolific users of TOXBASE with respect to the population in that region, with the North East and Yorkshire & the Humber being the biggest users of the system. The North West, East Midlands and Wales were all found to use the system at a higher rate than the West Midlands on average, although the 95% confidence intervals presented indicate that the difference in rates is not statistically significant.

Scotland was found to use the system at a lower rate than the West Midlands, though this result was not significant at the 5% level. Those regions in the south of England use the system less than average, with London being the smallest user of the system.

3.4 Discussion

This piece of work highlights that there are consistent trends in the usage of the TOXBASE database. One of the main features is that TOXBASE accesses have been on the rise since 2008, and have in fact been increasing since the online tool TOXBASE was indicated as the first port of call in cases of poisoning in 2005

[74]. One might attempt to make a causal link here, however there is no evidence to suggest that cases of poisoning have been increasing over the same period. In fact drug related deaths were decreasing in England and Wales until 2012 [23]. In contrast, drug related death in Scotland does appear to be increasing, although not to the same extent as the estimated trends in Figure 3.11 would suggest [22].

There is also no evidence to suggest that poisoning related admissions have increased at as high a rate as has been shown in access to TOXBASE [24, 75]. However, this does not account for presentations at emergency departments. These may be increasing in a way that is not reflected in the admissions data, perhaps due to better management leading to a lower proportion of admissions, a hypothesis that will be assessed in Chapter 7.

It is however of note that the second period of increase occurs in 2012, which is the same year that the NICE guidelines on the management of poisoned patients was published, recommending TOXBASE as a key source of information [45]. This is also the same year that new psychoactive substances were beginning to emerge [33].

The estimated seasonal trends indicate that there are three peaks in TOXBASE accesses, which seem to be consistent across location. These occur in February, July and October. The minima in accesses occur in between these, with an overall minimum around Christmas time. A "holiday effect" has been documented in a previous study on self-harm attempts using data collected from across Europe by the World Health Organisation [76]. This study found that there was evidence of a reduced number of suicides around holidays, most notably those of Christian origin. The peak in February may therefore be due to the fact that it lies between two of the main Christian holidays; Christmas and Easter. It may also be, in part, related to Seasonal Affective Disorder (SAD) which causes symptoms of depression around mid-winter [77].

The small increase in October is more than 2 months after the official start date for junior doctors, and therefore is unlikely to be caused by training. This peak does however occur slightly before daylight savings in the UK. It is also within a few weeks of the new university term. This effect may therefore be related to the "broken promises" principle outlined by Gabennesch, which suggests that vulnerable individuals may be strongly affected by disappointment after a new beginning [78]. There may also be some impact of the UK school holidays, which have a half-term break in February and October, in additon to the summer break.

Further, it is interesting that the overall seasonal pattern was found to be similar to that of antidepressants, which is suggestive of a link between access to TOXBASE and self-harm. The peak in summer was slightly more pronounced than that in the antidepressant series, and it is likely that there are other contributing factors. This summer effect has been seen by poisons centres in Sweden, where the suggestion is that increased outdoor activity in addition to the wide variety of plants, animals and agrochemicals involved leads to this increase [79]. The same can also be seen in records of calls to US poisons centres, this can be seen in Figure 3 of their annual report [67], although the seasonality in these calls was not formally assessed. This summer peak has also been seen in a study on seasonality in illicit drug overdose, and was found in both accidental and intentional overdose [80]. The authors attribute this to increased alcohol consumption in mid-summer, however this may also be related to increased drug use during festivals [69].

Seasonality in suicides and attempted suicides, of which poisoning plays a large part, has been previously researched. The results of these studies tends to vary, with different studies indicating different seasonal variation. In the studies found, the patterns were different to those found in this analysis. A literature review specifically centred around suicides by drowning found that there were

peaks in activity in spring and early autumn [81], a fact that was established in two English studies and one in the US [82, 83, 84]. An Irish study on suicide, however, tells a slightly different story. A peak in suicide was still found in spring, however rather than a high in activity in autumn, there was found to be a low [85].

A study carried out in Australia made a distinction between violent methods of self-harm and non-violent methods, the latter of which would include self-poisoning. This study found that there was seasonality in violent methods, but no evidence of a seasonal effect on non-violent methods, which may be in contrast to the results presented in this analysis [86]. There was, however, a study conducted in Norway, specifically discussing non-fatal opioid overdoses, found that there was a seasonal pattern in these cases which approximately corresponds with the seasonal trend found in this analysis, with low incidence rates in April and a peak in activity in August [87].

As with the seasonal pattern, the day of the week effect was found to be consistent across each of the regions examined. A study using the National Self-Harm Registry of Ireland found a similar day of the week effect in self-harm incidents [88]. In addition, a study on attempted suicides in Helsinki found that suicide attempts tend to cluster around the weekend [89]. This weekend effect may have a link to increased recreational drug use over the weekend [90] and it has been previously found that there is a greater number of poisoned patients seen at emergency departments between Friday evening and Tuesday morning [91]. The Irish National Poisons Information Centre noted that calls to their service were higher earlier in the week than at the weekend [92], contrasting the analysis described here.

This work has indicated that there are consistent trends in accesses to the TOXBASE database, with accesses higher at the weekend and at three periods through the year. That there are similar trends seen in some of the literature, as well as in the NHS 24 analysis in Chapter 2, may indicate a trend in cases of poisoning. The implications of this work in terms of the impact of TOXBASE on hospital admissions is not yet clear however, and this work will inform the linkage and analysis of the TOXBASE data and hospital admissions data in Chapter 7.

Chapter 4

Missing Data Methods

In the following chapters, routinely collected healthcare data will be linked with the TOXBASE data at hospital level. These data were obtained from three sources: the Information Services Division of NHS Scotland (ISD), the Health and Social Care Information Centre (HSCIC) in England and the NHS Wales Informatics Service. With the exception of the NHS Wales Informatics Service, it is the policy of these organisations to suppress small cell counts, in order to retain privacy. The policy in ISD is to suppress cell counts of 1, 2, 3 or 4, while HSCIC additionally suppress cell counts of 5. This issue will be described in more detail in Section 4.1.

In order for these analyses to be viable, it is important that the suppressed data are imputed in an appropriate way, in order to avoid bias in the results. Section 4.2 will provide an overview of some current methods, with focus on those methods being used as comparators for a potential solution for the scenario. The intuition and methodology behind the proposed solution for using these data are outlined in Section 4.3. A simulation study was conducted in order to compare some commonly used methods with the proposed solution, and the results of this are presented in Section 4.4.

4.1 Description of Problem

The specific problem which is the focus of this chapter is data which are missing not at random, given that only values within a particular range are missing.

The data in question were obtained from the Information Services Division of the NHS (ISD). These data concern monthly emergency admissions to hospital for cases of poisoning for 46 hospitals across Scotland between January 2008 and December 2015. The protocol used by ISD in giving out data requires the suppression of cells which contain values lower than or equal to 4, but with values of zero recorded. As poisoning is relatively rare, a large proportion of the data set was censored. Specifically, 15.6% of the data were censored, with smaller hospitals generally requiring more censoring. Out of the 49 hospitals, 22 had no missing data. These 22 hospitals are either hospitals which are so small as to have had no emergency admissions due to poisoning, of which there were three, or hospitals which are so large as to have always had more 5 or more admissions in a given month. Of the remaining hospitals, the number of missing time points ranges from 1 month out of a possible 96 up to 71 months for which there was missing data.

The ISD protocol is specific in which values are censored, so that it is known that all of the supressed cells must take the value 1, 2, 3 or 4. This additional knowledge allows a model to be built for data which are missing not at random (MNAR), which is potentially simpler than other complex MNAR models such as selection modelling and pattern mixture modelling, which rely on reasonably sophisticated statistical techniques. In contrast, the method outlined can be viewed as an extension to other imputation models which can take into account the known range of the missing values.

Although the ISD protocol is the focus of this study, the results are applicable to the HSCIC data from England, which additionally censor cell counts of 5.

4.2 Missing Data

Missing data in statistical analyses is fairly common, particularly in certain types of analysis. One type of data that commonly has missing values is survey data. This missingness can be caused by a variety of factors: bad question wording, some questions only need to be answered if certain conditions are met or sensitive questions are being asked.

Missingness is also fairly common in repeated measures studies, where participants have not been observed for a specific wave of the study, or where participants drop out of the study before completion. An example of analysis which can handle this type of non-response is survival analysis, which applies when the measurement of interest is a time to event.

In order to define different types of missingness, there are three well documented categories of missing data: missing completely at random (MCAR), missing at random (MAR) and missing not at random (MNAR) [93]. The value of the data that are missing completely at random do not depend on any variables observed or unobserved, and are therefore the easiest case to deal with. Data that are missing at random only depend on variables that have been observed, such as covariates in a regression. In this situation the missingness does not depend on the value that is missing; this is the case that is most commonly handled in the literature. The final type of missingness, missing not at random, occurs when the data that are missing depend on the missing value itself. This type of missingness is more difficult to analyse, and may lead to biased inference if handled incorrectly. This is the type of missingness which is present in the NHS data sets.

There are a variety of methods for dealing with missing data in the literature, although most are only suitable for MCAR or MAR data. A general rule is that the simpler the method, the stricter the assumption that has to be made on the missing data mechanism, with the simplest methods being suitable for only MCAR data. More complex methods, however, can be adapted to cope with the more general MAR or MNAR missingness. Missing data methods generally fall into three categories: naïve methods, imputation based methods and data augmentation. Some discussion will also be given to methods which are applicable in the particular case of censored data.

4.2.1 Naïve Methods

List-wise Deletion

The most straightforward naïve method is list-wise deletion, which can also be referred to as complete case analysis. This technique ignores any data value for which a particular value, or combination of values is missing [93]. Figure 4.1 shows how this would work for a data set with three variables. One of the issues with this method is that, where a lot of data are missing, the sample size can be substantially diminished.

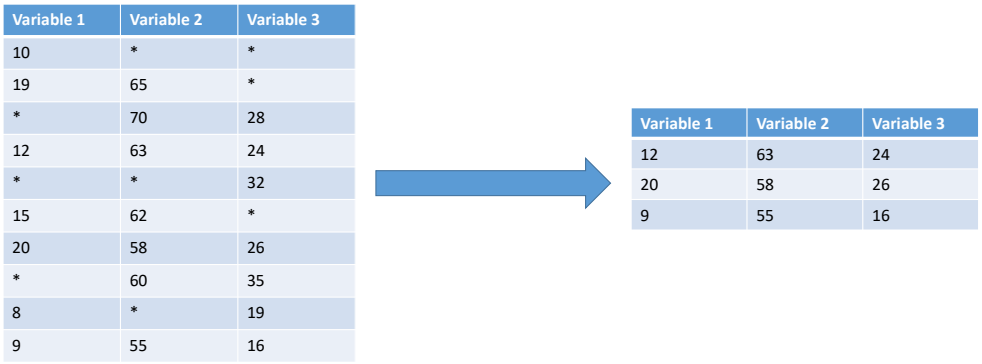


Figure 4.1: Image illustrating listwise deletion. Each row which contains at least one missing value is removed from the data

However, the main issue with this method is that it makes the assumption that the missing data are random draws from the same population as the observed data. As such this method only produces unbiased estimates and accurate standard errors in situations where the data are missing completely at random.

A related method of handling missing data is pairwise deletion. This is used to calculate pairwise summaries within the data set using all possible information [93]. Figure 4.2 demonstrates this. Each pair of variables is taken from the original data set, and listwise deletion is carried out on the pairwise data, producing the data sets shown on the right of Figure 4.2. Then, pairwise summaries can be calculated on each of these subsets of the data.

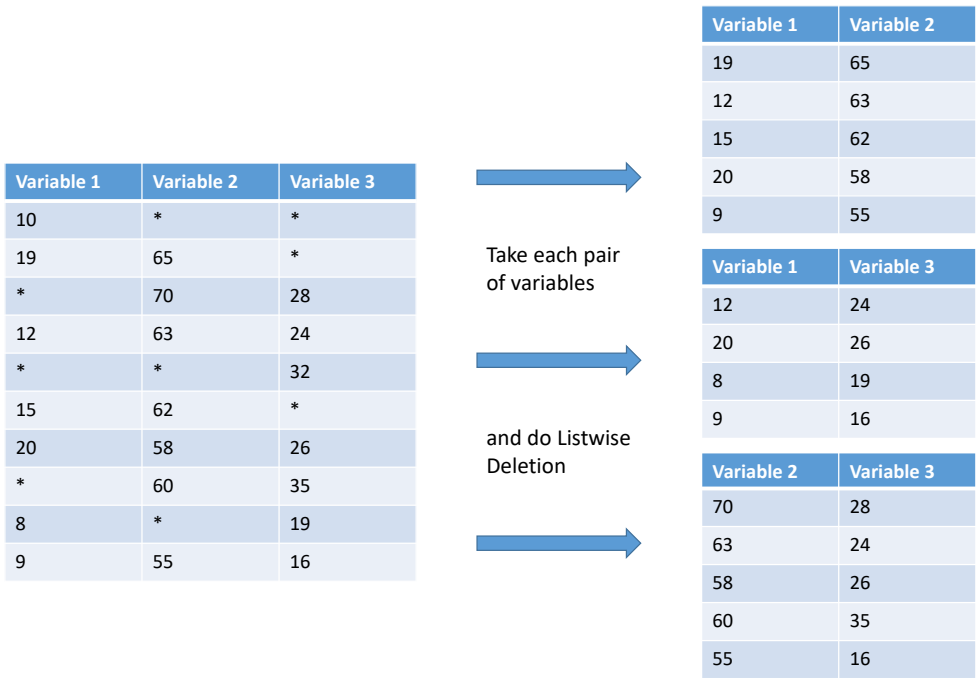


Figure 4.2: Image illustrating pairwise deletion. Each pair of variables is taken, and the observations with missing data in either of these are removed. Pairwise summaries can be calculated on each individual pair.

The logic behind this method is that the data being used for each calculation is maximised. However, each element of the covariance matrix is estimated with different subsets of the data, which can lead to the estimated matrix being non-positive definite.

4.2.2 Imputation Methods

Other common missing data methods come under the umbrella term of single imputation. These methods involve substituting each missing observation with a single value, which is usually estimated using values which have been observed.

Average Value Imputation

Average imputation provides a straightforward way of carrying out a single imputation. Using this method, the mean or the modal value is substituted in place of the missing value, meaning that this method has a straightforward implementation, which is illustrated in Figure 4.3.

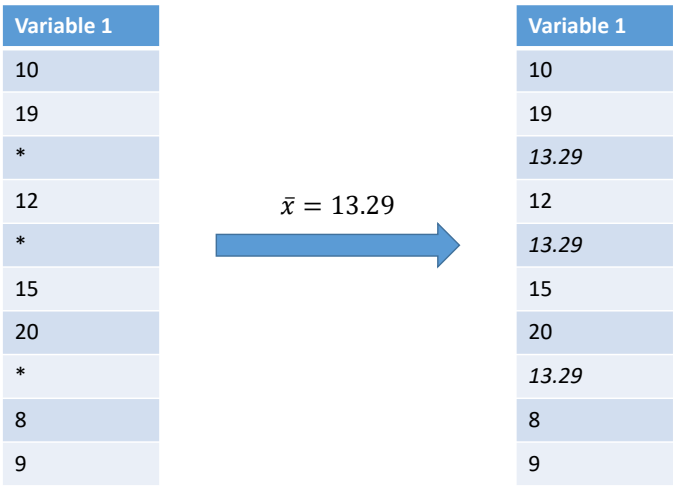


Figure 4.3: Illustration of Average value imputation. The mean of the observed data in the sample is taken and substituted in for the missing values

In order to substitute the average value in for the missing value, this method assumes that the missing data follow the same distribution as the observed data, leading to biased estimates where the data are MNAR. Treating each value as being fixed at the mean also tends to lead to the variability being underestimated [93].

Regression Imputation

Regression imputation extends the idea of average imputation by constructing a model based on the observed values to impute the missing value. This method relies on a model fitting procedure, and as such may be sensitive to model misspecification. As in the previous method, there is still a possibility of bias in the imputed values where the data are missing not at random. It is generally preferred to alter this method so that, rather than using the model prediction as a proxy for the missing value, the value is a drawn from the conditional distribution outlined by the model. An example of this would be fitting a linear regression model of the form:

$$\begin{aligned} y_i &= \beta_0 + \beta_1 x_i + \varepsilon_i \\ \varepsilon_i &\sim N(0, \sigma^2) \end{aligned}$$

The parameters β_0, β_1 and σ^2 would be estimated in the usual way using the complete data. The missing values would then be substituted by values drawn from the $N(\beta_0 + \beta_1 x_i, \sigma^2)$ distribution, where x_i is the observed covariate corresponding to the missing value. Sampling in this way eliminates any covariance distortion that would come from directly using the predicted response. This

method relies on the covariate value x_i being known for all of the missing cases, meaning that this method is unsuitable where there are missing values in both the response and the covariates.

Hot Deck Imputation

Hot deck imputation is another simple method of imputing missing values; it involves replacing each missing value, or recipient, with an observed value which has similar properties, referred to as the donor. There are two types of hot deck imputation; random and deterministic.

Random hot deck imputation is where the donor is randomly selected from a pool of possible donors. Figure 4.4 shows how this might work in practice, with each missing value being replaced by a randomly selected value from the observed data. In a situation with multiple groups, the pool of donors would be restricted to those within the same group.

Deterministic hot deck methods impute the missing values as from its observed nearest neighbour, based on some distance criteria applied to covariates [94]. This method is attractive in a situation where the missing data mechanism is ignorable, i.e. when the data are MCAR or MAR as the missing data are drawn from a pool of observations from the overall sampling distribution. However, for data that are missing not at random, this method is likely to be inappropriate since, by definition, the missing data are not from the same sampling distribution as the observed data [94].

A final single imputation method that is specifically mentioned in literature on longitudinal studies is last observation carried forward; a naïve method which imputes the missing values with the last observation before the missingness occurred. A related method, last observation carried backwards, works in a similar

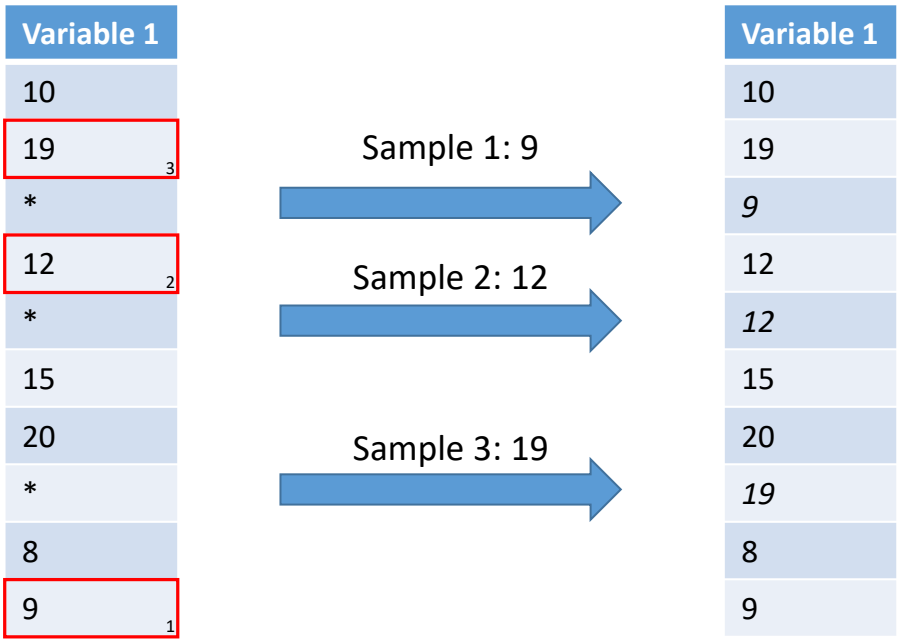


Figure 4.4: Image illustrating hot deck imputation. A sample has been taken for each of the missing values in the data set, and the missing values are replaced by these sampled values

way, but instead imputing missing values with the most recent observation following missingness. This method has been commonly used although some have denounced the method as producing unreliable results [95].

Multiple Imputation

Despite being relatively simple to implement, single imputation methods are somewhat unsatisfactory, in that the uncertainty estimates do not take into account the fact that the imputed values are guesses, and therefore carry additional uncertainty. In an attempt to account for this, the multiple imputation process was developed by Rubin in 1987 [96].

This method is an extension of single imputation from a conditional distribution, where instead of merely generating one value for each of the missing data points, some number ($m > 1$) possible values are generated. In general, m complete data sets are generated from the original data, each of which is analysed using standard complete data methods. As such, each data set produces its own estimate of the parameter of interest, which can be combined using "Rubin's Rules" to form one estimate [96]. A diagram describing the full multiple imputation process is shown in Figure 4.5 for three imputations. This process has been shown to produce unbiased estimates in some MNAR scenarios, it has also been shown to produce unbiased uncertainty estimates for data that are MAR.

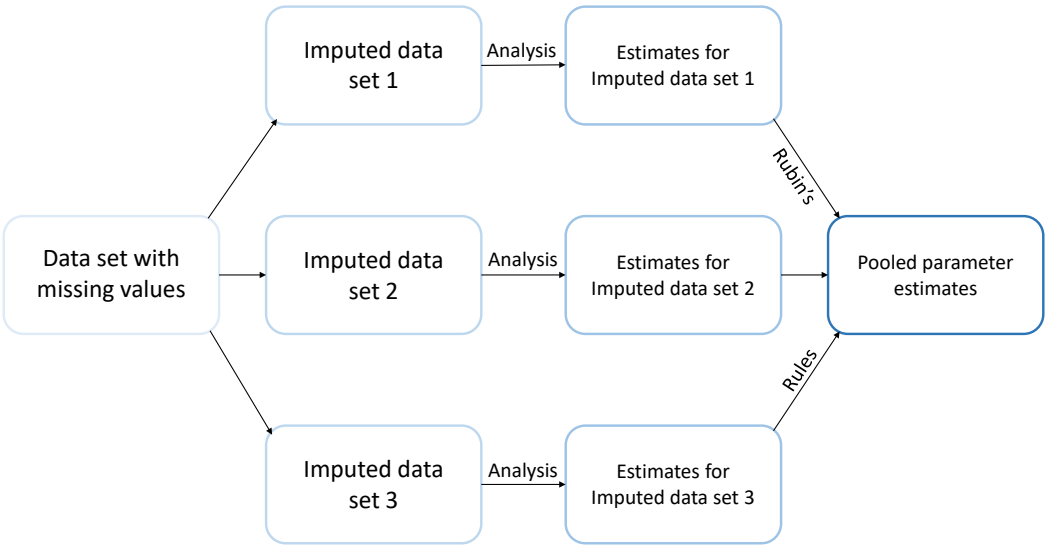


Figure 4.5: Image showing the process of multiple imputation. Some number of data sets is imputed from the original data, in this example three data sets are imputed; in general m data sets are imputed. The appropriate analysis is performed independently on each imputed data set and the results from each of these are combined using Rubin's Rules

4.2.3 Data Augmentation

Data Augmentation is defined as the method of constructing iterative optimization algorithms, which make use of unobserved or latent variables [97]. These methods can be implemented in a frequentist framework, using maximum likelihood, or can be implemented in a Bayesian framework. Due to the generic nature of data augmentation, these procedures can be developed to fit a wide range of scenarios provided that they can be formulated as a problem involving both observed values and latent variables.

Data augmentation via maximum likelihood was famously discussed in an article on the EM algorithm by Dempster, Laird and Rubin [98]. This differs from the imputation methods described previously, in that, instead of trying to fill in the missing data, the missing data are treated as random variables that have not been sampled and must be removed from the likelihood function as nuisance variables. EM algorithms aim to solve difficult incomplete data problems by iteratively solving a simpler complete data problem. Put simply, the missing values are first filled in with a (educated) guess. This complete data set is then used to estimate parameters under maximum likelihood, these parameter estimates are then used to update the missing values. This process is repeated until the algorithm converges. It is, in fact, unnecessary to estimate each missing data point individually; rather it is enough to estimate the complete data sufficient statistics.

The EM Algorithm can be modified for problems on a case by case basis, which can require a significant amount of work; although once the algorithm has been developed it can be used repeatedly for similar problems [97]. This means that these algorithms, if understood can be constructed to account for data which are not only MCAR and MAR, as is the case with several missing data methods, but also MNAR.

Bayesian data augmentation methods were introduced by Tanner and Wong in their 1987 paper [99]. Their algorithm consists of the following, simple representation of the posterior density:

$$p(\theta|y) = \int_Z p(\theta|z,y)p(z|y)dz$$

In the above, y represents the observed data, z denotes the unobserved data and θ represents the parameter(s) of interest. The predictive density of z can then be represented as:

$$p(z|y) = \int_\Theta p(z|\phi,y)p(\phi|y)d\phi$$

Where z and y are as before and ϕ represents some set of parameters that governs the missing data process. These equations can be combined to form

$$g(\theta) = \int\int p(\theta|z,y)p(z|\phi,y)g(\phi)dzd\phi$$

At each step of the algorithm, $g(\theta)$ is updated as follows:

$$g_{i+1}(\theta) = (Tg_i)(\theta)$$

where

$$(Tg_i)(\theta) = \int\int p(\theta|z,y)p(z|\phi,y)g_i(\phi)dzd\phi$$

This means that at each step of the algorithm, the unobserved data are drawn from the current approximation of $p(z|y)$ (the imputation step) then the posterior distribution is updated to be the mixture of $p(\theta|z^{(j)},y)$ where $z^{(j)}$ represents each sample drawn from $p(z|y)$ (the posterior step). This requires that the distribu-

tions $p(z|y)$ and $p(\theta|z,y)$ should be either calculated directly where possible, or be sampled from. This provides a reasonable alternative to maximum likelihood based approaches where at least one of the densities is intractable.

4.2.4 Censored Data Methods

Methods to cope with missing data due to censoring are generally very specific to the particular application. These methods are commonly seen in examples where a time-to-event is the variable of interest. An example of this is survival analysis, which is used to model the time to event, where some subjects may have dropped out or died before the particular event of interest has happened. Thus the time to event is only known up to a lower bound. The methods used in that instance are very specific to this particular application and therefore do not easily generalise to other applications.

One commonly used modelling technique in the presence of censored data is tobit regression. This method was developed by James Tobin in 1958 in the context of household expenditure [100]. This extends the concept of OLS regression to account for the fact that some data points are observed only to an upper bound. The method combines OLS regression with probit regression in order to get a realistic estimate of the relationship where there is a clustering of data around a limiting value. This method has thus far only been developed for normally distributed data and is only relevant where there is complete truncation at the end of a distribution (i.e. all values below a threshold are coded the same). This method is therefore not suitable for the scenario of interest in this analysis, which is made up of count data which are censored over an interval.

4.3 A Solution

This scenario is fairly common although does not seem to be well documented in the literature. The general consensus when handling data which are MNAR is that the missing data process is not known and therefore cannot be estimated from the data. The general advice seems to be to treat the data as though the values are MAR, although in this case that technique is unsatisfactory as we have some knowledge of the missing data process and can use similar data points to provide additional information of how to impute these data.

The aim of this work was to take advantage of this additional knowledge and impute values which are in line with the missing data process. In order to do this, the distribution being sampled from was truncated, only for those values that were missing. In this way, the imputed values are in line with the missing data mechanism, and thus the parameter estimates should be unbiased. The model is as follows. Where the data are observed the response follows a Poisson distribution with parameter lambda:

$$y_i \sim \text{Poisson}(\lambda_i)$$

In the above, y_i represents the number of admissions to hospital at time i and λ_i is the expected number of admissions at time i.

Then, where the data are censored, the data still follow the same distribution, but the pool of values that can be selected for the data is restricted to an interval that is known a priori (indicated by the $I(\text{lower}, \text{upper})$ notation), in the specific case of our data set between 1 and 4:

$$y_i \sim \text{Poisson}(\lambda_i) I(1, 4)$$

In each case lambda is estimated in a generalised linear modelling framework. For the simple case with only one site:

$$log(\lambda_i) = \beta_0 + \beta_1 x_i + \beta_2 m(i)$$

Where x_i represents year and $m(i)$ represents the month at time i, or some transformation of that month.

In the more complex case with multiple sites there is the addition of a random effect such that

$$\begin{aligned} y_{ij} &\sim \text{Poisson}(\lambda_{ij}) \\ \log \lambda_{ij} &= \beta_0 + \beta_1 x_i + \beta_2 m(i) + \xi_k \end{aligned}$$

Where x_i and $m(i)$ are as previously. In the above y_{ij} is the number of admissions at time i in hospital j with λ_{ij} representing the corresponding expected value. The term $\xi_j \sim N(0, \sigma^2)$ is a random effect for location. These models are fit in a Bayesian framework with the following priors:

$$\begin{aligned} \beta_0 &\sim N(0, 100) \\ \beta_1 &\sim N(0, 100) \\ \beta_2 &\sim N(0, 100) \\ \frac{1}{\sigma} &\sim \text{unif}(0.0001, 10) \end{aligned}$$

The prior distributions for the betas were selected to reflect the fact that we have no knowledge a priori as to how the number of admissions varies through time. These have therefore been chosen to have a wide distribution centred

around zero. The uniform distribution was chosen for the precision parameter of the random effect to reflect lack of prior knowledge as with the β parameters. The limits were chosen such that the maximum value that the standard deviation could take was $\sigma = 1000$, this was done to avoid problems with large number calculation within the `BUGS` software which was used to estimate the model parameters.

4.4 Simulation Study

This simulation study will focus on data sets similar to the NHS data described above. The initial focus will be on a simple case where there is one "hospital" with only partial missing data. Then focus will shift to a full data set with 46 "hospitals" some with missing data and some without missing data. The specific models used to generate the data will be discussed for each case individually.

Some common methods of missing data handling have been compared as part of a simulation study. These include more straightforward methods such as list-wise deletion, average value imputation using the median and random hot deck imputation. Since the data being used are count data, the median was rounded to the nearest whole number when using average value imputation. Results for a multiple imputation using a sample based on a poisson distribution with rate parameter equal to the median of the observed values will also be presented. The models for these methods have been fit using the relevant regression function in `R`, `glm()` in the single site example and `glmer()` in the multi-site example.

A Bayesian parameter estimation method was implemented in a similar way to the method described above, without incorporating the additional information provided by the missing data mechanism. Finally the results are presented for the method described, which will be referred to as the selective sampler. Both of the Bayesian models were implemented in `OpenBUGS`.

For each method, the bias, root mean squared error (RMSE) and coverage were calculated for each parameter. A negative bias would indicate that a parameter is being underestimated on average over the simulations, while a positive value indicates that the parameter is being overestimated on average. The root mean squared error provides a measure of variability around the bias and is calculated as follows:

$$\text{RMSE} = \sqrt{\sum_{i=1}^{n} \frac{(\hat{\theta}_i - \theta)^2}{n}}$$

$\hat{\theta}$ is the value estimated in the ith simulation, θ is the true parameter value and n is the number of simulations.

For the study, 1,000 data sets were simulated. The decision to simulate 1,000 data sets was a balance between having enough data to see how well each method performed, and the need for reasonable computing time, which was of particular concern in the multi-site study.

4.4.1 Simple example with one site

One hospital requires data for each month between January 2008 and December 2015, a total of 96 data points. A sinusoidal seasonal trend was enforced in addition to a linear long term trend, with coefficients as below:

$$\log \lambda_i = 1.2 + 0.3 \sin\left(\frac{2\pi m_i}{12}\right) + 0.2 y_i$$

where m_i represents the month corresponding to observation i and y_i represents the corresponding scaled and centred value of year. This model was used to generate 1,000 data sets.

These data were created to reflect the worst case scenario; that is a small hospital with a large proportion of missing data. The simulated data sets had an average 63.6 suppressed values out of a possible 96 values. The number of missing values across the 1,000 data sets ranged from 49 missing values to 76 missing values. The full distribution of the number of missing values can be seen in Figure 4.6.

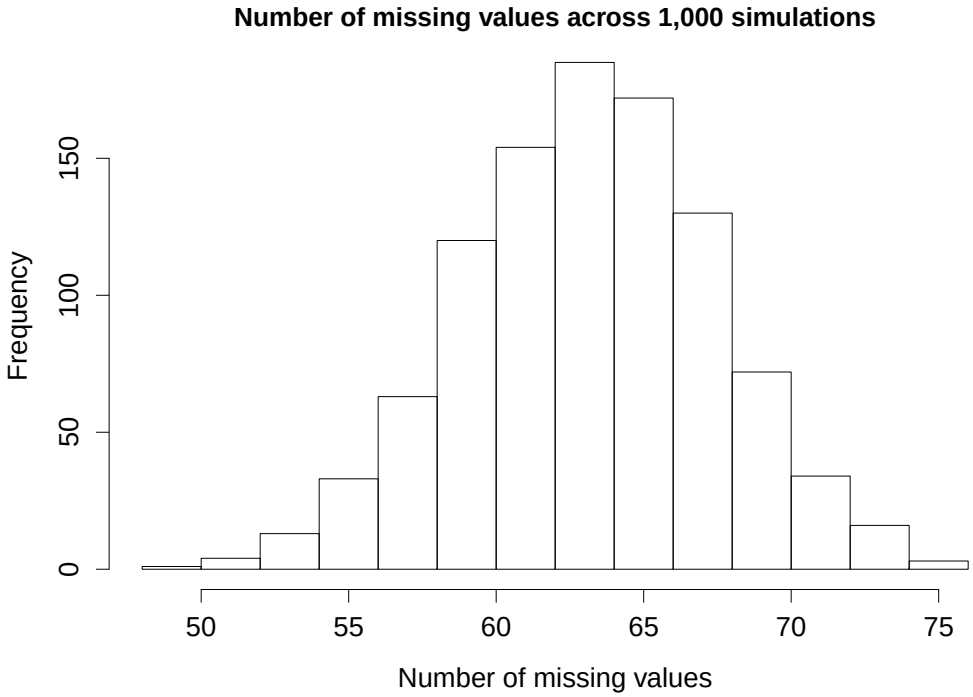


Figure 4.6: Histogram showing the number of missing values in the 1,000 simulated data sets

Each method was used to obtain parameter estimates for each of the 1,000 data sets, resulting in 1,000 different parameter estimates. These estimates were used to assess the mean bias, the mean squared error and interval estimates from each of the methods were used to assess the coverage of the method. In calculating the intervals, a 95% confidence was used, therefore the coverage should be approximately 95%. The results of these simulations are displayed in Table 4.1.

Table 4.1: Results of simulation study for small data set consisting of one small location

	Bias (RMSE)		
Method	β_0	$\beta_{\textbf{year}}$	$\beta_{\textbf{month}}$
List-wise Deletion	0.275 (0.307)	0.002 (0.113)	-0.019 (0.117)
Average Value Imputation	0.497 (0.503)	-0.220 (0.222)	-0.142 (0.148)
Hot Deck Imputation	0.465 (0.475)	-0.212 (0.219)	-0.134 (0.152)
Multiple Imputation	0.488 (0.494)	-0.269 (0.275)	-0.109 (0.115)
Regression Imputation	0.266 (0.299)	0.004 (0.113)	-0.017 (0.117)
Selective Sampler	-0.017 (0.069)	0.002 (0.063)	0.002 (0.086)
	Coverage (95%)		
Method	β_0	$\beta_{\textbf{year}}$	$\beta_{\textbf{month}}$
List-wise Deletion	29.3%	89.7%	94.4%
Average Value Imputation	0.0%	0.4%	31.8%
Hot Deck imputation	0.1%	2.1%	40.4%
Multiple Imputation	35.4%	100%	100%
Regression Imputation	31.3%	89.7%	94.4%
Selective Sampler	94.5%	95.2%	95.4%

These results appear to favour the selective sampler approach to parameter estimation. The bias is consistently small for estimates of all parameter values and the coverage is close to 95% in all three parameters. Additionally, the RMSE is small relative to the other estimates, meaning that the estimates are not varying wildly around the estimate from sample to sample.

Other methods appear to estimate certain parameters better than others. For example, list-wise deletion provides reasonable estimates for the year and month slope parameters but does not perform as well for the intercept term. The Bayesian data augmentation method that does not truncate the sampling distribution performs similarly to the complete case analysis, with low bias and reasonable coverage for the two slope parameters, but larger bias and low coverage for the intercept term.

Average value and hot deck imputation do not appear to provide reasonable estimates on any of the parameters, with larger bias than all other methods and low coverage across all three parameter estimates.

The average bias and root mean squared error, compared to the observations which were suppressed, were evaluated for the imputed values for all of the single imputation methods, and can be seen in Table 4.2.

Table 4.2: Average bias and RMSE in imputed values for the single site study

Method	**Average Bias**	**RMSE**
Average Value Imputation	3.018	3.202
Hot Deck Imputation	2.810	3.976
Multiple Imputation	2.792	2.833
Regression Imputation	1.572	2.230
Selective Sampler	-0.030	1.058

As with the parameter estimates, this suggests that the selective sampler performs better in terms of imputing the missing observations. This result is what would be expected, given that the method restricts the range of values that the missing values can take.

4.4.2 Full Example

This example more closely follows the true data set. The following model was used to generate data for each of 46 different locations:

$$\log(\lambda_{ij}) = 1.2 + 0.3 \sin\left(\frac{2\pi m_i}{12}\right) + 0.2 y_i + z_j$$

where m_i and y_i are defined as previously and $z_j \sim N(0, \sigma_z^2)$ represents a random effect for location. The parameter σ_z^2 in this example took the value 2. As previously, 1,000 data sets of this nature were generated, with each data set yielding estimates for each of the parameters in the model.

In these data sets there were 4,416 entries. On average, the simulated data sets had 1,488.16 suppressed observations. The simulation with the least suppression had 1,403 missing values, and that with the highest level of suppression had 1,569 missing observations. The distribution of the number of suppressed observations for the 1,000 simulated data sets can be seen in Figure 4.7.

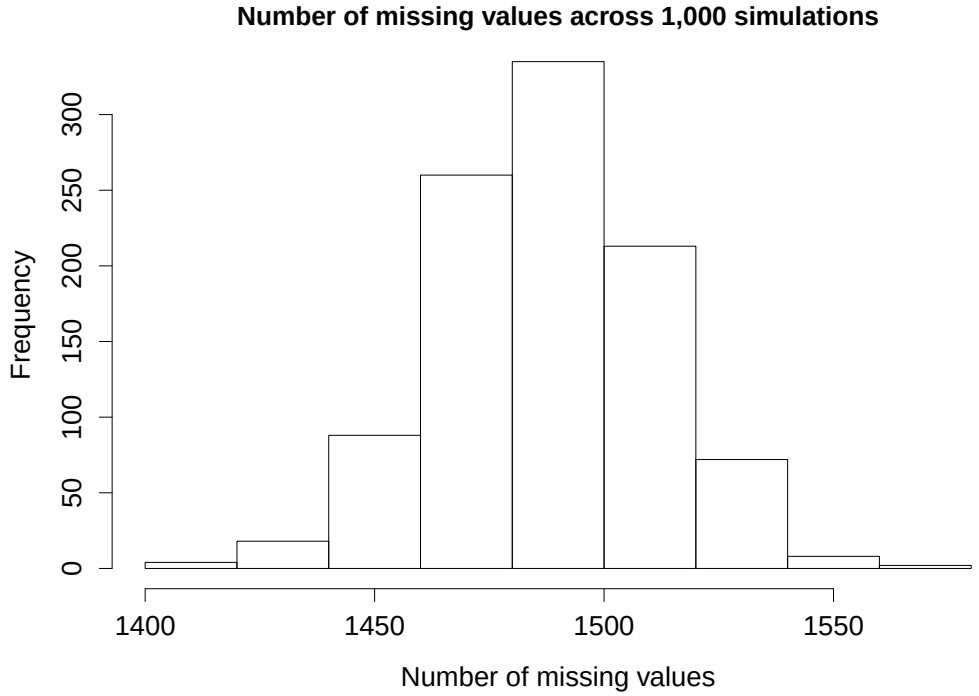


Figure 4.7: Histogram showing the number of supressed observations across 1,000 simulations of the multi-site data set

The different methods were again applied to each of the 1,000 simulated data sets, and compared on bias, MSE and coverage as previously. It is worth noting that each of average and hot deck imputation methods replaced missing values with averages of the complete cases from the relevant hospital, rather than from across the data set. The results are shown in Table 4.3.

Table 4.3: Results of simulation study for multi-site data set

	Bias (RMSE)			
Method	β_0	$\beta_{\mathbf{year}}$	$\beta_{\mathbf{month}}$	σ_z
List-wise Deletion	-0.697	-0.006	-0.004	1.116
	(0.500)	(5.1×10^{-5})	(5.8×10^{-5})	(1.127)
Average Value Imputation	-0.981	-0.036	-0.024	1.500
	(0.989)	(0.036)	(0.024)	(1.511)
Hot Deck Imputation	-0.790	-0.038	-0.025	-1.367
	(0.803)	(0.038)	(0.025)	(1.384)
Multiple Imputation	-0.781	-0.125	0.062	1.356
	(0.794)	(0.125)	(0.062)	(1.373)
Regression Imputation	-0.696	-0.006	-0.004	1.131
	(0.706)	(0.007)	(0.008)	(1.144)
Selective Sampler	0.140	-1.7×10^{-4}	-3.5×10^{-4}	-0.330
	(0.141)	(0.004)	(0.006)	(0.332)
	Coverage (95%)			
Method	β_0	$\beta_{\mathbf{year}}$	$\beta_{\mathbf{month}}$	σ_z
List-wise Deletion	99.9%	78.9%	90.6%	0.0%
Average Value Imputation	76.0%	0.0%	2.7%	0.0%
Hot Deck imputation	98.3%	0.0%	2.2%	0.0%
Multiple Imputation	100%	100%	100%	100.0%
Regression Imputation	99.8%	78.4%	90.4%	0.0%
Selective Sampler	100%	95.7%	94.8%	97.6%

As in the smaller example, the Bayesian data augmentation and complete case methods appear to produce estimates with similar performance criteria. Both of these produce reasonable coverage for the intercept and month parameters, but

with high bias and variance on the intercept term. The year slope has lower coverage at only around 78%. This may be due to the low variance estimate, which is leading to the resultant confidence intervals being too narrow. The bias on the estimates of this parameter, however was also very small.

As previously, average value and hot deck imputation do not perform well. Average value is arguably the better of the two, with higher coverage on the intercept parameter. The intervals computed using hot deck imputation produced 0% coverage across all of the parameters.

One thing worth noting is that all of the methods, with the exception of multiple imputation and the selective sampler, the examined method performed poorly in the estimation of the random effect variance parameter. The poorly performing methods had large bias, with hot deck imputation underestimating the value of σ_z by 1.367, which is a large problem, given that the true value of σ_z would be $\sqrt{2} \approx 1.4$. The coverage for these methods was 0%.

Multiple imputation had good coverage, but had relatively large bias and RMSE. In contrast, the selective sampler method performs well. The bias and RMSE are both low, particularly when compared to the other methods and the intervals estimated provided good coverage across all of the parameters. While this is true for all parameters, it is most evident in the estimate for the random effect variance, which was not particularly well estimated using the other missing data methods.

The average bias and RMSE in the imputations across the 1,000 simulated data sets can be seen in Table 4.4 for each of the single imputation methods.

Table 4.4: Average bias and RMSE in imputed values for the single site study

Method	**Average Bias**	**RMSE**
Average Value Imputation	0.663	2.922
Hot Deck Imputation	0.785	3.373
Multiple Imputation	0.609	0.780
Regression Imputation	0.178	1.996
Selective Sampler	-0.109	0.977

This indicates a conclusion similar to that seen in the single site example, although the results are similar across all four methods. This indicates that, with more data, the more basic methods are able to impute values which are similar to the true values.

4.5 Discussion

The aim of this work was to develop a robust method of estimating parameters in the situation where data obtained have been censored to prevent unintended disclosure. This will allow for future work to be carried out on data sets obtained from both the ISD and aggregated Hospital Episode Statistics from England.

The method discussed involved using the additional information that the missing values were all within a known range in order to better estimate parameters. The model was implemented in a Bayesian framework, with samples for the missing values being drawn from a truncated distribution, rather than a full distribution.

In addition to the method described, the model parameters were estimated using several other approaches to handling missing data and compared through a simulation study, first for a simple single site example and then for a larger multi-site example which was more comparable to the target data.

The results from the single site example indicated that some of the common methods, such as list-wise deletion and regression imputation, can produce reasonable estimates under data missing within a small range of values. These methods tended to be better at estimating the slope parameters, but the intercept parameter was not so well estimated. The average value and hot deck imputation methods did not estimate any of the model parameters well in the simple example, which in turn meant that the missing values were not well estimated. In contrast, the selective sampler suggested here produced unbiased estimates for the intercept terms as well as the two slope terms, with better imputation of the missing values compared to the other methods considered. In addition, it was shown that the method developed here had lower RMSE and provided better coverage.

The multi-site example proved more challenging for the simpler missing data methods. List-wise deletion and the data augmentation procedure produced reasonable estimates for some of the model parameters, where average value and hot deck imputation did not perform well at all. None of the simpler methods were able to reasonably estimate the variance parameter in the random effect. Only the selective sampler and multiple imputation methods were successful in producing reasonable estimates, in terms of bias, RMSE and coverage, for all parameters in the model.

Comparing the imputed values against the observed values which were suppressed confirmed that neither average value nor hot deck imputation were particularly robust to MNAR data, though the bias and RMSE were lower for these

methods compared to for the single site example. Regression imputation performed better, with relatively low bias. The selective sampler also performed well, though this was the only method to decrease in performance on the single site example. This is likely because the additional estimation of the random effect variance parameter impacted on its performance in estimating the intercept parameter, thus impacting the estimation of the imputed values.

One thing to note is that, while the method discussed performed well in this setting, the method relies on a regression model, which was known and used in the estimation process. As such this method may be let down where the model is misspecified. This is of little concern in the analyses to follow where there are large amounts of data which can be used in order to estimate the trends. It is, however, important to explore the trends within the hospital in order to ensure that the trends observed in the data are consistent across locations, in order to make the model specification as accurate as possible before imputing the missing values.

This chapter only examined one simulation for each case, however, the results indicate that, over 1,000 simulated data sets, this method was most appropriate for both parameter estimation and value imputation. This was particularly clear in the multi-site example, which is of most relevance to the analysis in chapters 5, 6 and 7, where the selective sampler provided a substantial improvement to the estimation of the random effect variance parameter compared to the other methods. These simulations are, therefore, sufficient to show that the selective sampler is appropriate for imputing missing values for the work presented in the remainder of this thesis.

Chapter 5

An Examination of Trends in TOXBASE access, Emergency Attendances & Hospital Admissions due to Poisoning

The focus of this chapter is on assessing temporal trends in emergency attendances and admissions due to poisoning across the UK between January 2008 and December 2015 and comparing these with trends in access to the TOXBASE database. The aim of this is to assess whether there are any consistent trends across each of the countries and across each of accesses, admissions and attendances respectively.

The data used in this chapter, and Chapters 6 and 7 is outlined in Section 5.1. The methods used in this chapter are outlined in Section5.2, though these are similar to those seen in Chapter 2. This will be followed by a presentation of the results for each of the three countries individually. The chapter will conclude with a discussion of these results.

5.1 Data

Requests were made to the Health and Social Care Information Centre (NHS Digital), the Information Services Division of NHS Scotland (ISD) and the NHS Wales Informatics Service to obtain data on admissions and attendances which were specifically recorded as being poisoning related. Here an attendance refers to any individual patient who presents at an emergency department, regardless of outcome, while an emergency admission is defined as an attendance where the patient requires a longer period of treatment (potentially overnight), which requires a more specific diagnosis reason according to ICD10. These data are not freely available to researchers, and each organisation differed in its request procedures.

The Scottish data were obtained by making contact with the unscheduled care team at ISD. Upon initial contact a discussion was opened in order to identify exactly what was required in the data extract. Following these discussions, a member of the team was able to produce the data extract.

The procedure to obtain the data from Wales was more formal than that in Scotland. Initial contact was made with the NHS Wales Informatics Service, who then issued a data request form. The completed form can be seen in Appendix B.

The process in England was similar to that in Wales. A form was filled out which detailed exactly what was required on the data extract. The completed form can be seen in Appendix C.

Previous work carried out on attendances at hospital due to accidental drug poisoning and overdose revealed inconsistencies in recording across time and within hospitals, which suggests that the data are not fit for purpose. As such

ISD were not able to release this data, and in Scotland only data on admissions and overall attendances are available. Both England and Wales provided data on the number of attendances due to poisoning.

In these data, admission reasons are coded using ICD10 [101]. The codes extracted were those related to poisoning by drugs, medicaments and biological substances (T36-T50, hereafter referred to as drugs poisoning), those related to the toxic effects of substances chiefly nonmedicinal as to source (T52-T60 & T63-T65, hereafter reffered to as other posoning). Codes corresponding to alcohol poisoning (T51) and food poisoning (T61 & T62) were excluded, as these topics are not covered within TOXBASE. For attendances, each of England and Wales have slightly different coding methods, though both have explicit codes for attendance due to poisoning, which were used in the data extraction [102, 103].

In England and Scotland, small numbers were subject to suppression. For England, all values from 1 to 5 inclusive were suppressed, while in Scotland all values from 1 to 4 were censored. Values were imputed for the Scottish and English data using the method outlined in Chapter 4, which selected the most likely value from the missing range using a Bayesian model, so that hospitals with observed values closer to the upper limit of the range were more likely to be imputed as larger values, while for those with observed values of mostly zero the imputations were more likely to be 1 or 2. An example of this can be seen for two Scottish hospitals in Figure 5.1.

In this figure, the top two plots show the observed (left) and imputed (right) values for a hospital which had the majority of its observed data at zero. The bottom two plots show the observed (left) and imputed (left) values for a slightly larger hospital where a greater proportion of non-missing values were non-zero.

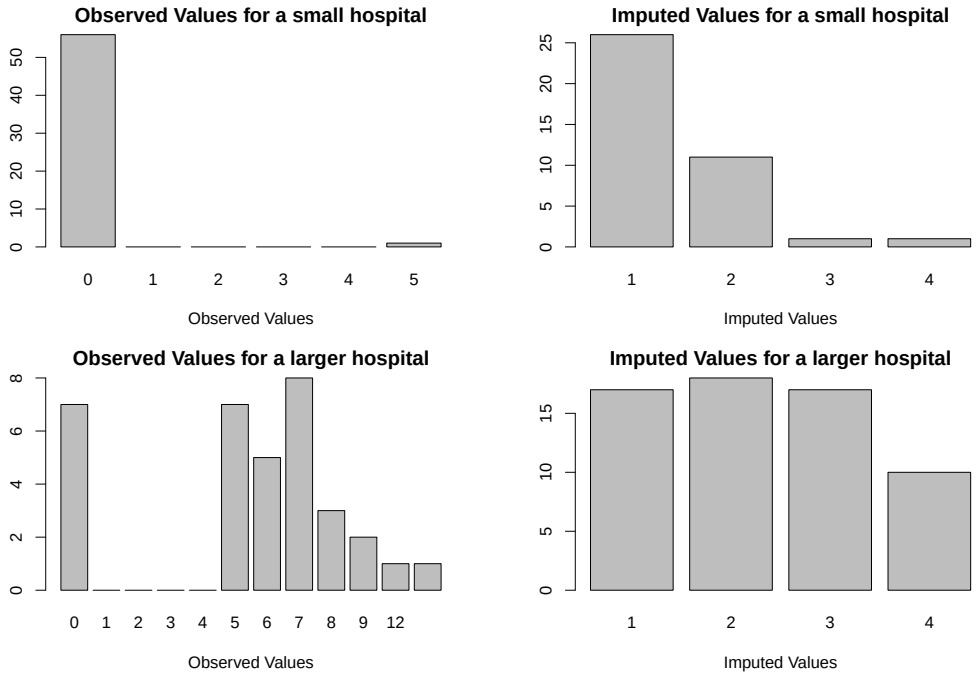


Figure 5.1: barplots showing the oberved (left column) and imputed (right column) values for a small (top) and larger (bottom) hospital in Scotland

This figure shows that some of the imputed values were 1 or 2, however further investigation indicated that these mostly occurred during longer periods of imputation.

These data were linked to the TOXBASE access database described in Chapter 3, this was done at this stage in order to ensure that the data used were consistent through this chapter and the following two chapters. This data set provides information on accesses at hospital level. The Scottish and Welsh data were similarly provided at hospital level, making linkage by hospital name straightforward, so that these data could be used with minimal processing. In the few cases where there were duplicate hospital names, these were edited in both the admissions data set and the TOXBASE access data set.

In England, however, hospitals are governed locally by NHS Trusts. Trusts can cover one site only, however, the majority of trusts cover multiple hospitals, and any data requirements are set at trust level. This means that in the majority

of cases in the English data, admissions and attendances are at trust level, though there are some data provided at hospital level. This meant that a large amount of data cleaning was required to ensure that these data were useable.

The full data cleaning and linkage process for the English data was complex and is summarised in Figure 5.2. However, for clarification, the process will be described in more detail.

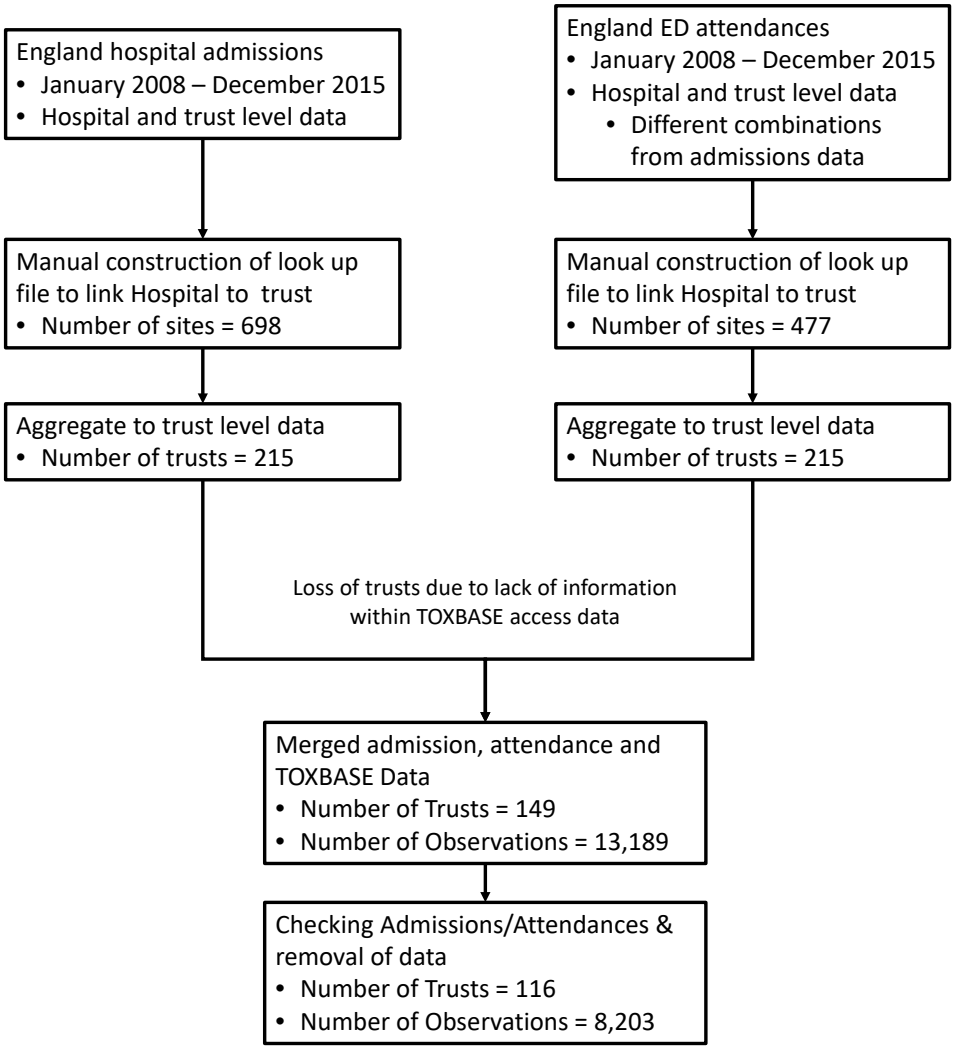


Figure 5.2: Figure showing the data cleaning and linkage process for the English data

In order to make the data as consistent as possible, any data provided at hospital level was aggregated up to trust level. This aggregation was possible through use of hospital codes, where the first three characters of the hospital code provide information on trust membership.

The data cleaning process was additionally complicated by structural changes which have occurred through the years, as trusts have closed, or merged. These structural changes were apparent, where either individual hospital codes changed, or specific trusts had stopped providing data. This was done after examination of each individual trust for gaps or sudden changes in the level of data provision. Each of these changes had to be individually investigated, by noting the time periods over which certain trusts were active, or over which there were changes in the level of admissions. The care quality commission (CQC) website was of particular use for this part of the process, particularly when data provision changed from a hospital level to a trust level, as it provides information on trust membership, past and present. Figure 5.3 shows the page for a hospital which is no longer managed by the same trust.

The hospital could be found via the URL, which contains the provider code and is consistent across all hospitals and trusts. The yellow box indicates that the provider in question no longer exists, either because it has closed or because its provider code has changed as a result of having it managing trust changed. The new provider code, and therefore trust, can be found by clicking on the "see new profile" link within the yellow box. In this way, the English admissions and attendances data sets were cleaned and aggregated ready for linkage with the TOXBASE database.

This linkage proved to be challenging, as the TOXBASE user table does not contain information on trust membership. A table was manually created, which provided the TOXBASE Hospital names which linked to each provider code from

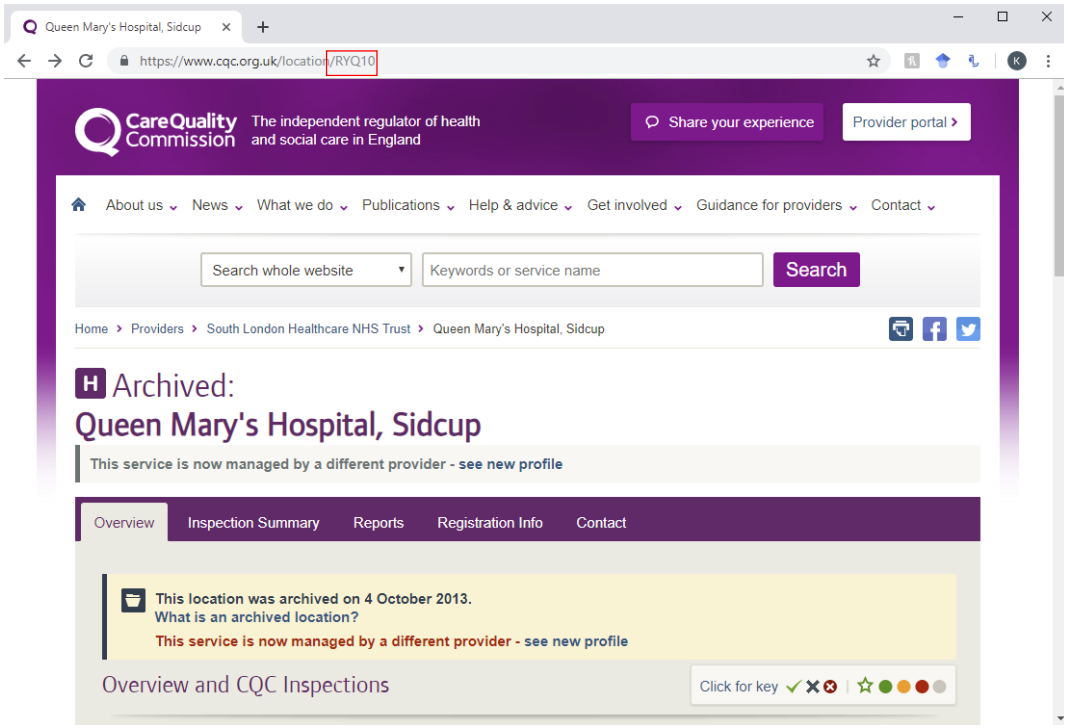


Figure 5.3: Figure showing the CQC webpage for a location which has changed management. The provider code forms part of the URL (highlighted by a red box), making these hospitals easy to locate. The new trust can be found by clicking on the "See new profile" link within the yellow box

the admissions and attendances data set at each date. Hospital name was used for linkage, as there were multiple user IDs linked to the same hospital and the date was required because of the changes in trust membership through time. The first few rows of this linkage table can be seen in Figure 5.4.

This table contains two columns corresponding to hospital codes within the admissions data, called NHS Code 1 and NHS Code 2. This was due to the fact that some trusts had multiple locations providing information at one time. It is worth noting that, despite it looking like there are dates with missing data, this is not the case, the data have been sorted by NHS Code 1.

The attendances data were not always provided in the same way as the admissions data, with some of the hospital level data having been provided at trust level, and some data which had been provided at trust level for admissions being

	A	B	C	D
1	Date	HospitalSurgery	NHS Code	NHS Code 2
2	2010-04	county hospital	5N972	
3	2010-05	county hospital	5N972	
4	2010-06	county hospital	5N972	
5	2010-07	county hospital	5N972	
6	2010-08	county hospital	5N972	
7	2010-09	county hospital	5N972	
8	2010-10	county hospital	5N972	
9	2010-11	county hospital	5N972	
10	2010-12	county hospital	5N972	
11	2011-01	county hospital	5N972	
12	2011-02	county hospital	5N972	
13	2011-03	county hospital	5N972	
14	2008-01	lymington new forest hospital	5QC00	RW100
15	2008-02	lymington new forest hospital	5QC00	RW100
16	2008-03	lymington new forest hospital	5QC00	RW100
17	2008-04	lymington new forest hospital	5QC00	RW100
18	2008-05	lymington new forest hospital	5QC00	RW100
19	2008-06	lymington new forest hospital	5QC00	RW100
20	2008-07	lymington new forest hospital	5QC00	RW100
21	2008-08	lymington new forest hospital	5QC00	RW100
22	2008-09	lymington new forest hospital	5QC00	RW100
23	2008-10	lymington new forest hospital	5QC00	RW100
24	2008-11	lymington new forest hospital	5QC00	RW100
25	2008-12	lymington new forest hospital	5QC00	RW100
26	2009-01	lymington new forest hospital	5QC00	RW100

Figure 5.4: Image showing a subset of the table used to link TOXBASE access data with admissions data

provided at hospital level for attendances. A similar data cleaning process to that used for the admissions data was followed, and a similar linkage file was produced for the attendance data.

On initial examination of the aggregated and merged data set, there were a few observations for some of the locations where the number of admissions due to poisoning was larger, and in some cases much larger, than the number of attendances due to poisoning (Figure 5.5).

In some instances this may have been due to admissions direct from ambulance, however it is unlikely that admission direct from ambulance would account for enough admissions to outweigh those cases which attended and were not admitted.

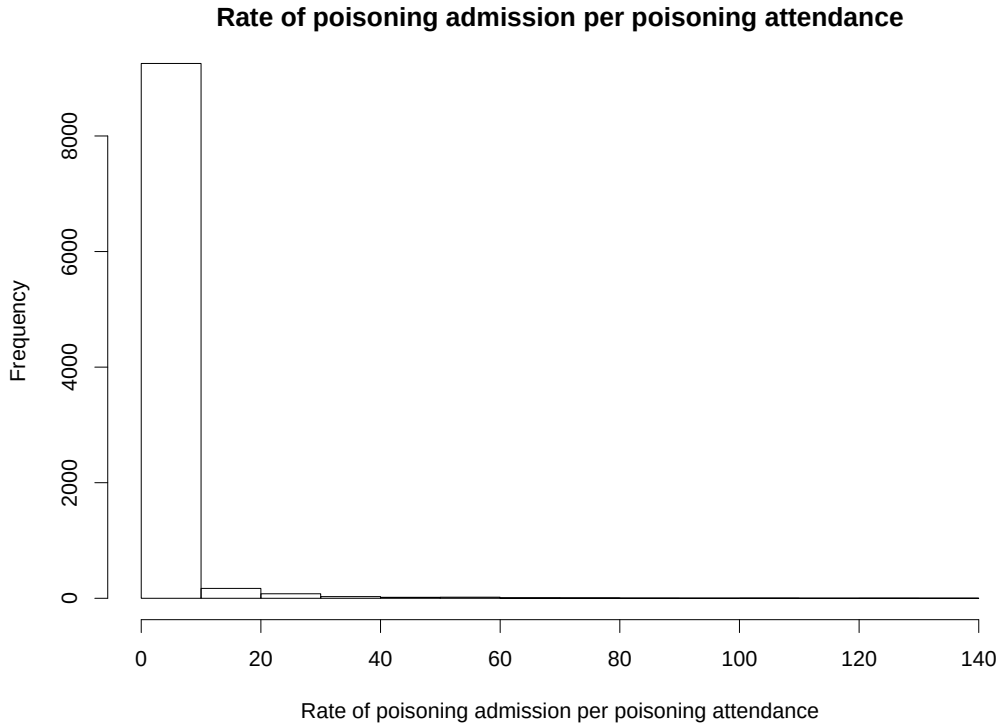


Figure 5.5: Histogram of the rate of admissions due to drugs poisoning per poisoning attendance

These cases were examined to determine whether there were any sudden changes in the level of admissions due to drug poisoning or poisoning related attendances which may indicate poor data quality. Where sudden changes in the series were found, it was likely that these were not accurately recorded at the presentation stage, therefore the decision was made to remove these data points from the analysis; where no sudden changes existed in the series, these data were retained for the analysis. As a result of this, 4,986 data points pertaining to one month in one trust were removed, leaving 8,203 observations across England between January 2008 and December 2015.

The removed data points were examined to ensure that there was no trend in these unreliable observations which may affect the overall result. Figure 5.6 shows the number of removed observations by year (left) and by month (right). These show that the number of observations removed was relatively consistent

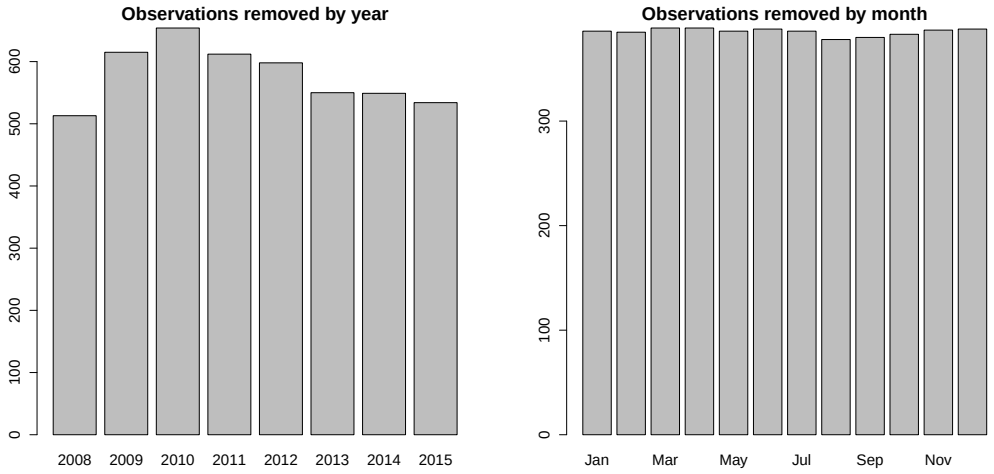


Figure 5.6: Plots showing the number of observations removed by year (left) and by month (right)

from month to month, suggesting no seasonality in the number of unreliable observations. There does appear to be some variation in the number of observations removed by year, with more observations removed in 2010 than in any other year. The increase in unreliable data in 2010 and, to a lesser extent, in the surrounding years may be linked to reforms set out in the 2010 white paper *Equity and excellence: Liberating the NHS* [104].

It is these data, along with the data for Scotland and Wales, that are used throughout this chapter, Chapter 6 and Chapter 7.

5.2 Methods

The methods used in this chapter were similar to those used in Chapters 2 and 3. Generalised additive models were fit via the `mgcv` package [59] in `R` [63].

Section 5.3.1 describes the trends in the absolute number of admissions due to drugs poisoning and TOXBASE accesses across Scotland, England and Wales individually. The models used in this section were of the form shown in Equation 5.1.

$$E(\log(y_l(t))) = \beta_0 + f(m(t)) + g(t) + \gamma_l \tag{5.1}$$

The term $y_l(t)$ represents the number of admissions due to drugs poisoning or the number of TOXBASE accesses at location l. β_0 represents an overall average value, while the smooth terms for seasonal and long term trends ($f(m(t))$ and $g(t)$ respectively) are deviations from this, where $m(t)$ is a month indicator for time t. Since this model describes the number of accesses and admissions respectively, no offset is included in this model, a hospital/trust level random effect (γ_l) was instead used to account for variation in hospital size. The output from these models is presented as the centred response, which is the response minus all other terms in the model. For example the centred response for the seasonal term is modelled on $\log(y_l(t)) - \beta_0 - g(t)$.

The models fit in Sections 5.3.2 and 5.3.3 are similar to that described in Equation 5.1, but with the addition of $\log(x_l(t))$ (Equation 5.2).

$$E(\log(y_l(t))) = \log(x_l(t)) + \beta_0 + f(m(t)) + g(t) \tag{5.2}$$

This additional term acts as an offset, and allows for the modelling of rates and proportions. In Section 5.3.2 this offset will represent the overall number of attendances, and will represent either poisoning attendances or overall attendances in Section 5.3.3. This will depend on whether there is evidence to suggest that it is appropriate to use overall attendances as a proxy for poisoning attendances.

5.3 Results

The results are split up into three main sections. The first (Section 5.3.1) will examine the trends in the number of poisoning admissions and the number of TOXBASE accesses for Scotland, England and Wales individually. This will be followed by an assessment of whether poisoning attendances occur as a constant proportion of all attendances (Section 5.3.2), hence allowing the use of overall attendances as a substitute for poisoning attendances, so that Scotland can be included in future analyses. Finally results on the rate of poisoning admissions and TOXBASE accesses, per poisoning attendance or overall attendance (as indicated by Section 5.3.2), for the appropriate countries.

5.3.1 Trends in the number of accesses and admissions

This section describes similarities in the trends in absolute numbers of TOXBASE accesses and admissions due to drugs poisoning.

Scotland

After the data linkage procedure, data on admissions and TOXBASE accesses were observed for at least one month in 42 hospitals, with 644 observations out of 3,982 having been imputed. Between January 2008 and December 2015, there were a total of 410,453 accesses to TOXBASE from Scottish emergency departments and 104,530 admissions due to drugs poisoning at these hospitals. In 2008 there were around 910 accesses and 317 admissions due to drugs poisoning on average per hospital. By 2015 access to TOXBASE had increased to 1,532 accesses per hospital, while admissions due to drugs poisoning remained consistent at 313 admissions per hospital on average.

The Scottish data were structured as shown in Table 5.1. The Year and Month columns were combined to form the Year.use column, such that Year.use = Year + (Month - 1)/12. The Accesses column contains a count of the number of accesses made to pharmaceuticals pages in TOXBASE, while Total_Accesses refers to the number of accesses made to any page in TOXBASE. Drugs, Alcohol and Other refer to admissions due to drug poisoning, admissions due to alcohol poisoning and admissions due to other types of poisoning respectively.

Table 5.1: Table showing the structure of the Scottish data

Hospital	Year	Month	Accesses	Total_Accesses	Drugs	Alcohol	Other	Admissions	Year.use
arran war memorial hospital	2008	1	0	0	0	0	0	36	2008.000
arran war memorial hospital	2008	2	0	0	0	0	0	37	2008.083
arran war memorial hospital	2008	3	0	0	0	0	0	45	2008.167
arran war memorial hospital	2008	4	0	0	1	0	0	51	2008.250
...	...	...	...	...	...	...	...	...	...
glasgow royal infirmary	2015	9	388	444	103	7	<NA>	3839	2015.667
glasgow royal infirmary	2015	10	390	409	108	13	<NA>	3915	2015.750
glasgow royal infirmary	2015	11	325	350	87	8	5	3987	2015.833
glasgow royal infirmary	2015	12	239	271	88	9	6	3962	2015.917

It was decided after examining the data that the proportion of missing values within the Other variable was very high at nearly 40% and other poisoning accounts for only a small number of admissions. These values were, therefore, not imputed. The response was taken to be the number of admissions due to drugs poisoning, since the TOXBASE database does not contain information on the treatment of alcohol poisoning.

The overall time series of accesses (top) and admissions (bottom) are shown in Figure 5.7. This shows a temporal trend in TOXBASE accesses which is similar to those seen in Chapter 3, with access to the database having largely increased between 2008 and 2015. The number of admissions appears to have remained relatively consistent through time, though admissions appear to have been slightly lower between 2008 and 2012, compared to the period from 2012 to 2015.

Histogram of the average accesses (left) and admissions (right) by hospital can be seen in Figure 5.8. There are a fairly large number of small, community hospitals in Scotland, which is reflected in the right skewed distribution of both TOXBASE accesses and admissions. A mid-sized hospital had, on average 26.6 admissions due to drugs poisoning and 90.2 admissions. The smallest hospital made no admissions due to drugs poisoning, while the largest hospital made 135.6 admissions on average per month.

These series are shown in Figure 5.9 for a subset of hospitals covering a small (top), a medium (middle) and a large (bottom) hospital, with the accesses series on the left and the admissions series on the right.

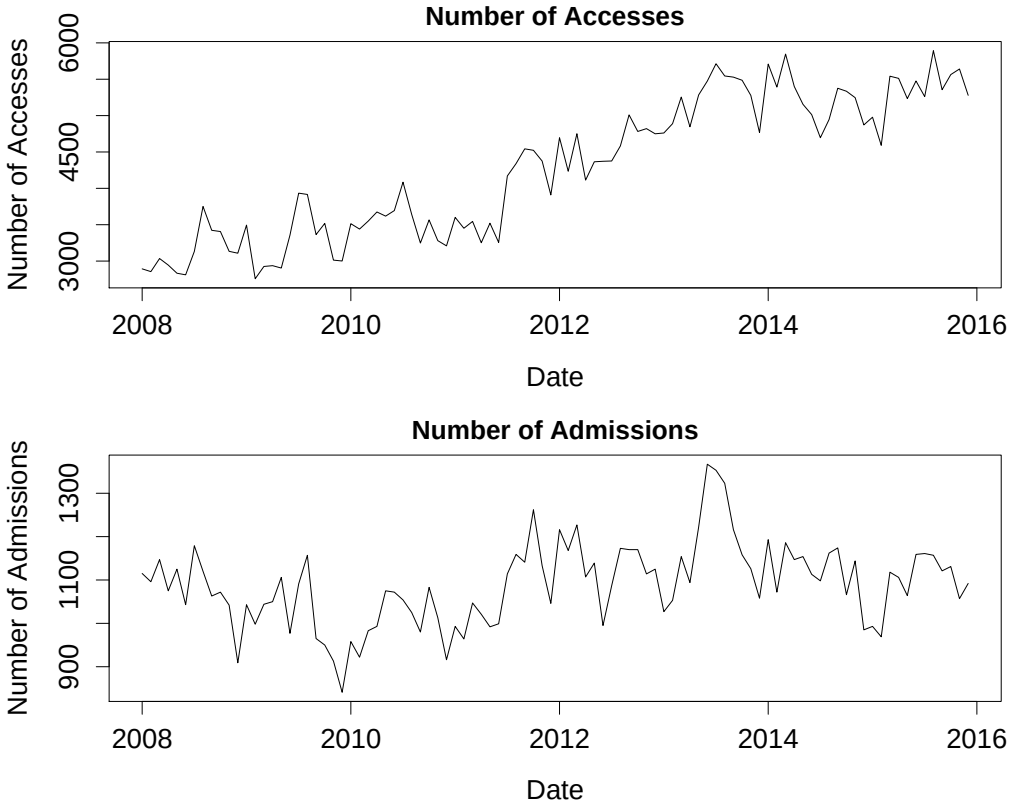


Figure 5.7: Figure showing the observed number of accesses (top) and admissions due to drugs poisoning (bottom) by month in Scotland

It is difficult to discern any trend in accesses and admissions in the small hospital. Though there is some suggestion of an increase in use of TOXBASE in more recent year, as evidence by fewer level period at zero, and larger peaks in access. There does not appear to be such an increase in the number of admissions for the small hospital.

In both the mid-sized and large hospitals, there appears to be a consistent trend in the number of TOXBASE accesses made over time. Both had relatively slow growth in usage of TOXBASE initially, though use of the system then began to grow at a greater rate from 2012 onwards. The trend in admissions, however does not appear to be consistent between the large and mid-sized hospitals. There

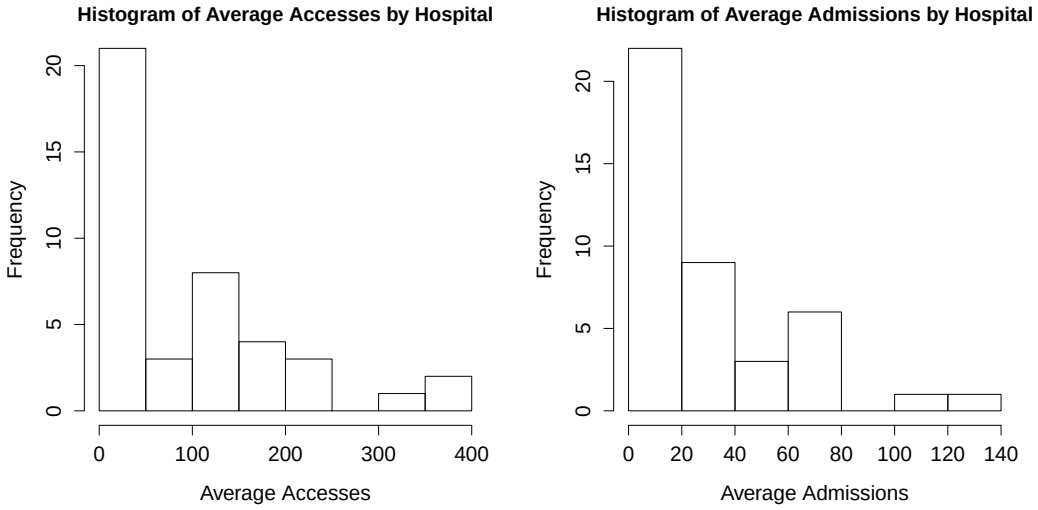


Figure 5.8: Histograms of the average number of monthly accesses (left) and admissions due to drugs poisoning (right) by hospital

is a slight increasing trend in the number of admissions in the larger hospital, while admissions appear to have remained similar across time in the mid-sized hospital.

In order to assess the overall trend in accesses and admissions across Scotland, a model of the form shown in Equation 5.1 was fit to the number of accesses to pharmaceutical pages in TOXBASE (`Accesses`) and the number of drugs admissions (`Drugs`), producing the trends shown in Figure 5.10.

The average seasonality shown in the left hand plot of Figure 5.10 shows that there is some small seasonality in TOXBASE use (grey lines), with a peak in September, which is slightly later than that suggested in Chapter 3, for Great Britain as a whole. However, this seasonality is not significant (P = 0.101). This can also be seen by the fact that a horizontal line can be drawn through the 95% confidence interval depicted by the dashed grey lines. There is, however seasonality in the number of admissions due to drug poisoning, with admissions peaking in August. At the peak, these admissions were 9.4% (95% CI: 6.7%, 12.2%) higher compared to January.

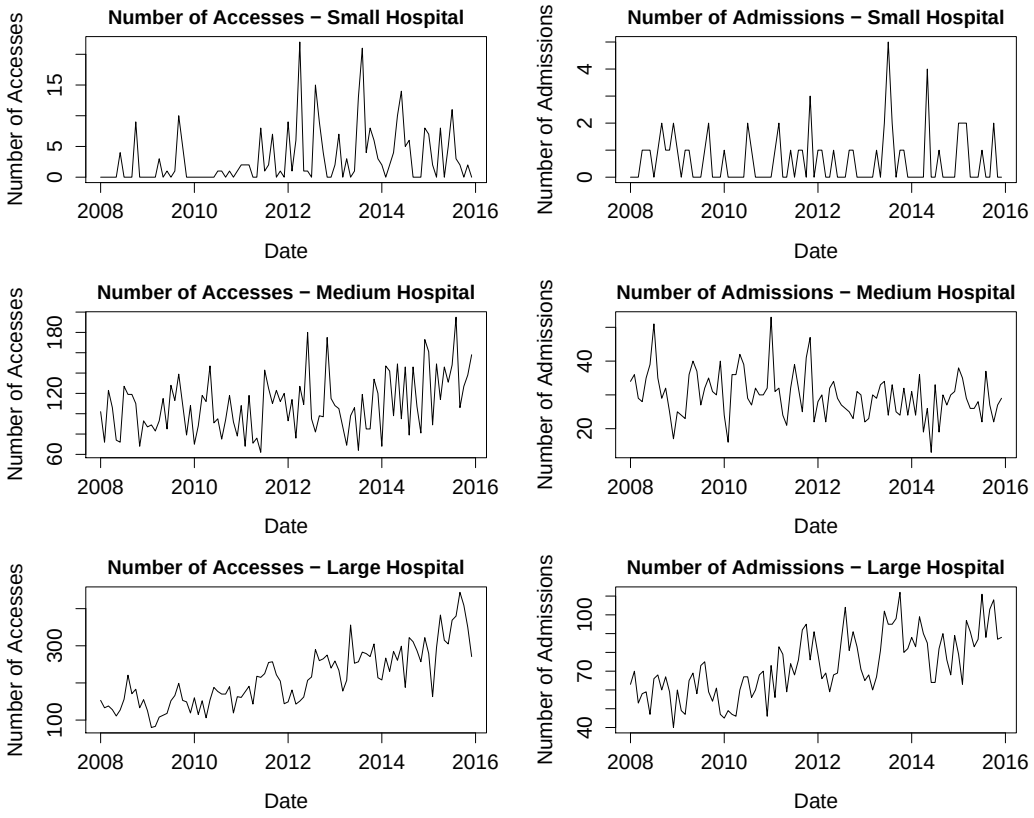


Figure 5.9: Time series showing the number of accesses (left) and admissions due to drugs poisoning (right) for a small (top), medium (middle) and large (bottom) hospital respectively

In the right panel of Figure 5.10, it can be seen that the number of admissions due to drugs poisoning was approximately the same in 2015 as it was in 2008. Overall the difference between January 2008 and December 2015 is equivalent to a non-significant 3.7% (95% CI: -0.9%, 8.1%) decrease. However, there has been some fluctuation in poisoning admissions during the study period, with slightly lower admissions in the period 2010/11, and slightly higher admissions in 2013/14.

This analysis also indicates that access to TOXBASE from Scottish emergency departments has increased by 58.8% (95% CI: 46.2%, 72.4%), which is roughly in line with what was seen in Chapter 3 for Great Britain as a whole.

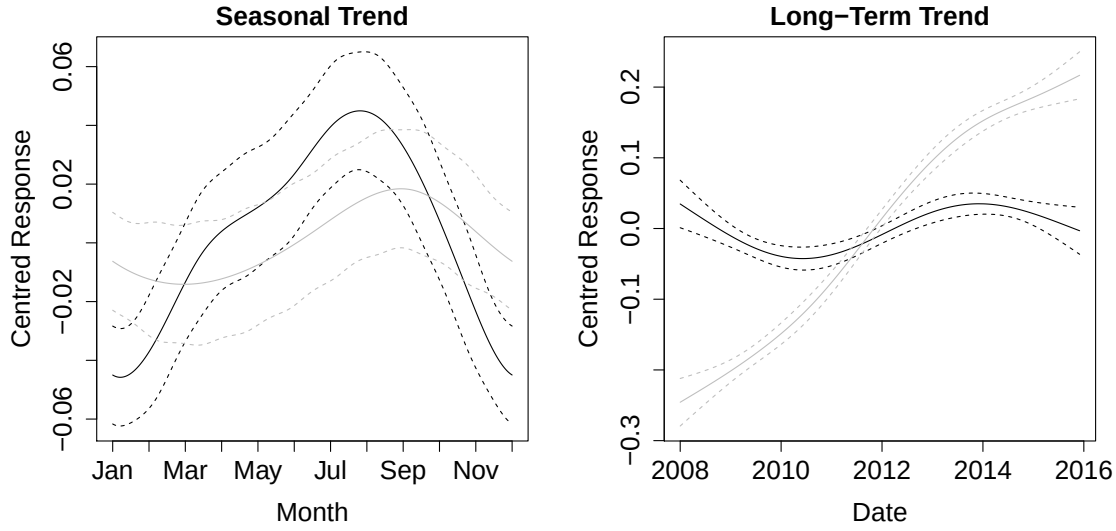


Figure 5.10: Seasonal and long term trends in admissions due to drug poisoning (black) and TOXBASE accesses (grey) in Scotland between 2008 and 2015

England

In the English data, 8,368 out of 87,138 observations of admissions and 5,419 out of 84,483 poisoning attendances were imputed. Following the data linkage with the TOXBASE data, there were observations for at least one month for 116 trusts. After accounting for those cases where the poisoning attendance data were unreliably recorded, 2,495,718 accesses were made to TOXBASE from emergency departments and 558,754 admissions due to drugs poisoning over the period from January 2008 to December 2015. In 2008 there were around 639 admissions and 2,205 TOXBASE accesses per hospital on average, compared to 762 and 4,165 in 2015. The data used in this analysis was structured as shown in Table 5.2 on the next page.

Table 5.2: Table showing the structure of the Scottish data

Hospital	Month	Year	All_Access	Pharma_Access	Admiss_Any	Attend_Total	Attend_Poisoning
R1F	January	2012	169	150	30	3212	61
R1F	March	2015	103	91	22	5372	69
R1F	December	2009	146	106	14	2834	40
R1F	November	2011	179	162	37	3340	54
...	...	...	...	...	...	...	...
RYR	August	2010	279	230	84	10978	171
RYR	May	2014	489	355	108	12111	170
RYR	October	2015	500	423	113	11654	169
RYR	April	2014	476	397	103	11421	162

Prior to analysis, Month and Year were combined in order to obtain a Year.use column for use in the model fit. All_Accesses and Pharma_Accesses are the number of TOXBASE accesses overall, and the number of accesses made to pharmaceuticals pages within TOXBASE. Admissions_Any is the number of hospital admissions due to drugs poisoning. There is additionally information on the total number of attendances (Attendances_Total) and the number of attendances attributed to poisoning (Attendances_Poisoning).

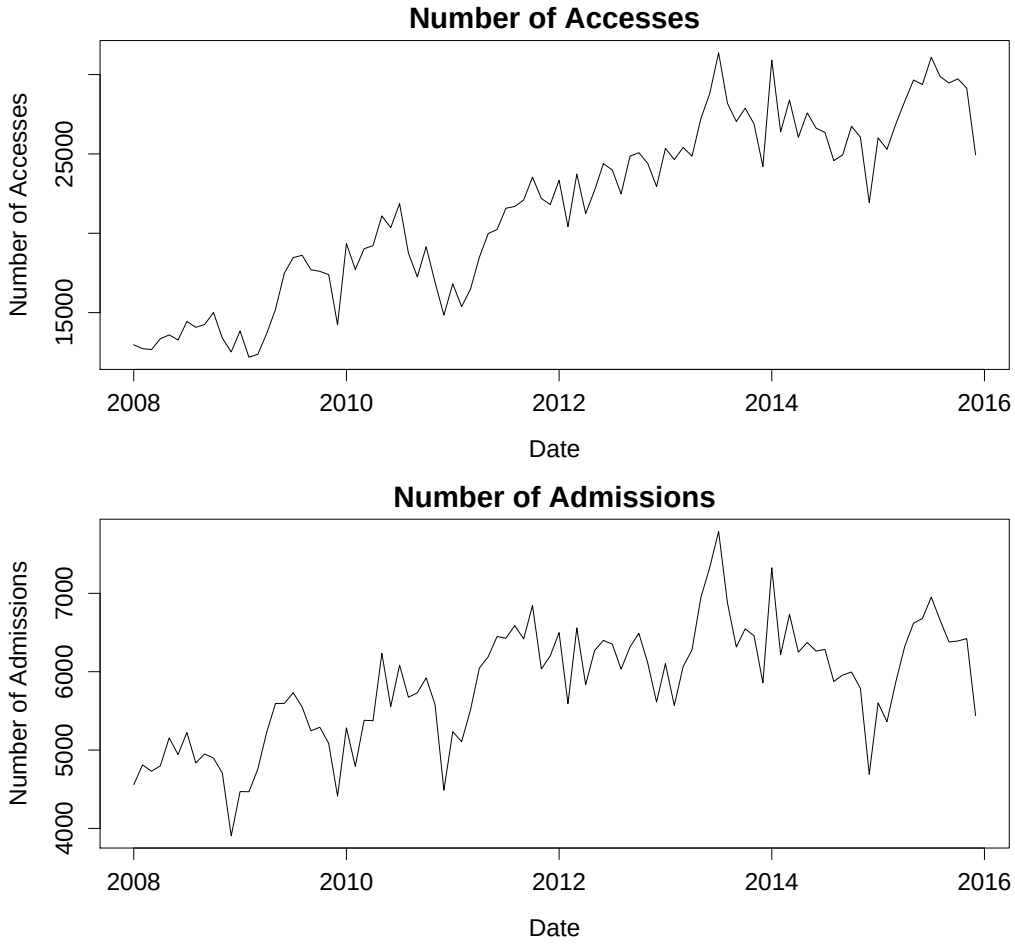


Figure 5.11: Plots showing the number of accesses to TOXBASE (top) and the number of admissions due to drug poisoning (bottom) by month in England

The time series of accesses and admissions across England can be seen in Figure 5.11. This indicates that access to TOXBASE has been generally increasing since 2008, with this increase seeming to slow in more recent year. There also appears to have been an increase in the number of poisoning admissions through time, although this increase appears less dramatic than that seen in the TOXBASE accesses.

Histograms of the average number of accesses and admissions per month can be seen in Figure 5.12. Since trusts are an amalgamation of one or more hospitals, there are relatively few trusts with a very small number of accesses or admissions. Average accesses tended to be around 200, with admissions being around 70 on average.

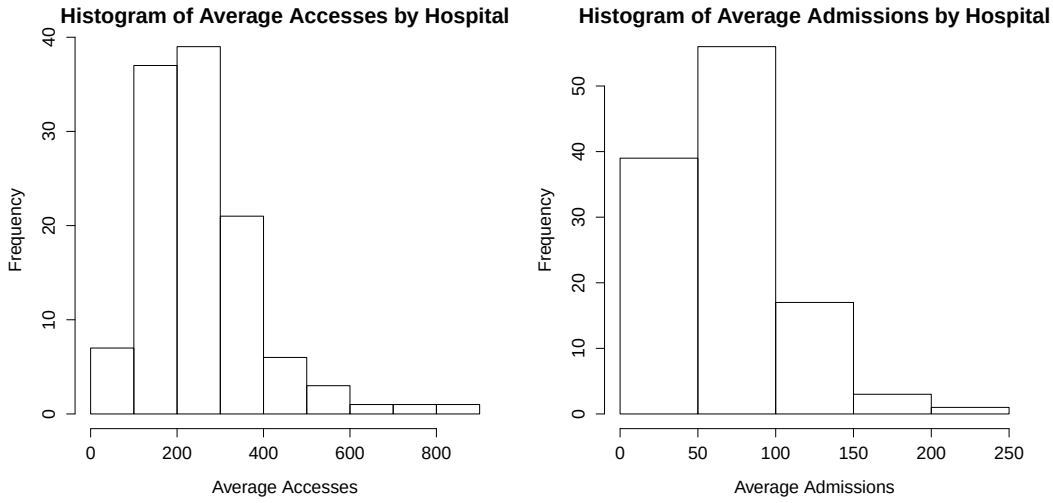


Figure 5.12: Plots showing the average number of accesses to TOXBASE (left) and the average number of admissions due to drug poisoning (right) by trust in England

As with the Scottish data, trends in admissions are presented in Figure 5.13 for a small, medium and large trust. These were selected from the pool of trusts which had observations available for all time points. These plots indicate that the

trend in admissions across all three trust sizes is similar, though there is perhaps a more pronounced increase in the number of admissions in later years in the small trust.

In terms of TOXBASE use, all three trust sizes vary in their trends, with the small trust showing an increase in TOXBASE access. Admissions due to poisoning in the medium trust appear to be relatively consistent across time, while there was an initial decrease in poisoning admissions in the larger trust before the number of admissions then increased from 2010 onwards.

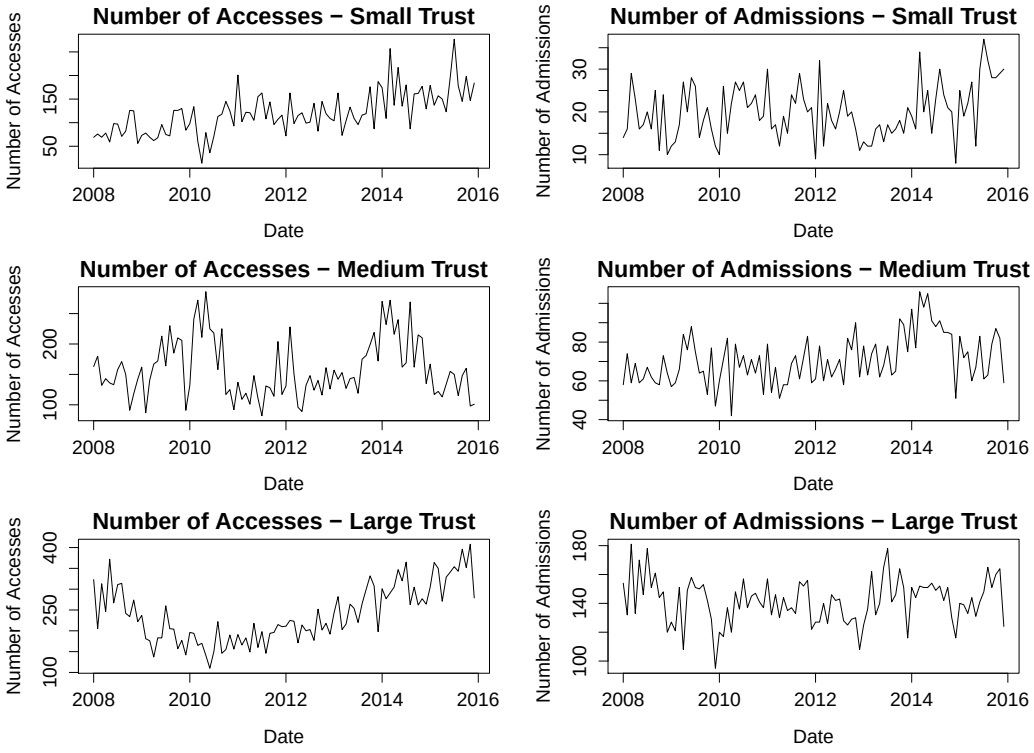


Figure 5.13: Plots showing accesses (left) and admissions (right) for a small (top), medium (middle) and large (bottom) trust

As in the Scottish analysis, a model was fit in order to extract the overall trends in the number of accesses made to TOXBASE and the number of admissions due to drugs poisoning (Figure 5.14). The plot of seasonal trends shows that the overall pattern in the two entities are very similar, with a peak in both accesses and admissions around June or July. The seasonal trends overall are al-

most identical; the difference from January until the peak in TOXBASE accesses was equivalent to 16.2% (95% CI: 14.2%, 18.3%), while in poisoning admissions, the same difference was equivalent to 16.1% (95% CI: 14.3%, 17.8%).

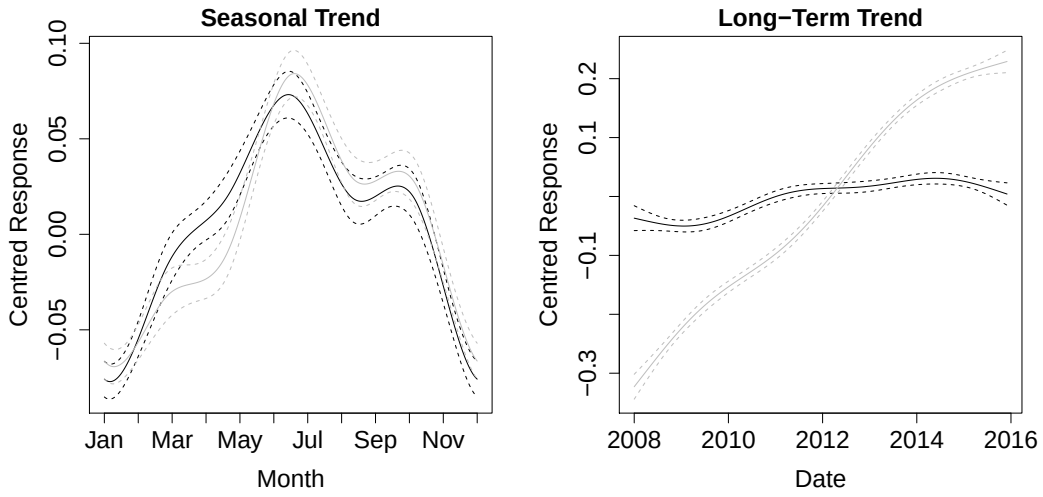


Figure 5.14: Seasonal and long term trends in admissions due to drug poisoning (black) and TOXBASE accesses (grey) between 2008 and 2015

The right panel of Figure 5.14 shows the long term trend in admissions and attendances. This shows that access to TOXBASE has increased overall. This increase means that TOXBASE use in December 2015 was on average 73.7% (95% CI: 68.2%, 79.4%) higher than what it was in January 2008. This plot also shows that admissions due to drug poisoning have remained approximately level, such that the change in the number of admissions corresponds to a 4.2% (95% CI: 1.3%, 7.1%) increase between January 2008 and December 2015. The difference between this percentage difference estimate and the difference in admissions described previously is likely due to differences in the amount of data between January 2008 and December 2015; 73 trusts out of 116 had accurately recorded data in January 2008, compared to 91 trusts in December 2015.

Wales

There were 13 hospitals in the Welsh data, which contained information on admissions and TOXBASE use for at least one month. Across these hospitals, there were 134,128 TOXBASE accesses and 22,614 admissions due to drugs poisoning between 2009 and 2015. There were approximately 207 admissions due to poisoning and 686 TOXBASE accesses per hospital in 2009 which increased to 291 and 2,239 respectively in 2015.

The Welsh data were aggregated in a similar way to the English data, and had similar column names. As previously, the Month and Year columns were combined to form a time indicator column (Year.use - not shown) for use in the model fit. TOXBASE accesses were stored in a column called All_Accesses and accesses to pharmaceuticals pages were stored in a column called Pharma_Accesses. Admissions_Any represented the number of admissions to hospital due to poisoning, and Attendances_Total and Attendances_Poisoning denoted the number of attendances overall and the number of attendances due to poisoning.

Table 5.3: Table showing the structure of the Scottish data

Hospital	Month	Year	All_Access	Pharma_Access	Admiss_Any	Attend_Total	Attend_Poisoning
morriston hospital	April	2009	69	57	42	6202	82
princess of wales hospital	April	2009	59	51	11	5707	61
university hospital of wales	April	2009	137	91	24	10472	42
withybush general hospital	April	2009	43	32	8	3330	33
...	...	...	...	...	...	...	...
withybush general hospital	December	2015	133	15	8	2991	32
wrexham maelor hospital	December	2015	259	40	53	5247	82
ysbyty glan clwyd	December	2015	164	11	34	4568	53
ysbyty gwynedd	December	2015	179	16	44	4119	49

The number of accesses to TOXBASE and the number of admissions due to drug poisoning in Wales for each month between January 2008 and December 2015 can be seen in Figure 5.15. This indicates the number of accesses to TOXBASE has been increasing fairly consistently over the period studied, with an increase in growth over 2014 and 15. The number of admissions has been fairly consistent over time, with the exception of 2009, where fewer hospitals provided data. Since the models were fit at hospital level, this will not affect the analysis.

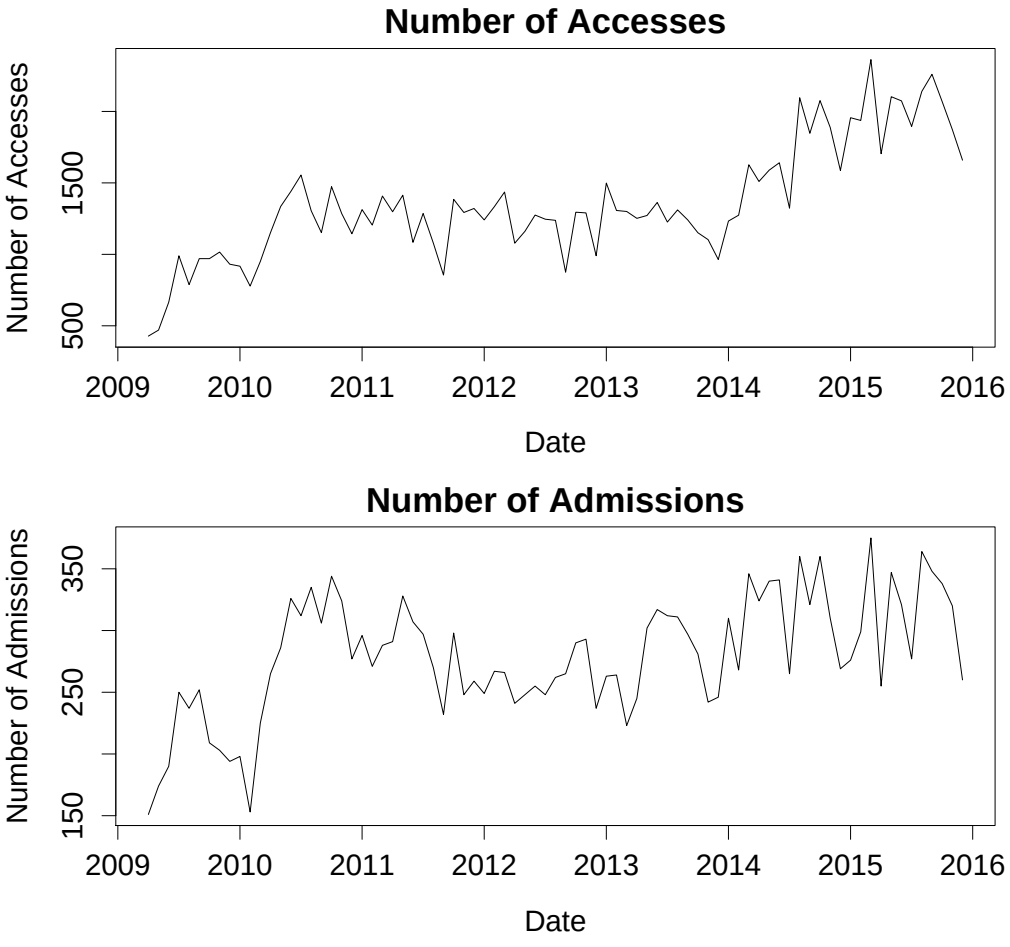


Figure 5.15: Number of TOXBASE accesses (top) and admissions due to drug poisoning (bottom) per month across Wales

The histograms in Figure 5.16 show the average number of TOXBASE accesses, and the average number of admissions due to drugs poisoning for each hospital. This indicates that all hospitals in Wales had at least 50 accesses to TOXBASE per month on average, with most having either between 50 and 100 average accesses per month, or having between 150 and 200 accesses per month. Most hospitals had more than 10 admissions due to drugs poisoning on average per month, with a large proportion of these having between 10 and 20 admissions per month on average.

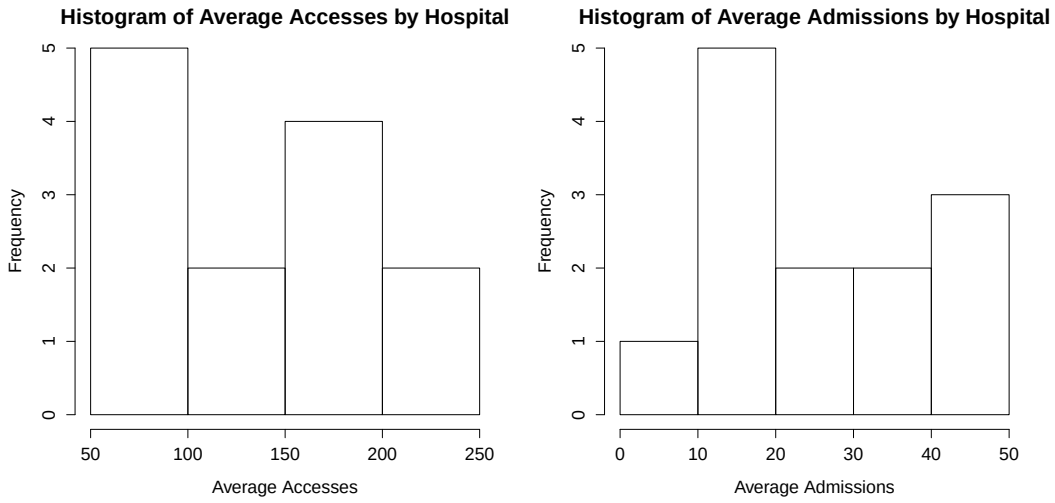


Figure 5.16: Histogram of the average number of accesses to TOXBASE (left) and admissions due to drugs poisoning (right) per month

The monthly number of accesses to TOXBASE and admissions due to drugs poisoning can be seen in Figure 5.17. Note that the dates across the three hospitals are not consistent, due to lack of data particularly in 2009 and 2010. These indicate that there has been fairly consistent growth in TOXBASE use across all three of the hospitals examined. There are similarities in the overall trends in admissions for the small and mid-sized hospitals, with admissions being relatively consistent through time. In the large hospital examined, there appears to have been sudden growth in admissions due to drugs poisoning at the end of 2013.

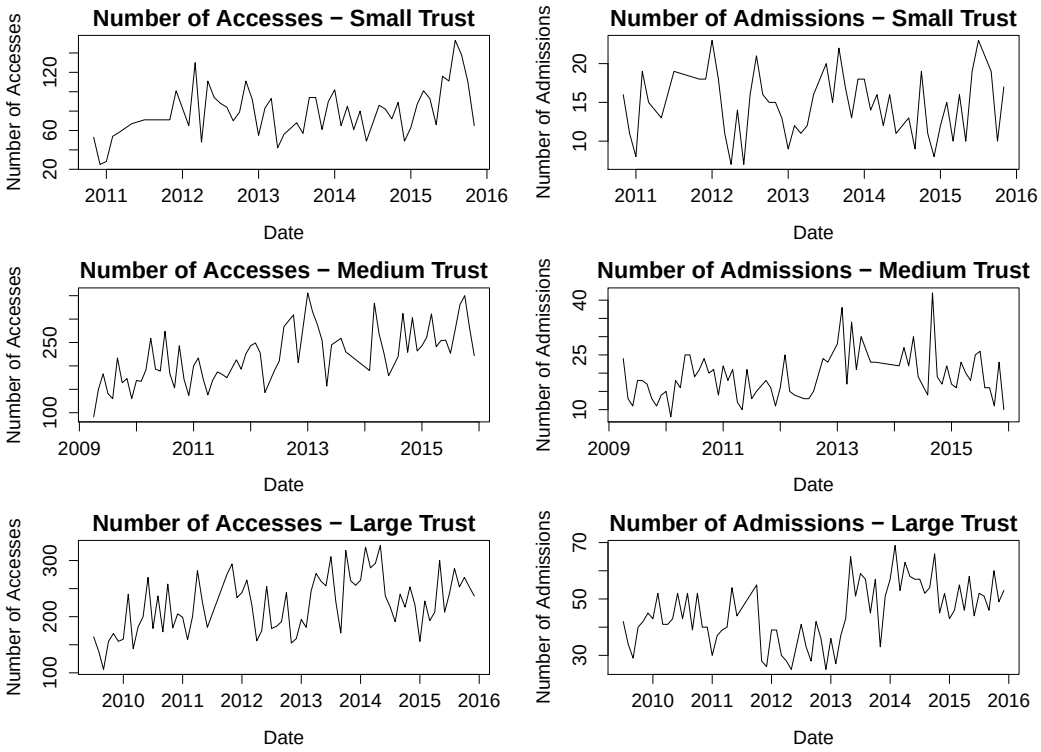


Figure 5.17: Series of the number of TOXBASE accesses (left) and the number of admissions due to drugs poisoning (right) for a small (top), medium (middle) and large (bottom) hospital

Figure 5.18 shows that similar seasonal trends exist in both TOXBASE accesses and poisoning admissions. Both trends peak in around June or July, though the TOXBASE seasonality is less defined that the seasonality in admissions. The change in TOXBASE accesses between January and the peak is equivalent to 8.0% (95% CI: 0.4%, 16.3%), while there was an increase of 18.0% (95% CI: 12.3%, 23.9%) in the number of admissions due to poisoning.

It is clear that both admissions due to poisoning and TOXBASE accesses have increased in Wales between 2009 and 2015. The increase in TOXBASE use is larger, with TOXBASE use in Wales being 2.2 (95% CI: 1.7, 2.9) times what it was in 2009. The change in admissions is smaller, with an increase of 25.7% (95% CI: 5.9%, 49.3%) between 2009 and 2015.

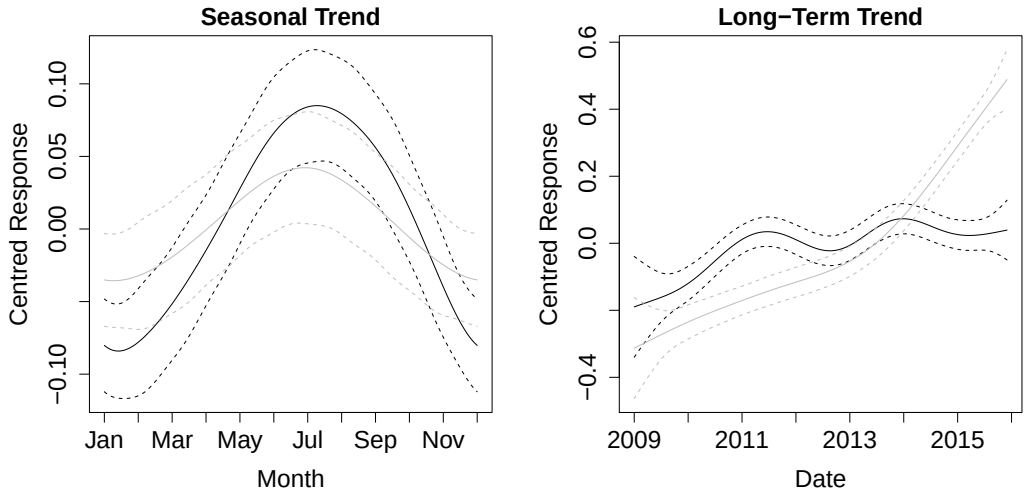


Figure 5.18: Seasonal and long term trends in admissions due to drug poisoning (black) and TOXBASE accesses (grey) between March 2009 and 2015

Tables 5.4 and 5.5 provide comparisons of the key values in the trends in access to TOXBASE and admissions due to drugs poisoning respectively.

The seasonality in Scotland appears to be quite different to that in both England and Wales, as can be seen in the second and third columns of Table 5.4. The peak in seasonality in Scotland is later, in September, compared to the peak in seasonality in either England or Wales, which was in June and July respectively. The magnitude of the seasonal effect was different in all three countries, with Scotland having the lowest level of seasonality, followed by Wales, and England had the largest estimated seasonal effect.

Table 5.4: Table comparing the trends in the TOXBASE accesses data from Scotland, England and Wales

	Month of Peak	**% difference to peak**	**% difference 2008-2015**
Scotland	September	2.5% (-0.4%, 5.4%)	58.8% (46.2%, 72.4%)
England	June	16.2% (14.2%, 18.3%)	73.7% (68.2%, 79.4%)
Wales	July	8.0% (0.4%, 16.3%)	120% (70%, 190%)

Scotland and England were more similar in terms of their long term trends, with increases in TOXBASE use of 58.8% and 73.7%. Wales demonstrated a much larger growth in TOXBASE use, increasing by around 120% over the period examined.

A summary of the seasonal trends in admissions due to drugs poisoning can be seen in the second and third columns of Table 5.5. This shows that the seasonality in Wales is similar to that in England, with a similar percentage difference from the start of the year to the peak in summer. The estimated peaks are in June and July for England and Wales respectively, though both of these are subject to error, and it is therefore not unreasonable that these peaks would occur in line with one another. The same is true for the peak in the seasonality in admissions in Scotland, although the effect size is around half of that for England and Wales.

Table 5.5: Table comparing the trends in the admissions data from Scotland, England and Wales

	Month of Peak	**% difference to peak**	**% difference 2008-2015**
Scotland	August	9.4% (6.7%, 12.2%)	-3.7% (-8.1%, 0.9%)
England	June	16.1% (14.3%, 17.8%)	4.2% (1.3%, 7.1%)
Wales	July	18.0% (12.3%, 23.9%)	25.7% (5.9%, 49.3%)

The long term trends in the number of admissions due to drugs poisoning is very different, with admissions decreasing in Scotland and increasing by an estimated 4.2% and 25.7% in England and Wales respectively. However, it is worth noting that this large increase in the number of accesses in Wales has a large uncertainty around it, due to a lack of data in some hospitals in 2009 and 2010 and that the actual percentage increase may be as small as 5.9%.

5.3.2 Trends in the proportion of poisoning related attendances

This section examines trends in the proportion of overall attendances in England and Wales which were recorded as relating to poisoning. This will provide information on whether it is appropriate to include the Scottish data in future analyses. If there is no evidence of trend in the proportion of poisoning attendances, then this indicates that poisoning attendances have occurred at a constant rate with respect to attendances overall, and it is, therefore, possible to use overall attendances as a substitute for poisoning attendances.

In England, there were 1,172,944 attendances recorded as being poisoning related, representing around 1.6% of the 74,876,814 recorded attendances for any reason between January 2008 and December 2015. Between April 2009 and December 2015, there were 3,872,345 attendances in Wales, of which around 1.2%, or 47,913 of which were poisoning related.

In 2008 there were on average 85,141 attendances per hospital in England, with 1,261 recorded as being poisoning related, so that poisoning accounted for around 1.48% of all attendances on average in England in 2008. This increased to around 1.65% in 2015, with 106,014 attendances overall and 1,745 poisoning related attendances. There was also an increase in attendances in Wales, with 40,081 attendances overall and 425 poisoning attendances per hospital in 2009, meaning that poisoning accounted for 1.06% in Wales in 2009. In 2015 poisoning accounted for around 1.35% of all attendances in Wales with 47,549 total attendances and an average of 641 poisoning attendances per hospital in 2015.

Figure 5.19 shows the raw series for poisoning attendances as a proportion of all attendances for England and Wales respectively. This indicates that there has been a steady increase in the proportion of poisoning attendances at emergency departments in both countries, though Wales continues to have a lower proportion of attendances compared to England.

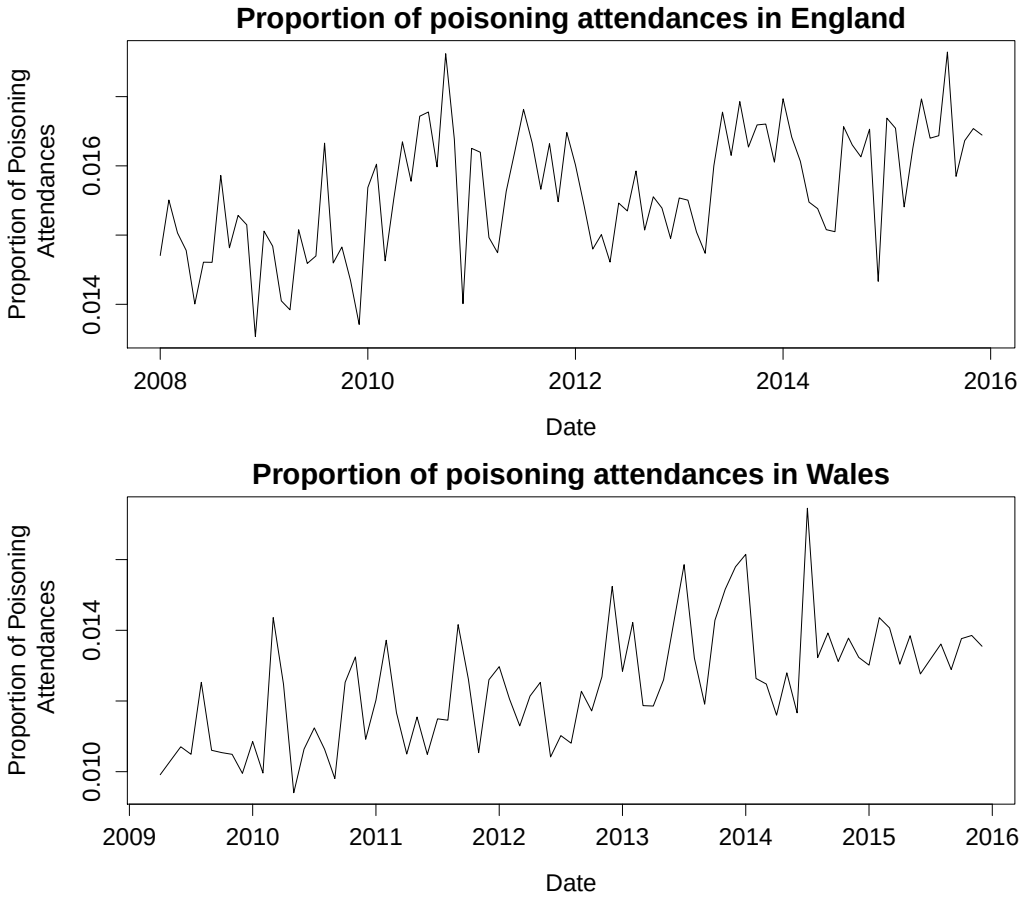


Figure 5.19: Plots showing the proportion of poisoning related attendances per month for England (top) and Wales (bottom)

Figure 5.20 shows the seasonal and long term trends in the proportion of poisoning attendances for England (left column) and Wales (right column). Focussing on the top row of this figure, it appears that poisoning attendances have been increasing at a faster rate than attendances overall. In Wales, poisoning

attendances have been consistently increasing as a proportion of all attendances, while in England an increase in the proportion of poisoning attendances can only be observed from late 2009 onwards.

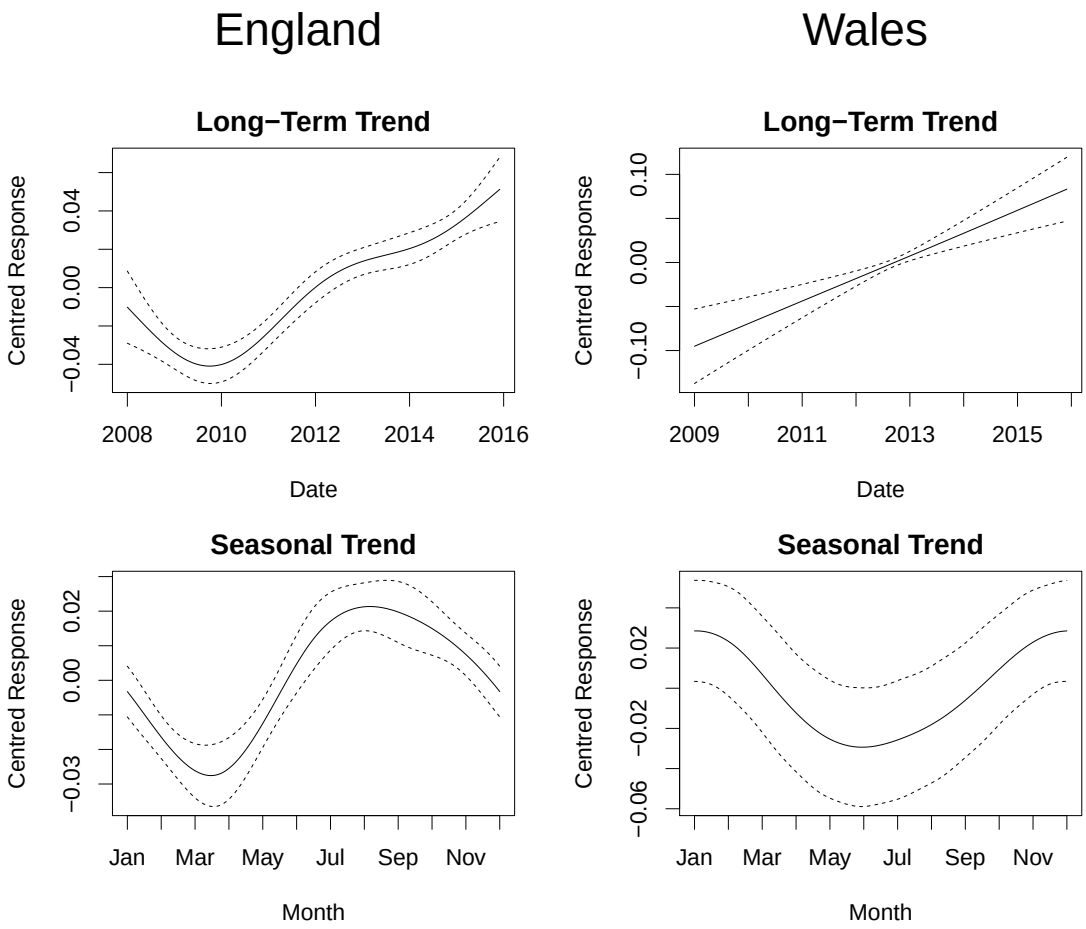


Figure 5.20: Seasonal and long term trends in the proportion of attendances related to poisoning for England (left) and Wales (right)

The seasonal trends are not consistent in England and Wales. In England the proportion of attendances related to poisoning is at a minimum in April and a maximum in August. In contrast, the proportion of poisoning related attendances in Wales is at a minimum in June, and a maximum across the months of December and January.

The key result from this analysis is that poisoning attendances do no occur as a constant proportion of attendances overall in either England or Wales. It would be possible to use one of the estimated trends in order to derive a multiplier for the Scottish data, however, it would be impossible to know which of the two estimated trends to use, and both are very different. Therefore, it is not appropriate to use overall attendances in place of poisoning attendances in the analysis to follow. Any analysis moving forward will include only data from England and Wales, due to lack of information on poisoning attendances in the Scottish data.

5.3.3 Trends in poisoning admissions and TOXBASE accesses as a rate of attendances

This section will assess the trends in the rate of TOXBASE accesses and admissions per poisoning attendance for England and Wales. The reason for doing this is that we would expect both TOXBASE access and admissions due to poisoning to increase with the number of poisoning related attendances, however it is not known whether the rate of access or admission per poisoning attendance have consistent temporal patterns, which may suggest influence beyond attendance numbers alone.

England

The left panel of Figure 5.21 shows the seasonal trends in the rate of TOXBASE access per poisoning attendance (grey) and the rate of poisoning admission per poisoning attendance (black). There is no significant seasonality in the rate of TOXBASE accesses, with the approximate significance test producing a p-value of 0.228. This non-significance can also be seen as a horizontal line can be drawn through the 95% confidence interval depicted by the grey dashed lines. This

suggests that TOXBASE is being used in line with attendances. In contrast, there is some residual seasonality in admissions due to drug poisoning after accounting for attendances. This is surprising, given that the seasonality in the number of admissions and TOXBASE accesses was so similar. The peak in the residual seasonality in admissions occurs sometime in April and corresponds to an increase of 2.2% (95% CI: 1.0%, 3.5%) from January. This peak coincides with a period of time where the seasonal trend in admissions is higher than TOXBASE accesses from Figure 5.14.

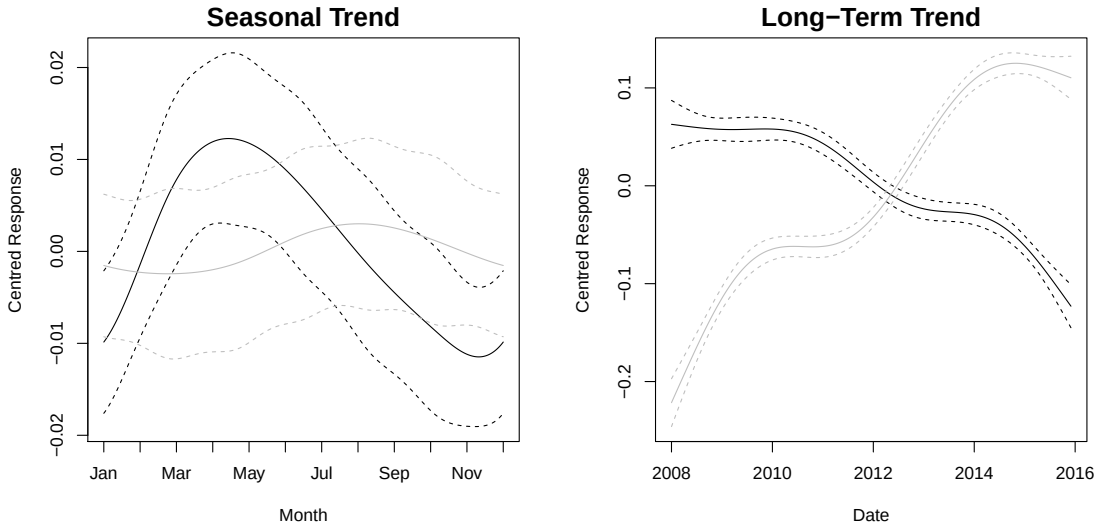


Figure 5.21: Seasonal and long term trends in admissions rate of due to drug poisoning (black) and TOXBASE accesses (grey) in England between 2008 and 2015

The plot of the estimated long term trends shows that TOXBASE use has been increasing at a greater rate than poisoning attendances, with an overall increase in the rate of TOXBASE access per attendance of 39.4% (95% CI: 34.1%, 44.8%) between January 2008 and December 2015. This plot also indicates that poisoning attendances have been increasing at a greater rate than admissions related to drugs poisoning, resulting in an overall decrease of 17.0% (95% CI: 14.2%, 19.6%) between January 2008 and December 2015.

Wales

As in England, it can be seen from Figure 5.22, that there is no residual seasonality in TOXBASE accesses as a rate of poisoning attendances, again suggesting that TOXBASE is used in line with poisoning related attendances. As previously, there is residual seasonality in poisoning admissions as a rate of poisoning attendances, which approximately corresponds with the seasonal trend seen in Figure 5.18, which peaks in August. The resulting change is a 9.8% (95% CI: 5.2%, 14.6%) increase in the rate of poisoning admissions per poisoning attendance between January and August.

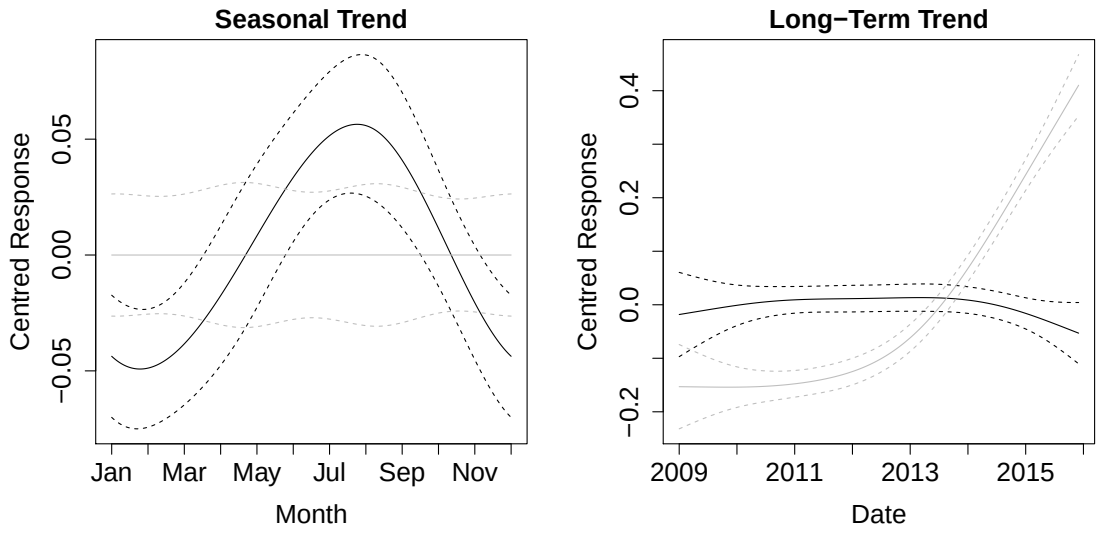


Figure 5.22: Seasonal and long term trends in the rate of admissions due to drug poisoning (black) and TOXBASE accesses (grey) in Wales between 2009 and 2015

As in England, it is clear from Figure 5.22 that TOXBASE use has increased at a greater rate than poisoning attendances overall. However, the trend is approximately level initially, suggesting that TOXBASE use was growing in line with poisoning attendances initially, and then something triggered a sudden growth in TOXBASE use in 2012. This may be linked to the release of the NICE recommendation that suggested that TOXBASE be used in cases of poisoning [45]. The

result was an overall growth in the rate of TOXBASE accesses of 76.9% (95% CI: 38.6%, 125.8%) between 2009 and 2015. This plot also indicates that admissions due to poisoning have increased at a similar rate to poisoning attendances with an overall non-significant change of -4.8% (95% CI: -16.5%, 8.4%) over the same period.

5.4 Discussion

This chapter has demonstrated the temporal trends that exist in TOXBASE accesses and in admissions due to drug poisoning. Initially trends in the number of admissions due to poisoning and in TOXBASE accesses were assessed for Scotland, England and Wales. These trends were estimated across each country, by fitting a model over individual hospitals, with differences in hospital size accounted for using random effects. This was done in order to retain a measure of the variability across hospitals or trusts within each country. The hospital random effect also helps to account for autocorrelation within hospitals or trusts.

The analysis indicated that there were similarities in the seasonal trend in both TOXBASE access and admission due to drugs poisoning across all three countries, and that TOXBASE use had generally increased at a greater rate than poisoning admissions.

The Scottish data were limited by the lack of data on hospital attendances related to specific reasons such as poisoning, due to recording inconsistency. In order to use the data available on admissions in Scotland, it would need to be assumed that poisoning attendances occur as a constant rate of overall attendances. The analysis at country level in Section 5.3.2 was sufficient to show that this is unlikely to hold, and therefore the Scottish data was excluded from the analysis in Section 5.3.3 and in Chapters 6 and 7. This rendered a more detailed regional

analysis unnecessary. The loss of Scottish data corresponded to a loss of 15% in admission numbers and around 14% in TOXBASE accesses, which represents a fairly small loss. There are around five times as many admissions in England compared to Scotland and around 6 times as many TOXBASE accesses.

After examining the rate of TOXBASE access per poisoning attendance in England and Wales, it was clear that there is some evidence to suggest that TOXBASE is being used in line with attendances due to the lack of seasonality in TOXBASE accesses after accounting for attendances due to poisoning. This suggests that locations are complying with the recommendation that TOXBASE should be consulted if a patient is suspected to be poisoned [45].

There was some residual seasonality in poisoning admissions, after accounting for poisoning attendances. This suggests that there is some other contributing factor to admission due to drug poisoning. One factor that was not accounted for was the hospital capacity for each month. It is known that hospitals are at higher capacity in the winter months [105], and therefore it may simply be the case that in winter hospitals do not have room, and are therefore less able to admit a borderline poisoning patient. The number of hospital beds have also decreased through time [105], suggesting that there may be some implication of this on the reduction in the proportion of poisoning attendances admitted overall.

The seasonality seen in the proportion of poisoning attendances and the rate of poisoning admissions per poisoning attendance may also indicate some seasonality within poisoning cases themselves, for example variation in the types of pharmaceutical being used at different times of the year, leading to more severe cases of poisoning in the summer months. However, without further data on either the substances involved or the severity of the case this hypothesis cannot be tested.

This chapter served as an exploration of the admissions and attendances data, with the specific goal of determining whether it would be feasible to incorporate the Scottish data into future analyses. For this reason, it was only the overall country trends which were assessed as part of this chapter. This was sufficient to show that the lack of data on the number of attendances due to poisoning meant that the Scottish data should not be used.

Chapter 6

Outliers in TOXBASE Use & Admissions due to Poisoning

This chapter explores the presence of unusual values, and includes an informal comparison of the attributes of outlying users and those within normal ranges. This will include characteristics such as hospital size, location and type for all hospitals and trusts. There will also be an assessment of TOXBASE Access and Admissions for English trusts using Summary Hospital-level Mortality Indicator (SHMI) and Care Quality Commission (CQC) ratings. Due to availability of data, these used the classifications for 2017/18, which were the most recent at the time of analysis.

The overall goal of the analysis in this chapter is to determine whether there are any potential features which are more common in those hospitals which either admit or use TOXBASE differently to the other locations. By using information on hospital size, location and type it may be possible to assess whether there are any consistencies in those hospitals which either use TOXBASE in an unusual way or those that admit in an unusual way. The hope is that, by compar-

ing TOXBASE access with SHMI and CQC ratings, it may be possible to see some relationship between TOXBASE and hospital quality, which may point to TOXBASE use having a positive impact on the treatment of poisoned patients.

6.1 Methods

This chapter used funnel plots [106] in order to assess whether a hospitals admissions or accesses were within the expected range. Funnel plots are scatterplots of an observed rate at a specific location against some measure of the population at that location. The idea behind using funnel plots is that uncertainty around a measurement tends to decrease where there is a greater population.

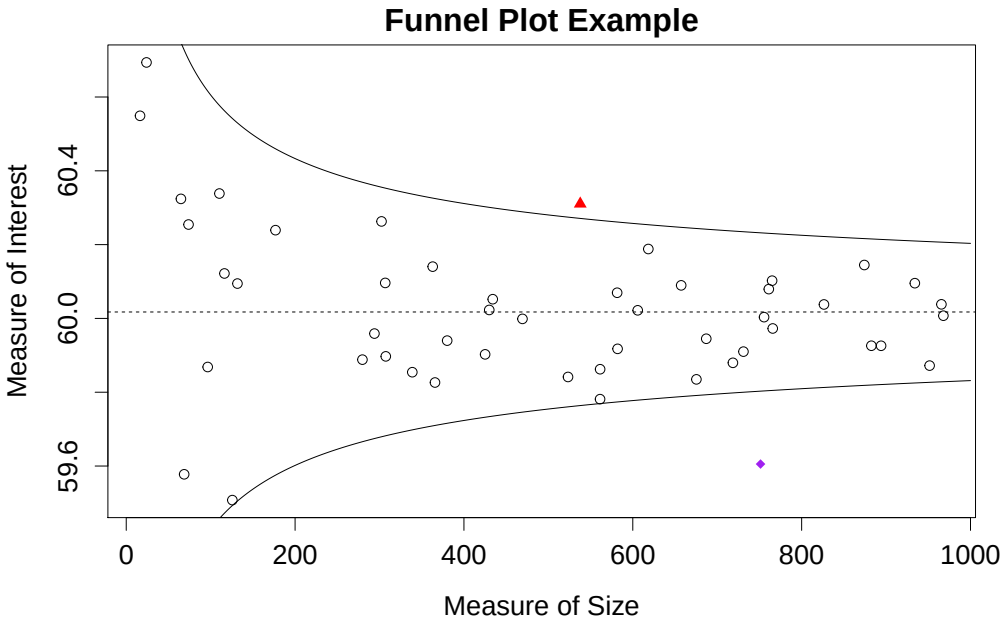


Figure 6.1: Example of a funnel plot. The dashed line represents the mean and the two solid lines represent the funnel limits. The red triangle is a point above the upper control limit and the purple diamond is a point that lies below the lower control limit

An example of a funnel plot can be seen in Figure 6.1. On the y-axis is some measure of interest, in this chapter this will be either the rate of TOXBASE accesses per poisoning attendance or the rate of admission due to drugs poisoning

per poisoning attendance. The x-axis represents some measure of size, which is usually the denominator in the rate of interest. Indeed, in these analyses the number of poisoning attendances will be plotted on the x-axis. Each point on the funnel plot represents a single location. For this reason, the monthly data described in the previous chapter were aggregated across time, to produce a data set with one observation per trust which would help to identify regions which were consistently accessing or admitting at an unusually high or low rate across the entire time period examined. The data were also assessed at annual aggregation to assess whether the unusual sites persisted across all years.

The straight dashed line on the funnel plot in Figure 6.1 represents the overall data average, which is constant over the range of the x-axis. In this analysis, this will be the average rate of TOXBASE access or the average rate of admission due to drugs poisoning. The two curved lines represent the control limits. A point lying above the upper control limit (red triangle in Figure 6.1) suggests that that location has an unusually high measure of interest, while a point lying below the lower control limit (purple diamond in Figure 6.1) indicates that a location has an unusually low measure of interest.

It is common practice for these control limits to be 95% or 99.8%, and can be thought of as being equivalent to significance testing. However these can be set to any desired coverage. In order to be conservative about the determination of outliers in this data, 75% limits were chosen as the boundaries at which an observation might be considered an outlier. The lower control limit for TOXBASE accesses was also amended to take into account the fact that hospitals should, according to NICE guidelines, be consulting TOXBASE for every poisoned patient. Therefore those locations which are found to have a rate of access less than one will be highlighted as an outlier, even if it lies within the 75% bounds.

These funnel plots make a distributional assumption in order to produce appropriate limits. As discussed previously in Chapters 3 and 5, these data exhibit overdispersion, meaning that the poisson distribution is not appropriate for constructing these intervals. In this instance, a quasi-poisson distribution was used in order to account for the relatively large amount of variability seen in these data [106].

The control limits for these analyses were calculated using an adapted version of Byar's method [107]. A quasi-poisson distribution is such that $E(Y) = \lambda$ and variance $\text{Var}(Y) = \phi\lambda$. The parameter ϕ accounts for overdispersion by multiplying the poisson variance. This parameter was estimated using an intercept only generalised linear model with a quasi-poisson distribution of the form:

$$E(\log(y_i)) = \log(n_i) + \beta_0$$

where y_i is the number of TOXBASE accesses of poisoning admissions and n_i is the number of poisoning attendances at that trust. The lower limit (M_l) is calculated as

$$M_l = \frac{P\left(1 - \frac{1}{9P} - z\sqrt{\frac{\phi}{9P}}\right)^3}{n}$$

where P is the expected number of cases, $\exp(\beta_0) \times n$, ϕ is the model estimated dispersion parameter and z is the appropriate quantile of the standard normal distribution. The upper limit (M_u) is calculated in a similar way:

$$M_u = \frac{(P+1)\left(1 - \frac{1}{9(P+1)} + z\sqrt{\frac{\phi}{9(P+1)}}\right)^3}{n}$$

Having used the funnel plots to identify unusual locations, an exploratory assessment of the data was conducted using several indicator variables which may impact the treatment of poisoned patients. The first consideration was hospital size, the total number of attendances and the number of hospital beds were compared based on outlier status.

Specific TOXBASE use was also considered, in order to assess whether outlying trusts had any differences in the profiles of poisoning cases that attended. Accesses to six specific drug groups were considered, which reflect some of the most commonly accessed product pages in the TOXBASE database: antidepressants, paracetamol, non-opioid drugs of abuse, opioids (including medications but excluding heroin), antipsychotics and heroin. The choice to separate heroin from the rest of the opioids was made due to the difference in usage of these two groups, since opioids are regularly used for pain management, while heroin is a drug of abuse.

Outliers were also compared by their university status, the thought being that medical students may be encouraged to use TOXBASE more as part of their training. University hospitals and teaching trusts were extracted by determining whether "University", "College" or "Teaching" appeared in the hospital or trust name.

Outlier status was examined by region in order to determine whether there was any spatial pattern in the number of outliers in terms of admissions or TOXBASE accesses. This may help to identify whether there are differences in practice over larger areas.

Finally, the TOXBASE access rates were compared with two indicators of care quality, in order to determine whether there was any association between trust level TOXBASE use and hospital performance. The two characteristics used were Summary Hospital-level Mortality Indicator (SHMI) and Care Quality Commis-

sion ratings. SHMI is calculated by taking the ratio between the actual number of patients who die in hospitals in a given trust and the number of patients that would be expected to die given patient demographics [108]. CQC ratings are awarded through continuous hospital level monitoring. The rating consists of four categories: inadequate, requires improvement, good and outstanding. These ratings are constructed based on five criteria: safety, effectiveness, care, responsiveness and leadership [109]. If TOXBASE were associated with better hospital performance, then it would be observed that those hospitals which use TOXBASE more to have a lower SHMI and that those hospitals which use TOXBASE more are more likely to be rated good or outstanding by the CQC.

6.2 Results

6.2.1 Admissions

Outlier Attributes

Figure 6.2 shows the funnel plot constructed for the rate of admissions due to drugs poisoning using the number of attendances due to poisoning as the size indicator required for the x-axis. This plot shows a fairly symmetric pattern of proportions admitted. The average (shown by the horizontal black line) is slightly below 0.5, which indicates that on average slightly less than half of all poisoning attendances are admitted, though some hospitals admit as many as 90% of their attendances, and some admit as little as 10% of their attendances.

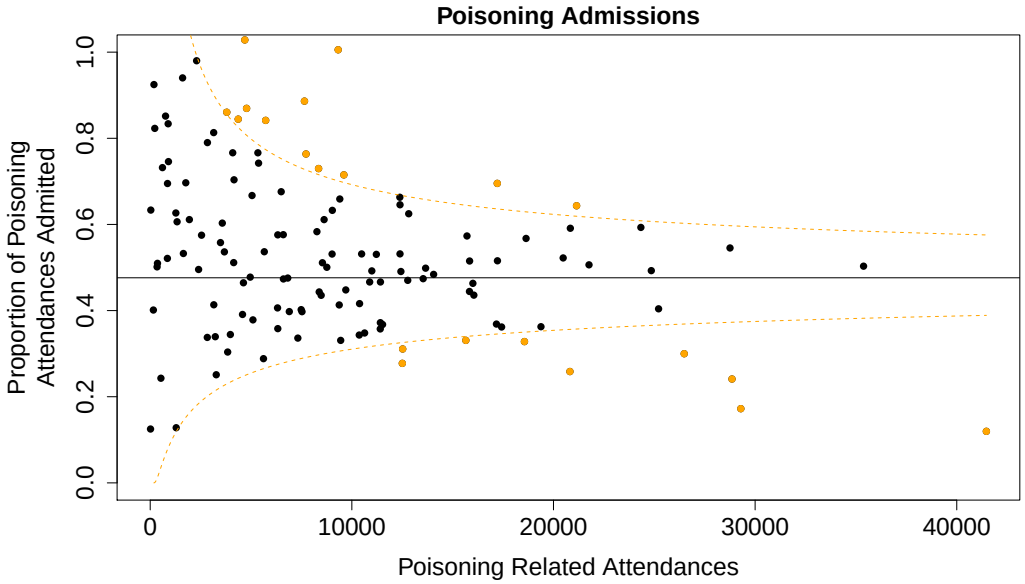


Figure 6.2: Funnel plot showing the distribution of admission rates with respect to the number of poisoning attendances. Unusual points have been highlighted in orange. The orange dashed lines represent the 25th and 75th percentiles

Those points in orange are those which are considered to be making admissions outside of the normal range. Out of 129 trusts, there are 21 trusts which are making admissions outside of the normal range; 12 are admitting patients more than would be expected, and the rest are admitting patients less than would be expected.

Figure 6.3 shows the variability in hospital size, as measured by number of overall attendances and number of beds, in those points within the normal range ("Normal"), those above the normal range of data ("High") and those below the normal range of data ("Low"). The plot on the left indicates that those hospitals which fall into the "Normal" category tend to have lower attendances on average than those hospitals in both the "High" and "Low" groups. However, attendances for those trusts within the normal range are skewed, with some "Normal" hospitals having large number of attendances. The "Low" group has slightly higher attendances on average compared to the "High" group.

These differences were assessed statistically via a linear regression on the log number of attendances against outlier status. The analysis of variance returned a p-value of 0.002, indicating that a significant difference exists between at least one pair of these categories. Plots of the difference estimates between each group are shown in the bottom left plot in Figure 6.3 along with confidence intervals corrected for multiple comparisons using Tukey's honest significant difference (Tukey's HSD). Noting that the effects are multiplicative, this demonstrates that there was a significant difference in attendances between the "Low" and "Normal" outlier categories, with the latter having attendances of around half of that in the former. The other comparisons were non-significant, though there is some suggestion that the "High" category has larger attendances than the "Normal" category on average.

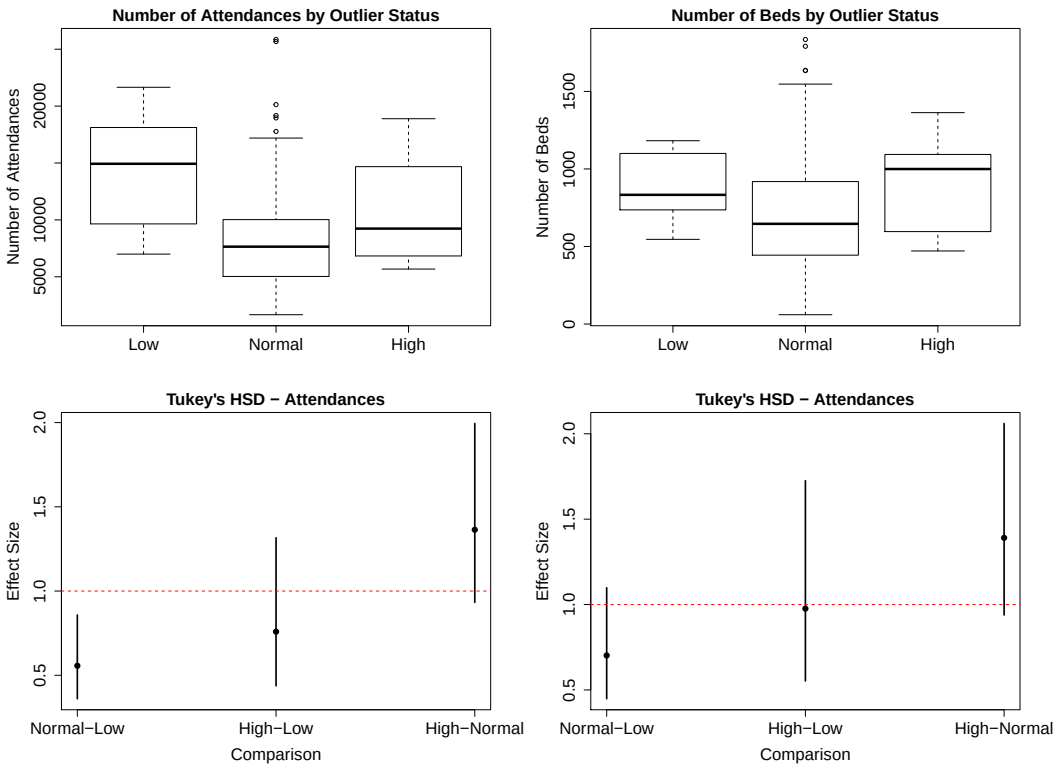


Figure 6.3: Figure showing the variation in hospital size for low admitting, normal admitting and high admitting hospitals. Hospital size was measured by the number of attendances (left) and number of hospital beds (right)

The number of beds (Figure 6.3, top right) in trusts in the "High" and "Low" groups tend to be marginally higher than those in the "Normal" group, with bed numbers being similar in the "High" and "Low" groups. The overall impression from these plots is that those hospitals which are unusual, in that they admit either more or less than expected, tend to be larger hospitals.

The formal assessment was carried out in the same way as for the attendances, though this time there were no significant differences at the 5% level (P = 0.0633). Since this P-value was still low, the Tukey HSD plot was produced as before (bottom right Figure 6.3). This shows close to significant differences in the comparisons of the "Normal" category with the "Low" and "High" categories, where the "Normal" category has a smaller number of beds on average compared to the other two groups.

The variability in the proportion of TOXBASE accesses for the most accessed pharmaceutical products in each of the three outlier status groups is shown in Figure 6.4. This indicates that there are some drug groups which are accessed consistently across the three groups, but also that there are some key differences in the way certain drug groups are accessed between low, normal and high admitting hospitals.

It can be seen that, on average, paracetamol, antidepressant and heroin pages are accessed consistently across the three hospital types, with slightly lower accesses made to antidepressants pages from those trusts with high admission rates. Access to antipsychotics pages is fairly consistent, though there appears to be a slight tendency for low admitting trusts to have higher accesses to these pages and for high admitting trusts to make lower use of these pages. This trend is more evident in accesses to non-opioid drugs of abuse pages, where the "Low" group tended to have a greater proportion of accessed to non-opioid drugs of abuse pages compared to the "Normal" group, and the "High" group accesses

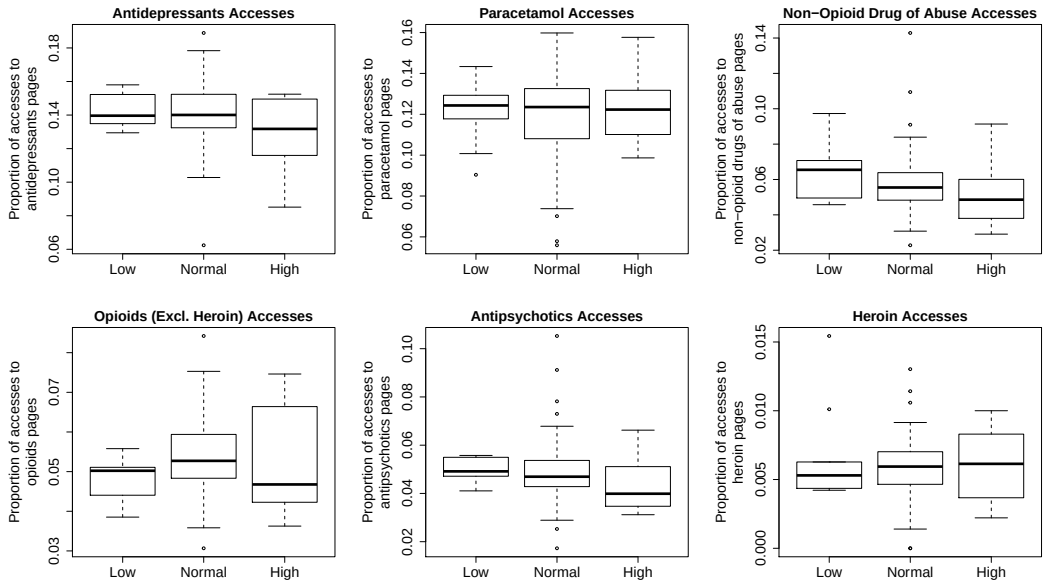


Figure 6.4: Figure showing the variation in access proportions for six commonly accessed drug groups for low, normal and high admitting hospitals. Top row from left to right: Antidepressants, Paracetamol, Non-Opioid Drugs of Abuse. Bottom row from left to right: Opioids (Excluding Heroin), Antipsychotics, Heroin.

non-opioid drugs of abuse pages at a lower rate than the "Normal" group. In contrast, the proportion of accesses to opioids pages was lower in the low admitting hospitals than in the hospitals admitting at a normal rate, for this category the variability in the proportion of accesses to opioids pages among high admitting trusts is large.

A similar process was followed for comparing TOXBASE use between the three outlier categories. The Tukey HSD plots can be seen in Figure 6.5. This indicates that, while none of the differences were significant at the 5% level, there are some differences which are close to being significant. In particular, the difference between the "High" category and the "Normal" category was very close to being significant, such that locations which had a high admission rate made fewer accesses to antidepressants pages as a proportion of their overall TOXBASE use.

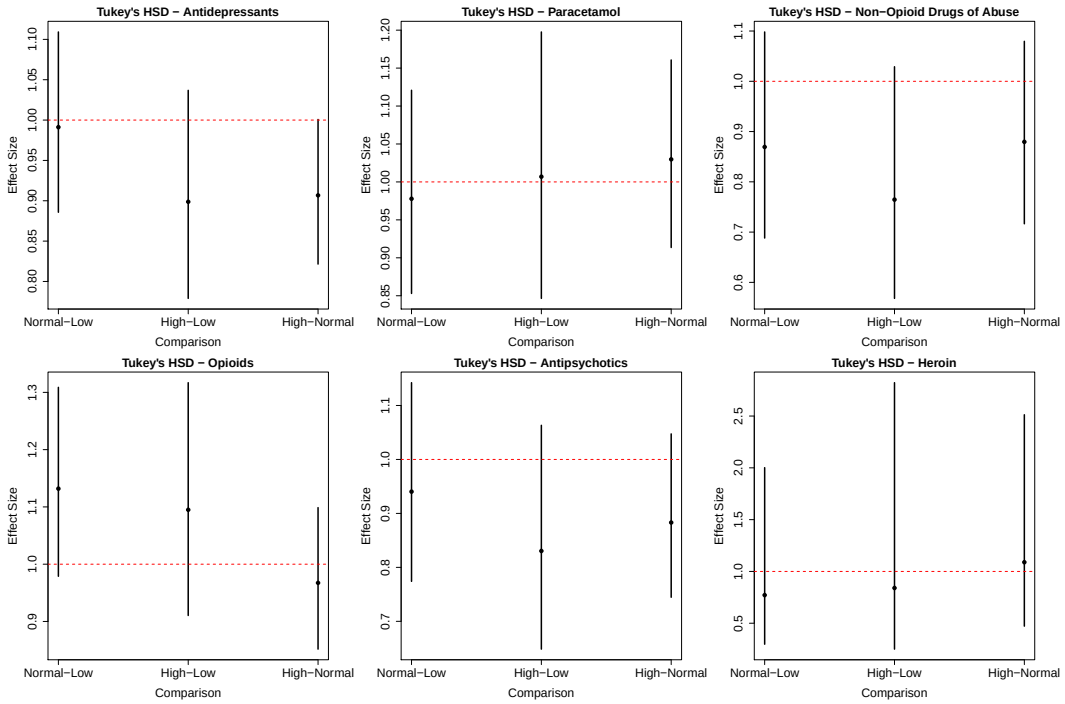


Figure 6.5: Plots showing estimated differences and confidence intervals for the comparison of the three outlier categories based on the proportion of accesses to specific groups of TOXBASE pages

There were a total of 33 university hospital trusts in the data. Table 6.1, shows how university status varied by outlier status. It appears that there may be a small tendency for more university hospitals to be outside of the normal range of data, however, there appears to be no distinction between university hospital and whether admissions are above or below the normal range of data. Further, a Fisher's exact test reveals that there is no statistically significant association between hospital type and outlier status (P=0.60)

Table 6.1: Table showing the number of university and non-university hospitals in low normal and high admitting hospitals.

	Low	Normal	High
Non-University Hospital	6	82	8
University Hospital	3	26	4

The outliers were, finally assessed by location. A map is shown in Figure 6.6 which displays the proportion of outliers in a given region. Regions which have an off-white colour have no outliers, while red means a large number of outliers with respect to the other regions.

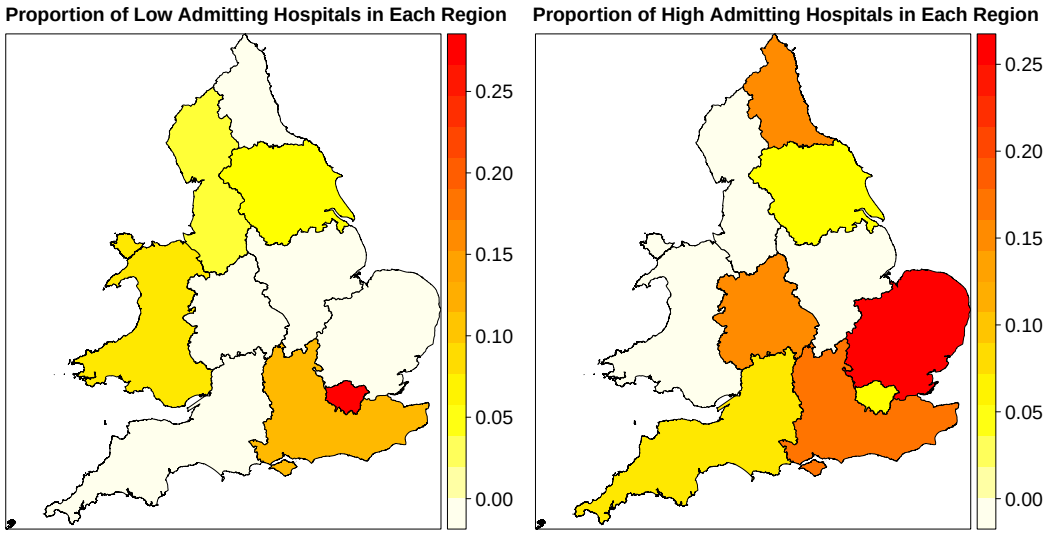


Figure 6.6: Plots showing the proportion of low (left) and high (right) admitting hospitals in each region. Region which are off-white have no low or high admitting hospitals, and red coloured regions have more low or high admitting hospitals compared to the other regions.

This shows that the South East and London tend to have more trusts with low admission rates, as do Wales, the North West and Yorkshire and the Humber, though to a lesser extent. Meanwhile, the East of England has the largest proportion of high admitting hospitals, with the South East, West Midlands and North East also having a relatively high proportion of high admitting hospitals. South West, Yorkshire and the Humber and London are also identified as having at least one trust with higher than expected admissions.

Outliers Through Time

The plots in Figure 6.7 show individual funnel plots for each year. Orange points indicate those hospitals which were found to be outliers in the previous analysis, rather than outliers in that year. This allows for the examination of how consistent these outliers were through time. Note that the number of orange points varies from plot to plot, due to the previously described data issues.

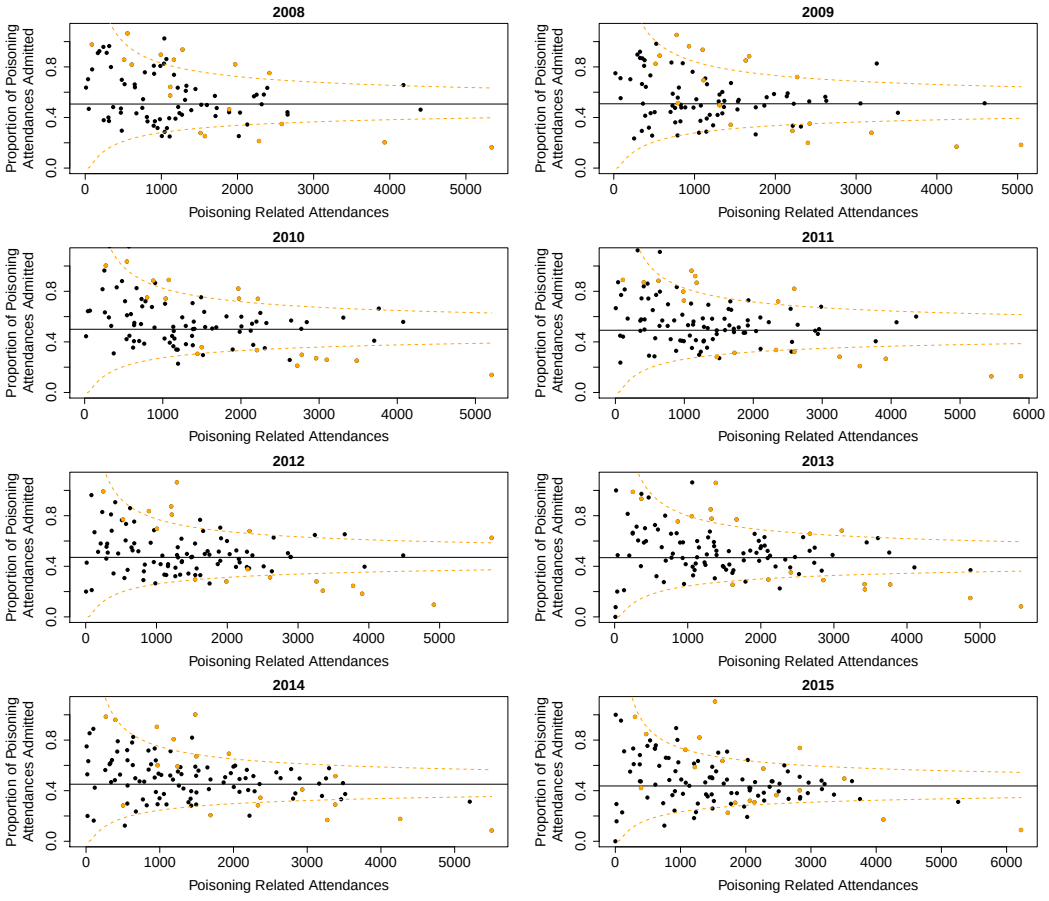


Figure 6.7: Funnel plots for each year, with orange points highlighting those locations which were identified as being outliers previously

This shows that a minority of those locations previously identified as outliers were within the normal range of admission rates, with the exception of 2015, where 12 out of the 19 points shown lie within the two boundary lines. This indicates that, while those that were identified as outliers are not always unusual,

they tend to be more unusual than other locations. Further, those points which were not identified as outliers, but which are outliers in at least one year appear to be fairly inconsistent, only appearing as an outlier in one or two years.

6.2.2 TOXBASE Accesses

Outlier Attributes

Figure 6.8 shows the funnel plot constructed for the rate of TOXBASE accesses per poisoning attendance. The size variable, as for admissions, represents the number of poisoning related hospital attendances. It is clear that the distribution of TOXBASE access rates are skewed towards larger rates, however the interest for this analysis lies in locations which do not use TOXBASE enough, and therefore only points that lie below the lower 75% confidence bound were considered as outliers.

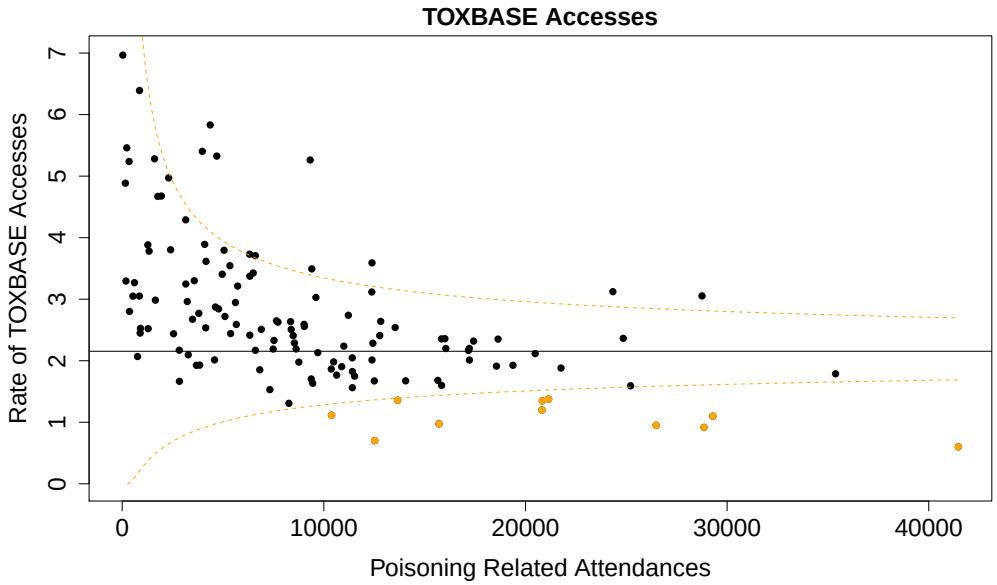


Figure 6.8: Funnel plot showing the distribution of TOXBASE access rates against the number of poisoning attendances. Hospitals with unusually low TOXBASE access rates are highlighted in orange. The orange dashed lines represent the 25th and 75th percentiles.

This plot indicates that there were slightly more than 2 TOXBASE accesses per poisoning attendance on average, with some accessing TOXBASE as many as 7 times per attendance and some locations accessing TOXBASE at a rate of less than once per attendance, meaning that in some hospitals there are attendances for which TOXBASE is not consulted.

In total, there are 11 hospitals or trusts which access TOXBASE at a lower than average rate. There are 7 locations which have both unusual admission rates and low TOXBASE access rates, and in 6 out of the 7 cases these were hospitals which admitted less than would be expected.

A similar analysis to that carried out in Section 6.2.1 was applied to the TOXBASE access data, this time with only two categories: those which accessed TOXBASE at a lower rate ("Low") and those which accessed TOXBASE at a normal or high rate ("Normal").

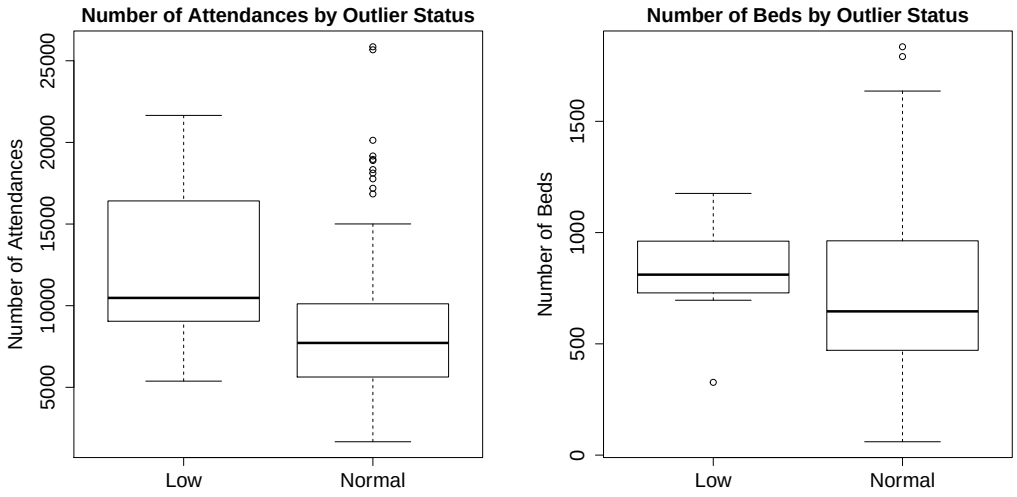


Figure 6.9: Boxplots showing the variability in hospital size for low accessing hospitals and normal or high accesses hospitals. Hospital size was measured using overall attendances (left) and number of hospital beds (right).

Figure 6.9 shows how hospital size, as measured by overall attendances and number of beds, varied depending on which category the particular trust falls into. This indicated that low accessing hospitals had slightly higher overall attendances

than "Normal" hospitals. However, the number of beds in each category was fairly similar, though those in the "Low" category had a smaller spread and lay slightly towards the top of the densest part of the distribution for "Normal" beds. This suggests that those hospitals which do not make as much use of TOXBASE tend to be busier, although not necessarily bigger hospitals.

Since there are only two categories present in this section, a t-test was carried out on the log number of attendances in order to make a formal comparison of the two groups. Comparing attendances between the "Low" and "Normal" categories indicated a significant difference (P=0.0097), with attendances in the "Normal" group being around 42% (95% CI: 12.1%, 73.0%) of that in the "Low" group. Looking at bed numbers, the test was just non-significant at the 5% level, with a p-value of 0.0522, however, this is very low and is very suggestive of a difference between bed numbers in low accessing hospitals and the other hospitals, such that low accessing hospitals have more beds on average than other hospitals.

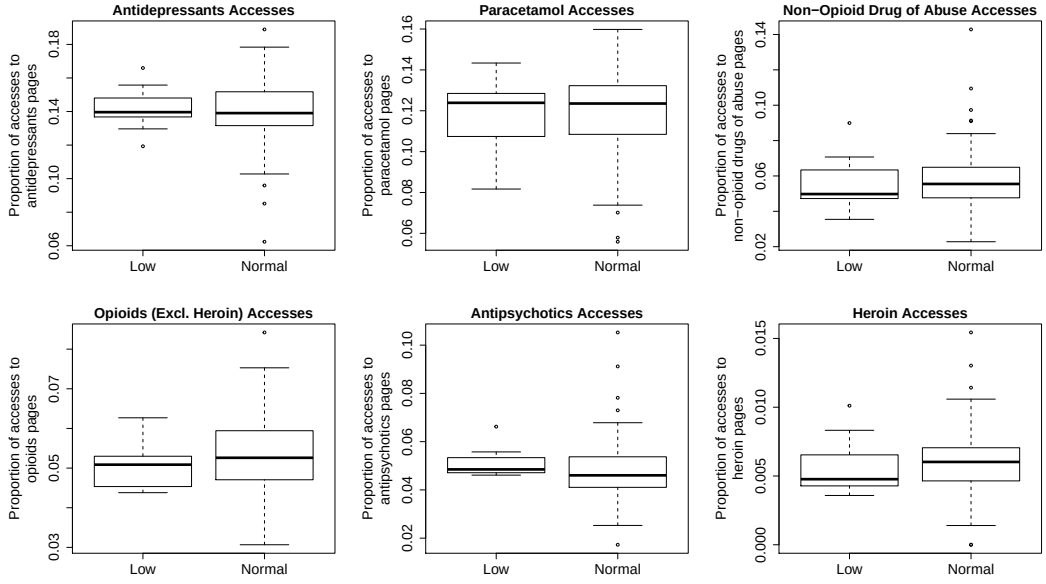


Figure 6.10: Boxplots showing the variability in TOXBASE access proportions to six commonly accessed drug groups. Top row from left to right: Antidepressants, Paracetamol, Non-Opioid Drugs of Abuse. Bottom row from left to right: Opioids (excluding heroin), Antipsychotics, Heroin

The TOXBASE access profile is shown in Figure 6.10. This indicates that those trusts which access TOXBASE at a low rate and those which access TOXBASE at a normal or high rate are very similar in terms of their accesses to specific drug groups. The only suggestion of a difference between the two groups is in accesses to opioids pages, where those with low TOXBASE access rates tend to make a marginally lower proportion of accesses to opioids pages, excluding heroin.

Table 6.2 shows the results from the hypothesis tests on each of the different types of pages examined. This shows that differences between the proportion of accesses to different page types were highly non-significant, with the exception of antipsychotics, which indicated that there were greater accesses to antipsychotics pages in those hospitals which were low users of TOXBASE overall.

Table 6.2: Table showing the P-values observed for the t-tests comparing TOXBASE access to specific groups of pages between low accessing and other hospitals

Drug Group	Difference	Confidence Interval	P-value
Antidepressants	1.023	0.932, 1.087	0.444
Paracetamol	0.998	0.887, 1.122	0.968
Non-Opioid Drugs of Abuse	0.989	0.833, 1.175	0.895
Opioids	0.947	0.867, 1.034	0.205
Antipsychotics	1.118	1.027, 1.217	0.013
Heroin	1.127	0.829, 1.530	0.436

Table 6.3 shows the distribution of university and non-university hospitals between low accessing and normal or high accessing hospitals. As with the admissions data, there is some evidence that there is a tendency for more university hospitals to be low accessing, however a Fisher Exact Test is not significant at the 5% level (P=0.47).

Table 6.3: Table showing the number of university and non-university hospitals for low accessing hospitals and normal or high accessing hospitals respectively.

	Low	Normal
Non-University Hospital	7	89
University Hospital	4	29

The map in Figure 6.11 indicates that there are more low accessing hospitals on the west, than in the east of the country. As with admissions, both London and the South East regions can be seen to have a high proportion of outliers relative to the rest of the country.

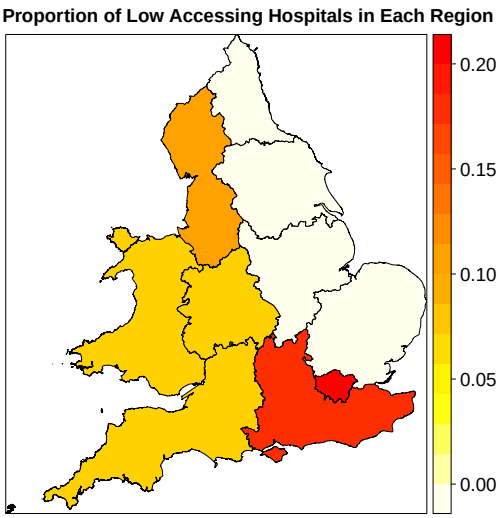


Figure 6.11: Maps showing the proportion of low accessing hospitals within each region. Region coloured in off-white have no low accessing hospitals and regions which are red in colour have a high proportion of low accessing hospitals with respect to the other regions.

TOXBASE use was also compared with SHMI and CQC ratings, in order to assess whether there was any correlation between TOXBASE use and hospital performance. These measures are unique to NHS England, and therefore the Welsh data have been omitted from this part of the analysis.

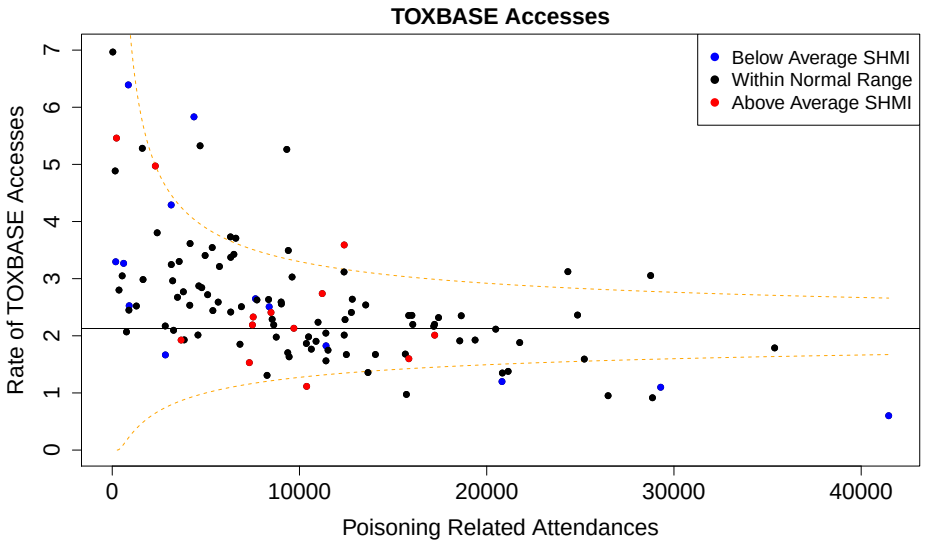


Figure 6.12: Funnel plot showing the distribution of the rate of TOXBASE accesses with poisoning attendances, coloured by SHMI

Figure 6.12 shows the funnel plot in Figure 6.8, but this time coloured by SHMI. SHMI is a relative index of mortality, black points are those with mortality rates within the normal range, green points are those whose SHMI lies below the normal range and red points are those whose mortality rate lies above the normal range. If there were association between SHMI and TOXBASE access, the coloured points would be clustered together. It is clear that this is not the case, indicating that TOXBASE use per attendance is not associated with SHMI.

The comparison of TOXBASE access and CQC rating is shown in Figure 6.13, and is similar to the previous figure. In this plot, points are coloured according to their CQC rating: black points have an "Inadequate" rating, dark blue points have a "Requires Improvement" rating, royal blue points have a "Good" rating and light blue points are rated "Outstanding".

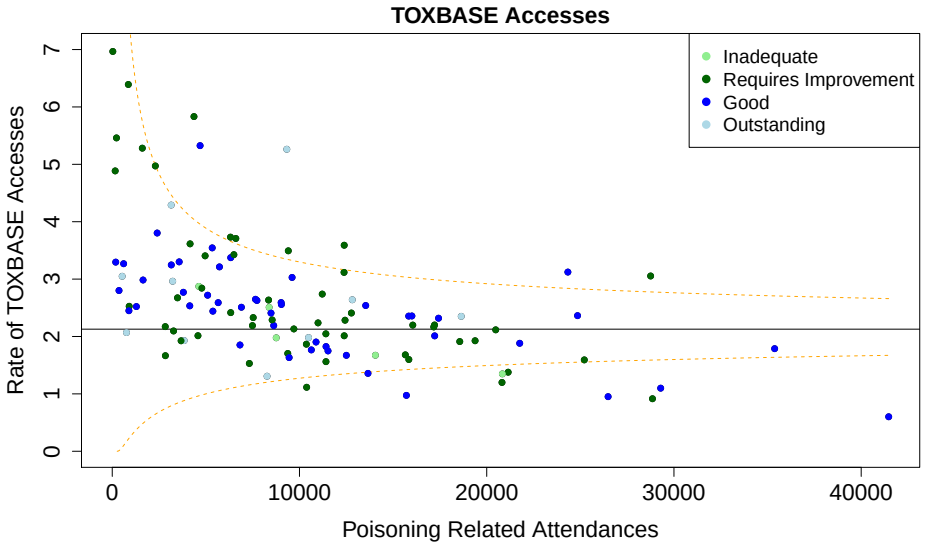


Figure 6.13: Funnel plot showing the distribution of the rate of TOXBASE accesses with poisoning attendances, coloured by CQC rating

As with Figure 6.12 if there were an association between TOXBASE use and CQC rating, clusters of similarly coloured points would be observed in Figure 6.13. As with SHMI, there does not appear to be any relationship between TOXBASE use and CQC rating.

Outliers Through Time

Individual funnel plots of TOXBASE accesses by year are shown in Figure 6.14, in order to assess how consistent those points identified as being unusual were through time.

The orange points indicate those points which were previously found to be unusual in their use of TOXBASE, and these are fairly consistently found to lie either outside of the boundaries, and if not they generally sit in the bottom half of the data. Those hospitals or trusts which were identified as unusual in their TOXBASE use appear to be more consistently unusual compared to those which were unusual in their admissions. Additionally, there are five years out of the

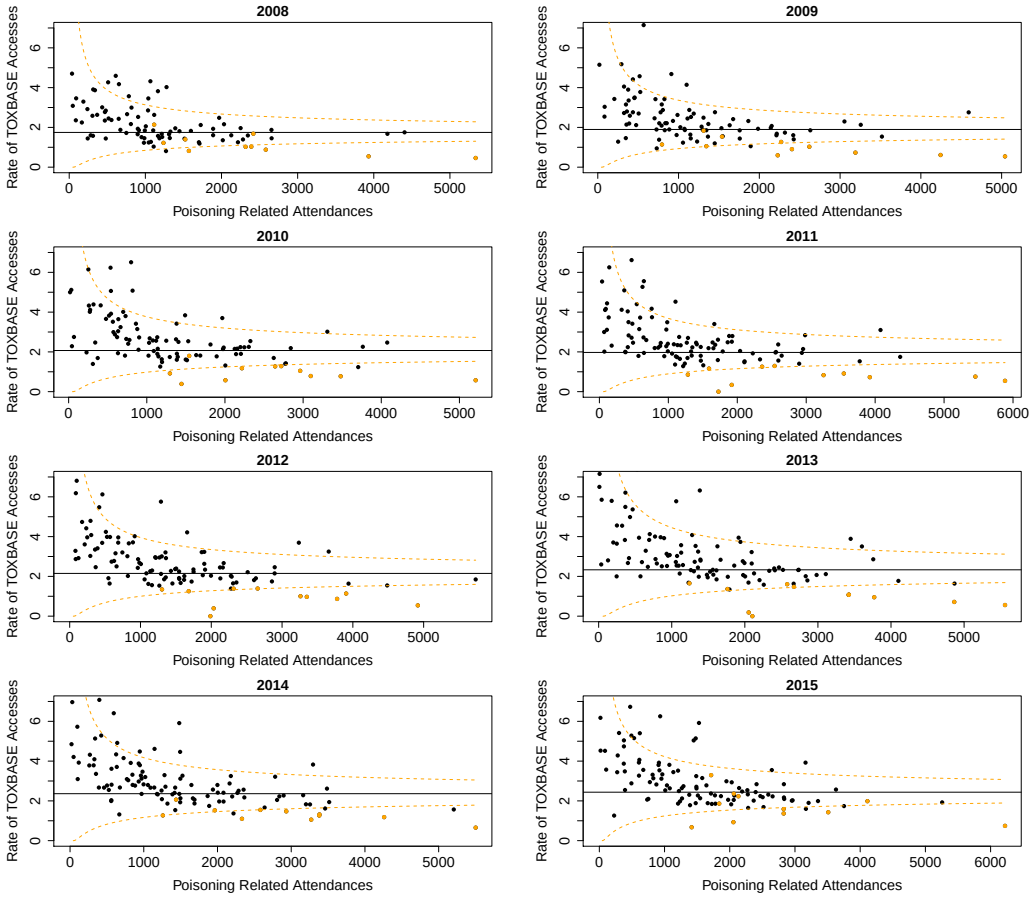


Figure 6.14: Funnel plots for each year, with orange points highlighting those locations which were identified as being outliers previously

eight where there is are additional outliers, though the location attributed to each of these outliers varies each time, with only one location appearing as an additional outlier in two years (2008 and 2014).

6.3 Discussion

The aim of this chapter was to identify unusual hospitals/trusts and to determine whether there were any common features of hospitals which were unusual, either in terms of admissions due to poisoning or with respect to their TOXBASE use.

Hospital size was considered as a feature, and there was some evidence to suggest that those trusts which were unusual in some way tended to be larger, both in terms of attendances and outlier status. This may reflect patterns in whether a given location is urban or rural, as urban hospitals tend to be larger, and therefore there may be many other outside factors which impact admission.

This location pattern was also reflected by regional differences in the number of unusual hospitals. London had a large number of trusts which were both low admitting and low accessing, suggesting that there may be some underlying protocol for poisoning treatment in London which is different from that in the rest of England and Wales. Meanwhile West Midlands had high admissions compared to the other regions, but was not unusual in its use of TOXBASE.

Since admission due to poisoning is predicated by both the severity and cause, it was important to examine how certain drug categories varied based on whether a trust was unusual or not. The results of this showed more differences between those hospitals which had unusual admissions than those which had unusual TOXBASE accesses.

A low proportion of accesses to opioids pages typically indicated hospitals with low admission rates, with a similar although less obvious trend in Heroin accesses. Opioids are highly addictive substances which are commonly abused. They are also used in pain management [110]. It is generally necessary to monitor patients who are undergoing treatment for an opioid overdose, and therefore where pages for opioids and heroin are less accessed, indicates fewer cases and hence fewer admissions [111].

Those hospitals which were low admitting also tended to have a higher proportion of accesses to non-opioid drugs of abuse pages, while those trust identified as high admitting tended to have lower proportional access to non-opioid drug of abuse pages. “Non-opioid drugs of abuse” is a blanket term which covers a range

of substances which are commonly abused. TOXBASE provides substance specific treatment paths, which may, with greater access, result in fewer unnecessary admissions for this type of poisoning.

A similar trend was found in accesses to antipsychotics pages, with low admitting hospitals having slightly more accesses than the other groups. High admitting hospitals had slightly lower accesses to antipsychotics pages. Antipsychotics, like non-opioid drugs of abuse, covers a wide range of prescribed drugs, which may be an indication that TOXBASE use may result in fewer unnecessary admissions.

There was no evidence to suggest that university hospitals admitted or used TOXBASE and a way that was different to other hospitals. This was surprising, as one might expect that medical students, or junior doctors, may be more rigorous in their patient treatment and therefore would make more use of TOXBASE. However, the tendency was for a greater proportion of low accessing locations to be university hospitals.

TOXBASE use was assessed to determine whether it could approximate hospital quality. This assessment indicated that there was no link between TOXBASE use and the two quality indicators SHMI and CQC rating. This is perhaps unsurprising, as poisoning accounts for around 1% of all hospital attendances, meaning that there is a large proportion of variability in these ratings that cannot be accounted for by TOXBASE use alone.

An additional assessment of the outliers involved examining funnel plots by year in order to assess whether those locations identified as being unusual were consistently unusual. One these plots, the majority of those locations identified as unusual in either TOXBASE use or admissions lay outside of the boundary lines. This indicated that those which had been previously identified were identified

because they were consistently unusual with respect to the rest of the hospitals and trust. This was found to be more true for unusual TOXBASE users than those locations which admitted at a higher or lower rate than expected.

This chapter gave some interesting insight in to hospital demographics and how they link with TOXBASE use and admissions. This has indicated a need for a more formal examination of the impact of accessing different drug groups on admission rates. It has also brought to light large regional variability in both TOXBASE use and admission rates, which will therefore be accounted for in the analysis in Chapter 7, where the association between access to TOXBASE and hospital admission due to drugs poisoning will be investigated.

Chapter 7

Impact of TOXBASE on the Treatment of Poisoned Patients

This chapter aims to answer the overall goal of the project, which is to determine whether there is any effect of TOXBASE use on patient care. The reason for this is that the main function of poisons information services is to improve the triage and care of poisoned patients. One would therefore expect that use of TOXBASE would have some effect on emergency admission due to poisoning. In fact it has been shown that poisons information used at primary care reduces poisoning attendances at emergency departments [112]. This study used TOXBASE to give users listed as "GP" the opportunity to answer questions both prior to accessing TOXBASE and after accessing TOXBASE, in order to assess whether there was a change in the recommendation. This analysis differs, in that the aim is to use routinely collected administrative data on admissions and attendances to assess, the effect of poisons information services within the emergency department.

The chapter focusses on data from England and Wales only, since the Scottish data had no information about poisoning specific attendances, and it was shown in Chapter 5 that it would not be appropriate to assume that poisoning attendances occur as a constant rate of overall attendances throughout the year, the data used in this chapter are the same data that were used in Chapter 5.

7.1 Methods

The first step in this analysis was to fit models which would evaluate the overall trends in admissions based on TOXBASE access and attendances. Initially, a model was fit which contained three terms: a main effect for the number of poisoning attendances, a main effect for the number of accesses to TOXBASE and an interaction between the two. The model also incorporated a random effect for site, as shown in Equation 7.1

$$E(\log(y_{ij})) = \beta_0 + \beta_1 x_{ij} + \beta_2 a_{ij} + \beta_3 x_{ij} a_{ij} + \gamma_i \tag{7.1}$$

Here, y_{ij} was the number of admissions due to drugs poisoning at time j in location i. TOXBASE accesses are denoted using the variable x and poisoning attendances were represented by the variable a. The random effect for site is denoted by γ_i.

Using this model, it was difficult to make sensible inference about the associations between the three model components. It was decided that additional information had to be incorporated into the model in order to separate out spurious association from true associations. To that end, further models which incorporated temporal terms to capture the features which were found to exist in Chapter 5. These trends were incorporated in a variety of ways, though each of these models seemed to be overcomplicated, which led to fitting the model

outlined in Equation 7.2. Instead of modelling the number of admissions due to drugs poisoning using TOXBASE accesses and poisoning related attendances, this model aimed to assess changes in the rate of admissions due to drugs poisoning per poisoning attendance. This took into account the rate of TOXBASE access per poisoning attendance and hospital size, which was measured using overall attendances.

$$E(\log(y_{ij})) = \log(a_{ij}) + \beta_0 + \beta_1 x_{ij} + \beta_2 s_{ij} + \beta_3 x_{ij} s_{ij} + f(t_j) + g(m_j) + \gamma_i \quad (7.2)$$

In this equation, $E(\log(y_{ij}))$ is the expected log number of admissions due to drug poisoning and $\log(a_{ij})$ is the log number of poisoning related attendances for trust i at time j. The two independent variables x_{ij} and s_{ij} are the corresponding rate of TOXBASE accesses per poisoning attendance and overall number of attendances respectively. The beta estimates will give information on how admissions due to drug poisoning varies with attendance and TOXBASE use; positive values of β_1 or β_2 would indicate that admissions increase with attendances and TOXBASE use, while negative values would indicate a decrease in admissions with increasing attendances and TOXBASE use. The coefficient β_3 is more complex in its interpretation, and can be interpreted as the change in admissions with attendances for different levels of TOXBASE use. Finally, the model contains two smooth components $f(t_j)$ and $g(m_j)$ which are the long term and seasonal trends respectively, where t_j is an overall time indicator and m_j is a month indicator. As previously, γ_i denotes the random effect for location.

The data used for this analysis are the combination of the English and Welsh data shown in Chapter 5, with observations on trust, month, year, number of TOXBASE accesses (both overall and to pharmaceutical pages), number of admissions and number of attendances (both overall and due to poisoning).

The model described in Equation 7.2 was compared to a simpler model which excluded the interaction (Equation 7.3) in order to assess how important the interaction term was in the analysis. The fits of these two models were compared using the deviance explained.

$$E(\log(y_{ij})) = \log(a_{ij}) + \beta_0 + \beta_1 x_{ij} + \beta_2 s_{ij} + f(t_j) + g(m_j) \tag{7.3}$$

In addition to assessing the trends overall, the association between admissions and six commonly accessed drug groups was examined. These are the same drug groups discussed in Chapter 6: antidepressants, paracetamol, non-opioid drugs of abuse, opioids (including medications but excluding heroin), antipsychotics and heroin. These relationships were evaluated using simple generalised linear models of the form shown in 7.4.

$$E(\log(y_{ij})) = \log a_{ij} + \beta_0 + \beta_1 x_{ij} \tag{7.4}$$

For these models, $E(\log(y_{ij}))$ represents the expected log number of admissions due to drugs poisoning and a_{ij} is the number of attendances due to poisoning in trust i at time j as before. The independent variable x_{ij} is the proportion of accesses made to pages in TOXBASE concerning the specific product type of interest in trust i at time j.

Spatial autocorrelation in the trends was examined prior to analysis for the rate of admissions due to drug poisoning and the rate of TOXBASE accesses per poisoning attendance respectively. This was done using the sample variogram

and Monte Carlo envelopes. The sample variogram is calculated by splitting the maximum distance into bins and for each bin calculating the average semivariance between the pair which fall into each bin as below.

$$\hat{\gamma}(h_{\pm\delta}) = \frac{1}{2N(h_{\pm\delta})} \sum_{i,j \in N(h_{\pm\delta})} (z(s_i) - z(s_j))^2$$

In the above, $h_{\pm\delta}$ denotes the range of distances covered by the bin and $N(h_{\pm\delta})$ represents the number of observations within the bin. Then s_i and s_j are a pair of points which lie within the bin and $z(s_i)$ and $z(s_j)$ are the observations which correspond to those locations.

Since each point on the sample variogram corresponds to one of the bins, and each point represents an average based on pairs of locations. The variogram is therefore more variable at larger distances where there are fewer pairs contributing to the average semivariance.

Monte Carlo envelopes are constructed using the same intuition behind the test for Moran's I outlined in Chapter 2. The data are permuted in order to produce a set of data which are spatially random, and the sample variogram is calculated. This is repeated a large number of times, resulting in a range of sample variograms which correspond to spatial randomness. Then, assuming $\alpha\%$ significance, the $\alpha/2$ and $1-\alpha/2$ percentiles are computed for each bin. The series of each of these percentiles are then the lower and upper envelopes respectively.

If the points in the sample variogram lie outside of these envelopes, particularly at the beginning of the variogram, then there is evidence to suggest that there is spatial autocorrelation present in the data.

7.2 Results

The overall rates of admission and TOXBASE accesses per poisoning attendances were examined by region, as shown in Figure 7.1.

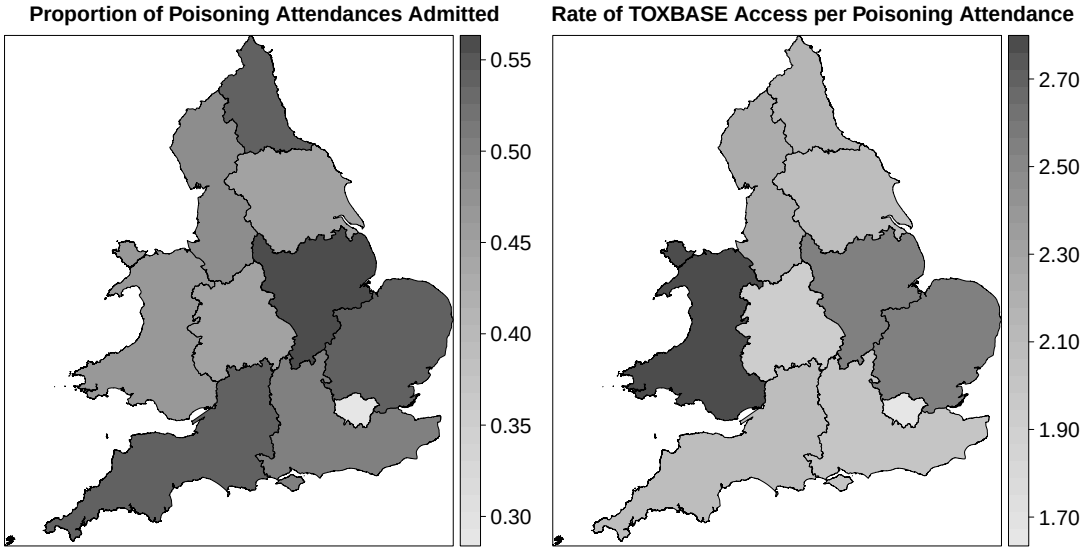


Figure 7.1: Maps of the proportion of poisoning attendances resulting in an admissions (left) and the rate of TOXBASE access per poisoning attendance (right)

The key feature that can be seen in these maps is that London appears to be unusual in its admissions. In fact, the proportion of poisoning attendances admitted was found to be 0.284 for London, compared to 0.451 in Yorkshire and the Humber, which had the lowest admission proportion except from London.

In terms of TOXBASE use, London was also low with respect to the rest of the country, with a rate of TOXBASE accesses of 1.634, the West Midlands was closest to London in terms of TOXBASE use, with a rate of access of 1.923. Wales appeared to have comparatively high TOXBASE use, with a rate of 2.799 accesses per poisoning attendance. For reference, the next highest rate of access was in the East Midlands with 2.573 accesses per attendance. Since London was unusual in both admissions and accesses, these findings led to the decision to

analyse the data for London hospitals and other England and Wales hospitals (excluding London) separately, in addition to constructing an overall England and Wales model.

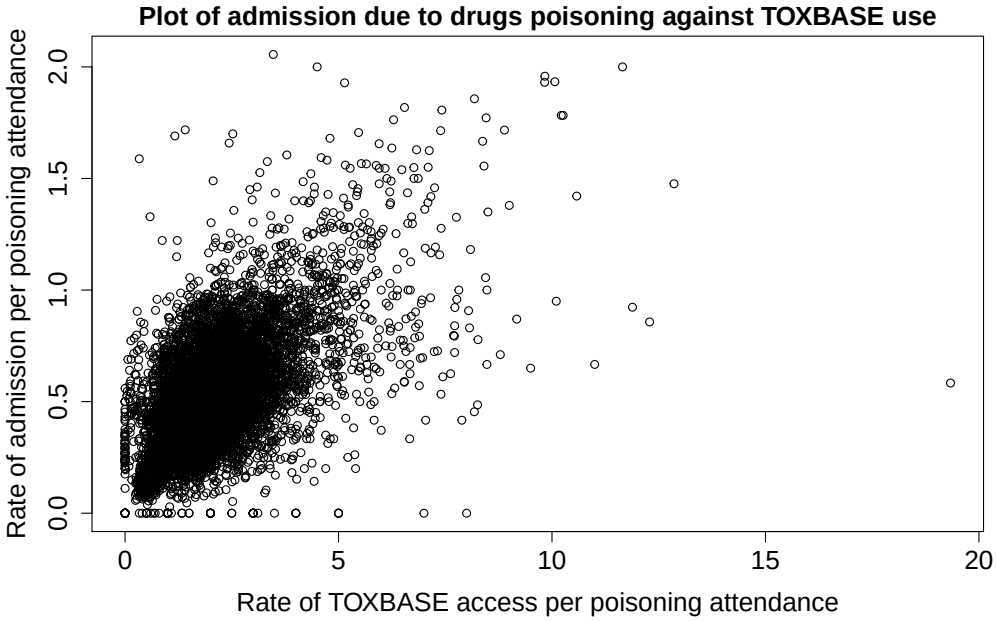


Figure 7.2: Plot of the rate of admission due to drug poisoning per poisoning attendance against the rate of TOXBASE access per poisoning attendance

The rate of admissions due to drug poisoning was plotted against the rate of access to TOXBASE, as shown in Figure 7.2. This shows a positive correlation between the rate of admission due to drugs poisoning and the rate of TOXBASE access per poisoning attendance. This is a surprising result, due to the fact that the generalised additive models in Chapter 5 indicated that the rate of TOXBASE access per poisoning attendance was increasing, while the rate of admission due to drug poisoning per poisoning attendance was decreasing.

This was explored further, and the data were aggregated across time, to produce one estimate for the rate of admission and the rate of TOXBASE use for each trust. These data were plotted and are shown in Figure 7.3. These data show the same positive correlation as previously, indicating that hospitals or trusts with higher rates of TOXBASE access tend to have a higher rate of admissions.

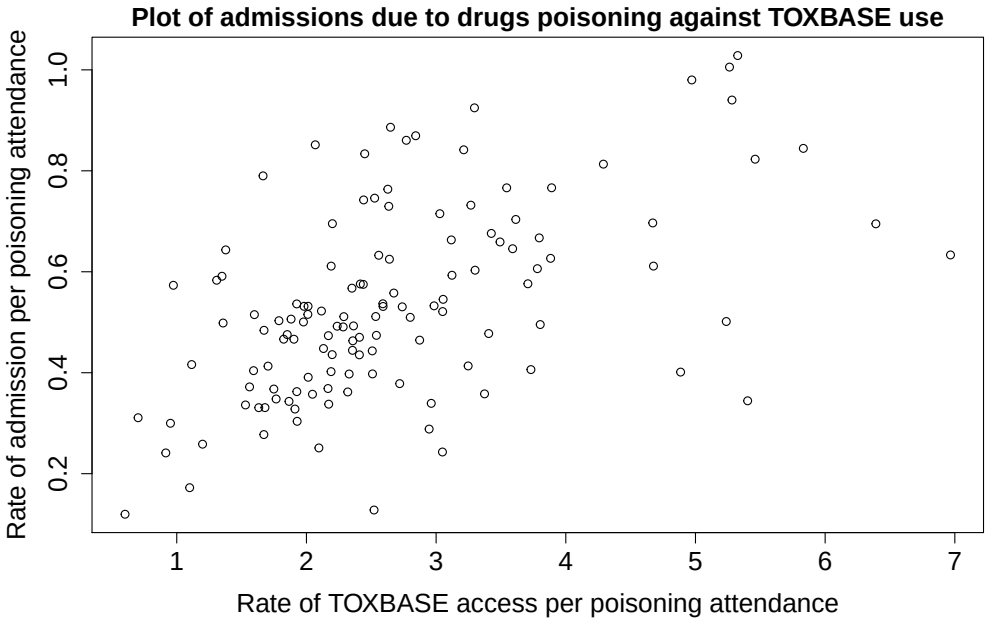


Figure 7.3: Plot of the rate of admission due to drug poisoning per poisoning attendance against the rate of TOXBASE access per poisoning attendance. Observations are aggregated so that one point is one hospital

The data were then aggregated across hospital, in order to produce one estimate of the rate of admission and the rate of TOXBASE access per poisoning attendance for each month in the data. These data were once again plotted (Figure 7.4, left), and a negative correlation was found. This indicates that there is negative temporal correlation between the rate of access to TOXBASE and the rate of admission due to drugs poisoning, which corresponds to the findings in Chapter 5, where TOXBASE accesses were found to increase over time, and admission rates were found to decrease.

In order to examine this further, the plot was altered, so that each year was coloured differently, and had a different plotting symbol (Figure 7.4, right). It can be seen from this that points belonging to earlier years tend to be located on the left hand side of this plot, while points corresponding to later years fall more on the right hand side of the plot. The overall negative correlation can still be

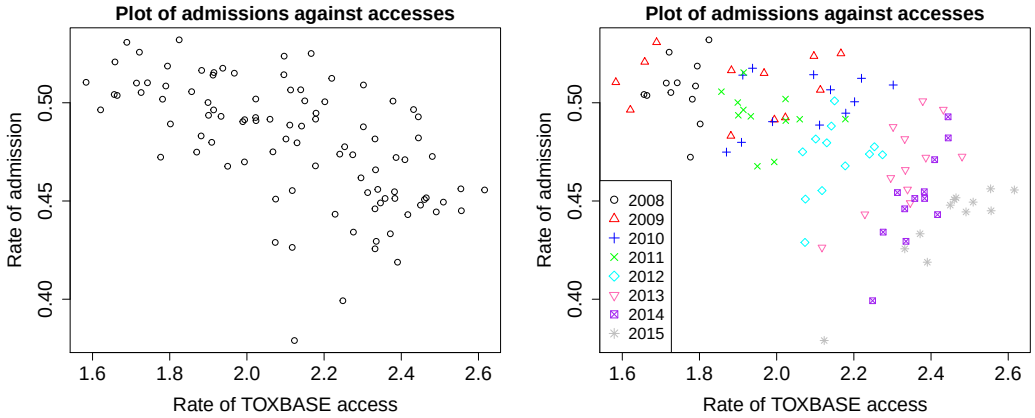


Figure 7.4: Plot of the rate of admission due to drug poisoning per poisoning attendance against the rate of TOXBASE access per poisoning attendance. Observations are aggregated so that one point is one month

seen in this plot, however it also demonstrates a positive correlation within each year. This suggests that, while temporally there is negative correlation, overall the rate of admissions tends to increase as TOXBASE use increases.

An assessment of spatial autocorrelation was also carried out as part of this initial examination of the data. This was done using the sample variogram and Monte Carlo envelopes as described previously. These have been plotted for both the rate of admission due to drugs poisoning and the rate of TOXBASE access and are shown in Figure 7.5.

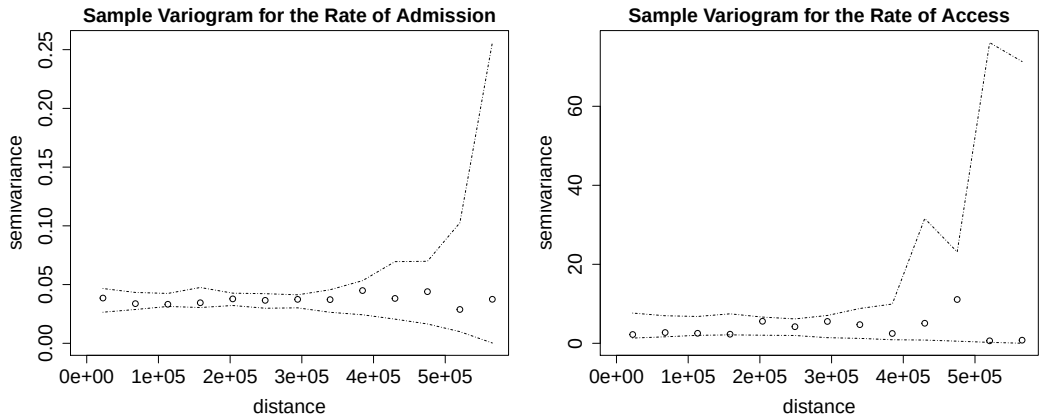


Figure 7.5: Plots of the sample variogram (points) and the corresponding Monte Carlo envelopes (dashed lines) for the rate of admissions and TOXBASE access per poisoning attendance

In both of these plots, the points lie within the envelopes, which indicates that there are no concerns with spatial dependence in these data.

7.2.1 Interaction Model

The first model fit was that outlined in Equation 7.1, which aimed to assess the association between admissions due to drugs poisoning, and poisoning attendances and TOXBASE access respectively. This produced the results shown in Table 7.1.

Table 7.1: Table showing the coefficients for the First model fit to admissions in England & Wales overall, England & Wales excluding London and London (standard errors for the estimates are given in parentheses)

	England & Wales overall	England & Wales excl. London	London
Attendances	1.82×10^{-3}	2.24×10^{-3}	1.34×10^{-3}
	(8.23×10^{-5})	(9.25×10^{-5})	(2.42×10^{-4})
Accesses	9.56×10^{-4}	9.81×10^{-4}	1.09×10^{-3}
	(3.82×10^{-5})	(3.91×10^{-5})	(1.56×10^{-4})
Attendances×Accesses	-1.34×10^{-6}	-1.51×10^{-6}	-2.71×10^{-6}
	(1.25×10^{-7})	(1.27×10^{-7})	(5.85×10^{-7})

The positive coefficients in the first two rows of Table 7.1 indicate that the number of admissions due to drugs poisoning increases with both the number of poisoning attendances and the number of TOXBASE accesses, which makes sense as these both capture an element of hospital size, where larger hospitals will admit more patients. The negative coefficient for the interaction term indicates that this effect levels off, so that as attendances and accesses increase, the growth in admissions reduces. This initially seemed to make sense and indicated a potential positive effect of the use of TOXBASE on admissions. However, thinking about

this in more depth, this negative effect may be confounded by additional factors. In particular hospital capacity, as discussed in Chapter 5 relating to the temporal trends, may be a contributing factor to the interaction coefficient, with larger attendances meaning more admissions overall and therefore less ability to admit borderline poisoning cases.

In an attempt to account for this, the temporal trends were incorporated as in Equation 7.2. The interaction term was initially included to allow for the investigation of whether the impact of TOXBASE use varied depending on the size of the hospital, as measured by the number of overall poisoning attendances. The coefficient estimates for models fit to England and Wales overall, England and Wales excluding London and London are shown in Table 7.2.

Table 7.2: Table showing the coefficients for the interaction model fit to admissions in England & Wales overall, England & Wales excluding London and London (standard errors for the estimates are given in parentheses)

	England & Wales overall	England & Wales excl. London	London
Attendances	-2.570×10^{-5}	-2.069×10^{-5}	-3.453×10^{-5}
	(1.57×10^{-6})	(1.92×10^{-6})	(3.55×10^{-6})
Accesses	0.1107	0.1198	0.0565
	(4.48×10^{-3})	(4.58×10^{-3})	(2.04×10^{-2})
Attendances×Accesses	4.073×10^{-6}	2.931×10^{-6}	8.405×10^{-6}
	(4.52×10^{-7})	(4.76×10^{-7})	(1.77×10^{-6})

The coefficient for the interaction term was found to be positive in all three of the cases considered, implying that if a large hospital were to increase its rate of TOXBASE use, the predicted increase in the admission rate is greater than that

for a small hospital which increased its rate of access to TOXBASE by the same amount. However, the size of this effect means that this has very little influence, particularly when compared to the estimated coefficient for TOXBASE use.

Since the interaction effect was found to be small, a model which did not contain this interaction effect was also fit. This new model had a deviance explained of 75.2%, this was not a substantial change when compared to the deviance explained for the previous model which was 75.4%. This supports the evidence shown in Table 7.2, that the interaction term does not provide a great deal of information about the relationship between TOXBASE use and admission due to drug poisoning. The coefficient estimates for the model without the interaction term is shown in Table 7.3.

Table 7.3: Table showing the coefficients for the model with no interaction fit to admissions in England & Wales overall, England & Wales excluding London and London (standard errors for the estimates are given in parentheses)

	England & Wales overall	England & Wales excl. London	London
Attendances	-1.719×10^{-5}	-1.354×10^{-5}	-2.185×10^{-5}
	(1.28×10^{-6})	(1.55×10^{-6})	(2.49×10^{-6})
Accesses	0.1447	0.1435	0.1438
	(0.0025)	(2.58×10^{-3})	(9.14×10^{-3})

The coefficient estimates in Tables 7.3 suggest a decrease in poisoning admissions with larger hospitals. This decrease is such that, for every 100 additional attendances (for any reason), the rate of poisoning admissions per poisoning attendance would decrease by around 0.2%. These results also show a tendency for admission rates to increase with increasing TOXBASE use, so that for every additional access made to TOXBASE, the rate of admission due to poisoning per poisoning attendance increases by 15.6% (95% CI: 15.0%, 16.1%). The size of this effect was similar across all three of the scenarios considered.

The seasonal and long term trends estimated by this model are shown in Figure 7.6. This shows trends which are similar to those seen for the rate of admission due to drugs poisoning in England in Chapter 5, with a peak in admission rates in April, and an overall decreasing trend the rate of admission between 2008 and 2015.

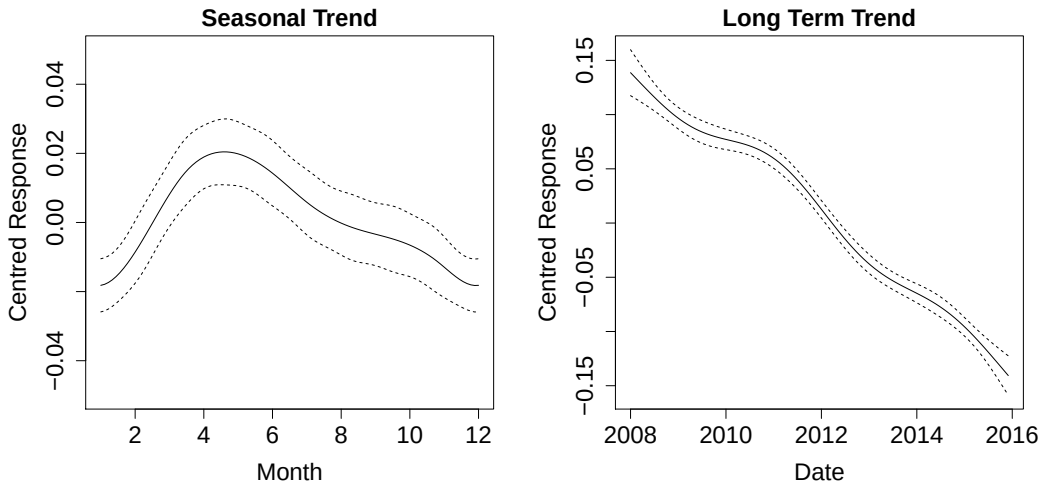


Figure 7.6: Plots showing the estimated seasonal (left) and long term (right) trends estimated by the model in Table 7.3 for England and Wales overall

This analysis does suggest some link between TOXBASE accesses and admissions, although it is difficult from this to determine whether TOXBASE has an impact on patient decisions or whether it is simply the case that more TOXBASE accesses indicate more complex poisoning cases, which are more likely to require admission. In order to examine the relationship between TOXBASE use and admission due to drug poisoning in more detail, a drug group specific analysis was conducted.

7.2.2 Drug Group Specific Analysis

The regional variation in the proportion of accesses to each category can be seen in Figure 7.7. This indicates that there are differences in the way in which TOXBASE is used across England and Wales.

Specifically, the South West region of England seems to have a comparatively high proportion of accesses to both antidepressants and paracetamol, which are the two most commonly accessed drug groups, but appears to be consistent with the rest of the country in accessing the other drug groups. There appears to be a decreasing trend in the proportion of accesses to non-opioid drugs of abuse pages moving from the north of England to the south. Additionally there is a decreasing trend in the proportion of accesses to opioids pages from the north east of England to the south west.

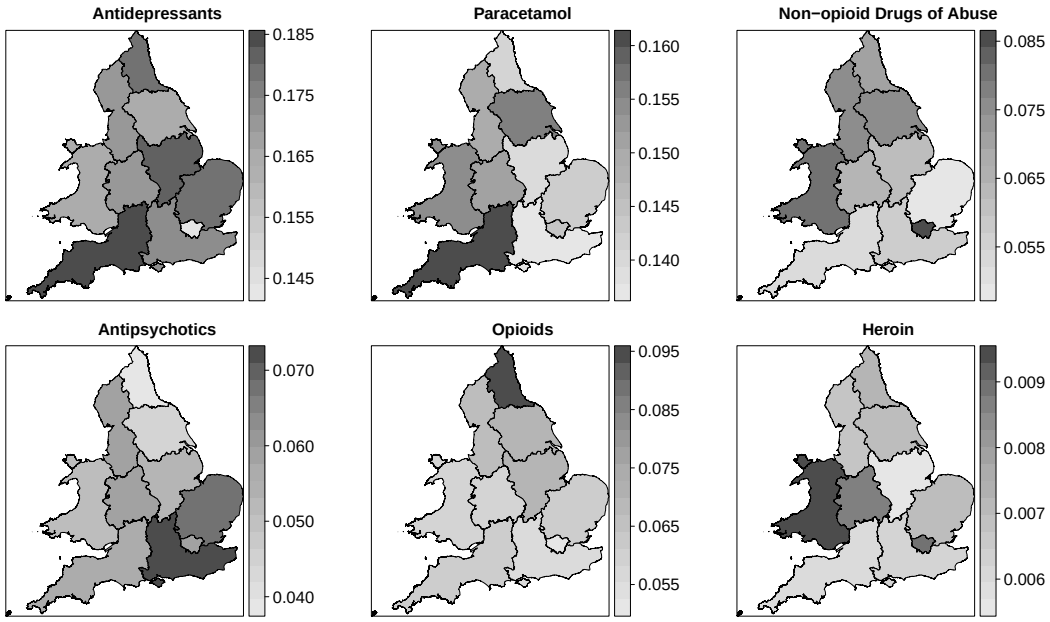


Figure 7.7: Maps showing the proportion of TOXBASE accesses made to each category of pages in the various regions in England

London appears to be very different in its usage of TOXBASE compared to the rest of England and Wales, as was found when looking at TOXBASE usage as a whole. In particular, the proportion of accesses to antidepressants pages is low in comparison to the rest of the country. Accesses to pages relating to non-opioid drugs of abuse, and heroin were higher in London compared to the other regions in England and Wales.

Figure 7.7 also represents the proportion of accesses accounted for by each of the drug groups under consideration. Antidepressants and Paracetamol represent the largest proportions of accesses, each accounting for approximately 15-16% of pharmaceuticals accesses. Non-opioid drugs of abuse represented 7% of accesses, while antipsychotics and opioids each accounted for around 6% of accesses. Heroin represented a relatively small proportion of all accesses at just under 1% of accesses to pharmaceuticals pages.

Due to sparseness of data, it was not possible to obtain information on how many admissions and attendances were attributed to each specific group under assessment. Therefore this section focusses on how the number of drugs admissions varies as a proportion of poisoning attendances with respect to the proportion of accesses made to each specific drug group.

Univariate models of the form described in Equation 7.4 were fit, relating the proportion of admissions to the proportion of accesses made to pages in each of the categories assessed, the results of which are presented in Table 7.4. This shows the estimated change in the proportion of admissions with an increase of 0.01 on the proportion of accesses to each type of page.

These results highlight three categories as being of particular interest. There appears to be a significant association between paracetamol accesses and admissions in all three of the area splits considered, where an increase in the proportion of paracetamol pages corresponds to an average decrease in the proportion of ad-

Table 7.4: The estimated percentage change in the admission rate per attendance for every increase of 0.01 in the proportion of accesses to pages within specific drug categories. Effects which were significant at the 5% level have been italicised

	England & Wales overall	England & Wales excl. London	London
Antidepressants	+0.01% (-0.11%, +0.14%)	+0.04% (-0.09%, +0.17%)	-0.21% (-0.71%, +0.30%)
Paracetamol	*-0.11% (-0.23%, -0.01%)*	*-0.18% (-0.30%, -0.06%)*	*+0.52% (+0.03%,+1.01%)*
Non-opioid Drugs of Abuse	*-0.15% (-0.29%, -0.01%)*	-0.04% (-0.19%, +0.10%)	*-1.02% (-1.53%, -0.50%)*
Antipsychotics	+0.00% (-0.19%, +0.20%)	+0.07% (-0.14%, +0.28%)	-0.49% (-1.21%, +0.24%)
Opioids	-0.02% (-0.22%, +0.17%)	+0.03% (-0.17%, +0.23%)	-0.59% (-1.14%, +0.23%)
Heroin	+0.44% (-0.15%, +1.03%)	+0.16% (-0.45%, +0.78%)	*+2.03% (+0.11%,+3.99%)*

missions of 0.11% (95% CI: 0.01%, 0.23%) in England and Wales overall and 0.18% (95% CI: 0.06%, 0.30%) when London is excluded from the analysis. In contrast, there is an estimated increase of 0.52% (95% CI: 0.03%, 1.01%) in the proportion of admissions with an increase in the proportion of accesses to paracetamol pages when London is considered on its own.

There is also a significant relationship between accesses to non-opioid drugs of abuse pages and admissions in London, which is likely influencing the relationship in England and Wales overall. This relationship is such that an increase in accesses to this category of pages indicates a, estimated decrease of 1.02% (95% CI: 0.50%, 1.53%) in the proportion of admissions in London.

There was a significant estimated relationship between access to heroin pages and admission in London only, where an increase in the proportion of accesses to pages related to heroin corresponds to an increase in the proportion of admissions due to drugs poisoning.

7.3 Discussion

The goal of this chapter was to provide evidence of a link between TOXBASE access and patient management decisions, with particular focus on whether TOXBASE use influences the decision to admit versus not admit a poisoned patient.

The first result of note was that trusts located in London seemed to have lower admissions and lower TOXBASE accesses when compared to the rest of the UK. This may be due to the fact that London was previously served by a poisons centre that was not affiliated with TOXBASE, a fact which changed when the NHS was restructured in 2005. However, since London would have previously had other sources of poisons information, it may be that these are still utilised by some practitioners.

The results of the analysis do indeed suggest that there is a link between TOXBASE use and admissions, with admissions generally increasing with increased usage of TOXBASE. However, this may be caused by greater admissions indicating a greater proportion of severe or complex cases of poisoning, which in turn would naturally require more use of TOXBASE, particularly where multiple teams are involved in the treatment of the patient.

In an attempt to assess the relationship between TOXBASE accesses and admissions further, TOXBASE access was separated into six distinct categories of pharmaceutical in order to approximate the case mix. Each of these was included in a univariate model to determine whether there were differences in

the estimated relationship between admissions and TOXBASE use for each drug group. The analysis found that there were three key groups which had significant association with admissions.

The first of these was paracetamol, where an increase in the proportion of accesses corresponded to a decrease in the proportion of admissions due to drugs poisoning in England and Wales overall, but an increase in London. This effect may be due to the fact that London has its own poison management system, and therefore do not use TOXBASE except in particularly difficult cases which are likely to require admission.

Interestingly, greater access to non-opioid drugs of abuse pages in London indicated a statistically significant decrease in the proportion of patients admitted due to drugs poisoning, an effect which was found to be non-significant in England and Wales excluding London. Finally, greater access to heroin pages in London corresponded to an estimated increase in the proportion of admissions, an effect which was not significant in England and Wales excluding London.

Considering the estimated effects as a whole, both positive and negative effects were estimated which gives some indication that TOXBASE does impact patient management decisions. However, with no data on the specific agents involved in the admission cases, it is impossible to determine whether TOXBASE use improves patient care in cases of poisoning, or whether these relationships can be explained by some other confounding variable.

Further analysis with more detailed admission and attendance data would be required to show this with any certainty. It would be of particular interest to obtain data on more specific toxicant information, and case severity, which could potentially be linked with TOXBASE access data at patient level. ICD10 classifications do allow for more specific classification, therefore more substance specific information would be possible to obtain. In order to do analysis at

this level, it would be beneficial to use TOXBASE user sessions, rather than straightforward TOXBASE accesses, which would allow for data linkage via time stamp, since TOXBASE access data do not correspond to specific patient IDs.

This analysis was also limited by incompleteness in the way in which attendances were recorded, as previously discussed, which meant that some data had to be excluded from the analysis. In doing this, it has been assumed that these errors have occurred across hospitals and is that these recording errors did not occur in a systematic way. This is likely to be the case as was discussed in Chapter 5, where the coding errors appeared to occur randomly throughout the year, though there was a slight increase in poorly recorded data in 2010 which corresponded to NHS restructuring. Since the analysis was carried out at hospital/trust level and was done using rates, the results are unlikely to have been affected.

Chapter 8

Conclusions

The aim of this thesis was to examine demographics of poisoning, along with assessing how the NPIS database TOXBASE is used in hospital emergency departments.

The thesis started by outlining previous findings on the demography of poisoning. This indicated that demographic variables such as age, deprivation and gender were significant factors in cases of poisoning. Young children were more likely to be involved in cases of accidental poisoning, as they explore the world around them, while young adults were most likely to be involved in intentional poisonings [3, 7]. Females were also more likely to be involved in cases of poisoning compared to males [3, 7, 5, 8, 9]. A greater risk of poisoning for more deprived areas has been demonstrated over a number of studies [9, 11, 16, 17, 19].

These findings informed the choice of independent variables in Chapter 2. In this analysis, the proportion of the population of pensionable age was found to indicate a lower proportion of poisoning related calls, which seems reasonable given that adults in that age group are lower risk for poisoning compared to younger age groups. The urban/rural score was found to indicate a higher proportion of poisoning related calls, which is potentially due to these populations making

more use of NHS 24 due to distance to other medical services. Interestingly, a higher proportion of the population living in a deprived area was associated with a lower proportion of calls. This may be linked to urban/rural score, in that deprivation is higher in cities than in more rural areas, and therefore members of deprived populations may prefer to attend hospital, rather than call NHS24.

Chapter 2 also demonstrated that there was a consistent seasonal trend in poisoning related calls to the Scottish health information. The estimated seasonality indicated a peak in calls in summer, which coincides with a similar result found by the American Association of Poison Control Centres [67]. The same seasonality was then found in TOXBASE accesses in Chapter 3 and in poisoning admissions in Chapter 5, which is a strong indication that this seasonality is driven by poisoning cases, rather than being driven by other external factors.

Over the long term, it was shown in Chapter 3, that TOXBASE use has been increasing, at least since 2008. A trend which does not follow trends in admissions over the same time period [24, 75]. Although there was a levelling off in 2010, which lasted through until 2012 when the NICE guideline on the treatment of poisoning indicated that TOXBASE use was best practice in the treatment of poisoned patients [45]. It is also interesting that the second period of increasing TOXBASE use began around the time that new psychoactive substances emerged [33].

In Chapter 5, it was found that TOXBASE use, after accounting for attendances due to poisoning showed no seasonal pattern, which is an indication that TOXBASE is being used consistently with patient attendance from poisoning. There was, however, seasonality in admission due to poisoning after accounting for attendance at emergency departments. This meant that, due to the lack of

attendance data for Scottish hospitals, analysis on both unusual hospitals and on the impact of TOXBASE use in the emergency department was restricted to include locations in England and Wales only.

Focusing on outliers (Chapter 6), there were a few commonalities in those locations which were unusual in some way. Those which were either low or high admitting compared to the rest of the trusts tended to be larger hospitals, as did those hospitals which were low in terms of their TOXBASE use. This may also reflect regional differences, as the larger hospitals based in London tended to be both low accessing and low admitting. The observed regional variation may be attributable to differences in patient management. For example some locations may have their own management plan for paracetamol, which is one of the most commonly ingested substance in cases of poisoning. This may also help to explain some of the differences which were seen in the proportion of accesses to different drug groups, which showed that those hospitals which were unusual in some way tended to have low proportional accesses to paracetamol pages.

It was also shown in Chapter 6 that TOXBASE use did not correlate with hospital performance as measured by SHMI and CQC rating. The likely cause of this is that poisoning, while a relatively common attendance, only accounts for about 1% of emergency department activity, leaving a large amount of additional factors which may also influence the performance measures.

The goal of the analysis outlined in Chapter 7 was to assess whether there was any impact of TOXBASE on the decision to admit (or not admit) patients who present at emergency departments with a suspected case of poisoning. The first stage in this analysis was to fit an interaction model to assess the relationship between TOXBASE use and admissions, while additionally assessing whether this relationship differed for hospitals with greater attendances. The estimated effects were consistent across England and Wales. After accounting for the tempo-

ral patterns in the data, the model indicated that the rate of admissions tended to increase as the rate of access to TOXBASE increased. A very small negative effect of the number of overall attendances was also indicated, which likely corresponds to capacity limitations, where a hospital seeing a large number of attendances across a variety of causes, they are less likely to admit borderline poisoning patients.

A secondary analysis was conducted in Chapter 7, which examined the relationship between the proportion of poisoning attendances admitted and the proportion of accesses made to specific drug categories. This found that there were differences in the estimated relationships between admissions and accesses to individual drug groups. This analysis also highlighted that London seems to use TOXBASE in a way that is inconsistent with the rest of England and Wales. The variation in these drug group level effects are thought to indicate some effect of TOXBASE on clinician decisions, however this analysis was limited by the lack of detail in the admissions and attendances data.

This thesis has shown some overall trends in poisoning, which seem to be consistent in NHS 24 calls, TOXBASE use and attendances due to poisoning. The analysis conducted over the course of the project has provided evidence to suggest a link between TOXBASE use and patient management within emergency departments, however limited data meant that a direct impact could not be established.

The data limitations experienced are relatively common in administrative data. The missingness due to suppression is particularly common is health care data, as a means of protecting individual identities. This issue was fairly straightforward to overcome using the method outlined in Chapter 4.

The linkage of the healthcare data with the TOXBASE access data was not straightforward. Linkage by hospital code was not possible, as this information is not stored by NPIS, however linkage by hospital name was possible, and this is what was done for data provided by ISD in Scotland and the NHS Wales Informatics Service after limited cleaning.

The English data posed more of a problem due to the administrative structures which govern NHS England, combined with the differences in methods of data provision across hospitals and trusts in England. This was overcome by manually creating linkage tables for admissions and attendances respectively, which allowed for the TOXBASE data table to be linked with the administrative data from NHS England. It is worth noting, however, that this was a lengthy process and that decisions made about these data were made after significant research.

Even after this process these data required additional cleaning, due to inconsistencies within the attendance numbers supplied and the admission numbers. After examination of these cases, some were removed due to obvious data quality issues, those with no obvious issues were retained, although it is possible that these data are not accurate, and there is no way to check these data retrospectively.

Using these data in Chapter 7 allowed for the identification of associations between TOXBASE use, attendances at hospital emergency departments and admissions due to drugs poisoning. This analysis could not, however, assess any real impact of TOXBASE use on admissions due to poisoning as was initially hoped. Instead this analysis has indicated a starting point for more detailed analysis of TOXBASE use in relation to emergency department activity, which would require more detailed admissions and attendances data. In particular, data on the specific substance ingested and some measure of case severity would go a long way in reducing the confounding effects seen in the analysis.

Future work might also consider utilising the TOXBASE data in a different way. If session numbers, rather than access numbers, could be extracted then this would potentially answer the question of whether the relationship between admissions and TOXBASE use is caused by case complexity. Using session data may also allow for data to be linked at patient level, although without a specific patient identifier this would be very difficult, and would effectively rely on linking by time rather than patient ID which may bring some subjectivity into the analysis.

Other studies might consider designing an experiment whereby hospitals can be compared, for example if some hospitals were asked to intentionally increase their use of TOXBASE when encountered with a poisoned patient and the others were used as a control group. There may, however, be ethical issues in this type of study, particularly given that TOXBASE is recommended as a standard of care for poisoned patients. This type of analysis could also be implemented as an interrupted time series analysis, where all hospitals are asked to increase their use of TOXBASE on a specific date, the patterns before and after that date could then be compared to assess the difference. These types of experiment would come with their own issues, not least how to monitor hospitals to ensure that they are complying with the instruction to increase TOXBASE use.

One way to minimise this issue could be to compare admissions data from prior to the introduction of TOXBASE to each hospital, with admissions data after TOXBASE has been introduced. Though TOXBASE was introduced around 20 years ago in some locations, which may make it difficult, both in terms of the reliability of TOXBASE data, and the reliability of administrative data.

Appendix A

Region to Postcode

Aberdeen Inner:

AB10, AB11, AB24, AB25

Aberdeen Outer:

AB13, AB14, AB15, AB16, AB21, AB22

Aberdeenshire North:

AB23, AB41, AB42, AB43, AB44, AB45, AB53, AB54

Aberdeenshire South:

AB12, AB30, AB31, AB32, AB33, AB34, AB35, AB36, AB39, AB51, AB52, AB99

Angus:

DD7, DD8, DD9, DD10, DD11

Argyll & Bute Islands:

PA42, PA43, PA44, PA45, PA46, PA47, PA48, PA49, PA60, PA61, PA62, PA63, PA64, PA65, PA66, PA67, PA68, PA69, PA70, PA71, PA72, PA73, PA74, PA75, PA76, PA77, PA78

Argyll & Bute Mainland:

PA20, PA21, PA22, PA23, PA24, PA25, PA26, PA27, PA28, PA29, PA30, PA31, PA32, PA33, PA34, PA35, PA36, PA37, PA38, PA41

Arran & Cumbrae:

KA27, KA28

Badenoch & Strathspey:

PH19, PH20, PH21, PH22, PH23, PH24, PH25, PH26

Caithness & Sutherland

IV24, IV25, IV27, IV28, KW1, KW2, KW3, KW5, KW6, KW7, KW8, KW9, KW10, KW11, KW12, KW13, KW14

Clackmannanshire:

FK10, FK11, FK12, FK13, FK14

Dumfries & Galloway:

DG1, DG2, DG3, DG4, DG5, DG6, DG7, DG8, DG9, DG10, DG11, DG12, DG13, DG14, DG16, DG2, DG3, DG4, DG5, DG6, DG7, DG8, DG9

Dundee City:

DD1, DD2, DD3, DD4, DD5

East Ayrshire:

KA1, KA16, KA17, KA18, KA2, KA3, KA4, KA5, KA6

East Dunbartonshire:

G61, G62, G64, G66

East Lothian:

EH21, EH31, EH32, EH33, EH34, EH35, EH36, EH39, EH40, EH41, EH42

East Renfrewshire:

G46, G76, G77, G78

Edinburgh East:

EH1, EH2, EH3, EH5, EH6, EH7, EH8, EH9, EH15, EH16, EH17, EH99

Edinburgh West:

EH4, EH10, EH11, EH12, EH13, EH14, EH28, EH29, EH30, EH91, EH95

Eilean Siar:

HS1, HS2, HS3, HS4, HS5, HS6, HS7, HS8, HS9

Falkirk:

EH51, FK1, FK2, FK3, FK4, FK5, FK6

Fife East:

DD6, KY1, KY6, KY7, KY8, KY9, KY10, KY14, KY15, KY16

Fife West:

KY2, KY3, KY4, KY5, KY11, KY12, KY99

Glasgow East:

G21, G31, G32, G33, G34, G40

Glasgow South:

G5, G41, G42, G43, G44, G45, G51, G52, G53, G58

Glasgow West:

G1, G2, G3, G4, G11, G12, G13, G14, G15, G20, G22, G23, G90

Helensburgh & Lomond:

G83, G84

Inverclyde:

PA13, PA14, PA15, PA16, PA18, PA19

Inverness & Nairn:

IV2, IV3, IV4, IV5, IV12, IV13, IV63, IV99, PH32

Lochaber:

PA80, PH30, PH31, PH33, PH34, PH35, PH36, PH37, PH38, PH39, PH40, PH41, PH42, PH43, PH44, PH49, PH50

Midlothian:

EH18, EH19, EH20, EH22, EH23, EH24, EH25, EH26, EH37

North Ayrshire:

KA11, KA12, KA13, KA14, KA15, KA20, KA21, KA22, KA23, KA24, KA25, KA29, KA30, PA17

North East Moray:

AB55, AB56, IV30, IV31, IV32

North Lanarkshire East:

ML1, ML2, ML4, ML6, ML7

North Lanarkshire West:

G65, G67, G68, G69, G70, ML5

Orkney Islands:

KW15, KW16, KW17

Perth and Kinross:

KY13, PH1, PH2, PH3, PH4, PH5, PH6, PH7, PH8, PH9, PH10, PH11, PH12, PH13, PH14, PH15, PH16, PH17, PH18

Renfrewshire:

PA1, PA2, PA3, PA4, PA5, PA6, PA7, PA8, PA9, PA10, PA11, PA12

Ross and Cromarty:

IV1, IV6, IV7, IV8, IV9, IV10, IV11, IV14, IV15, IV16, IV17, IV18, IV19, IV20, IV21, IV22, IV23, IV26, IV54

Scottish Borders:

EH38, EH43, EH44, EH45, EH46, TD1, TD10, TD11, TD12, TD13, TD14, TD15, TD2, TD3, TD4, TD5, TD6, TD7, TD8, TD9

Shetland Islands:

ZE1, ZE2, ZE3

Skye and Lochalsh:

IV40, IV41, IV42, IV43, IV44, IV45, IV46, IV47, IV48, IV49, IV51, IV52, IV53, IV55, IV56

South Ayrshire:

KA7, KA8, KA9, KA10, KA19, KA26

South Lanarkshire East:

ML3, ML8, ML9, ML10, ML11, ML12

South Lanarkshire West:

G71, G72, G73, G74, G75, G79

Stirling:

FK7, FK8, FK9, FK15, FK16, FK17, FK18, FK19, FK20, FK21, G63

West Dunbartonshire:

G60, G81, G82

West Lothian:

EH27, EH47, EH48, EH49, EH52, EH53, EH54, EH55

West Moray:

AB37, AB38, IV36

Appendix B

Wales Information Request Form

ADMITTED PATIENT CARE (APC) DATA REQUEST SPECIFICATION FORM

Please note that the fields marked with an asterisk (*) are mandatory. If these fields are not completed your information request may be delayed. For information, hover over the fields in blue.

***Requester's Name:**	Kate Pyper
***Organisation Name:**	University of Strathclyde
Organisation Address:	Click here to enter text.
***Tel No:** 07722271041	**Fax No:** Click here to enter text.
***E-Mail Address:**	Kate.pyper@strath.ac.uk
Date Required:	Click here to enter a date.
Previous Request Number (if applicable):	29373

Group Patient Data By *(Select all that apply):

Total:	☐
Local Health Boards:	☐
Hospital Sites:	☒
Local Health Board Resident Population:	☐
Local Health Board GP Registered Population:	☐
Ward of Residence:	☐

***Patient Coverage:**	All Data – No Filter
***Patient Classification:**	All Cases
***Admission Method:**	Emergency Only
***Activity Count Currency:**	Admissions
Admission/Discharge Based Analysis:	N/A
Diagnosis Criteria:	None
Diagnosis Details	Click here to enter text.
Procedure Criteria:	None
Procedure Details	Click here to enter text.
***Time Period Criteria:**	Calendar Years
From 01/01/2008	**To** 31/12/2015

Split Data By *(Select all that apply)*:

Months:	☒
Quarters:	☐
Sex:	☐
Age Bands:	Other Age Bands
Other Please Specify:	0-14, 15-44, 45+

Data Items Required:

The following data items are considered confidential: Name, Address, Full Postcode, DOB, CRN, GP Code, Consultant Code and Consultant Name.

Aggregated counts of admissions and A&E attendances at hospital level

***Purpose of Data:**

Data will be linked at hospital level with accesses to the TOXBASE Database in order to determine whether use of TOXBASE has an impact on the number of emergency admissions due to poisoning.

Changes to Original Request (if applicable):

Click here to enter text.

Any Other Information:

Click here to enter text.

***Specification Agreed By:** Kate Pyper ***Date:** 02/07/2018

Information Requests are usually analysed and quality assured within seven days, however particularly large analyses may be take longer.

Microsoft Excel Spreadsheet (.xls) is the standard output format however particularly large outputs may be Microsoft Access Databases (.mdb); this is dependent on file size. If the information is required in a different format please indicate this in the *Any Other Information* field above.

If a completed specification form is not returned within seven days it will be assumed the information is no longer required.

For further information or help on completing this form please e-mail pdit.requests@wales.nhs.uk or alternatively telephone (029) 2050 2363

Appendix C

England Information Request Form

NIC number (NIC-23511-T9B9Z) – Tabulation Specification

A&E

Data Source

HES_AE_0708, HES_AE_0809, HES_AE_0910, HES_AE_1011, HES_AE_1112, HES_AE_1213, HES_AE_1314, HES_AE_1415, HES_AE_1516.

ORG_0708, ORG_0809, ORG_0910, ORG_1011, ORG_1112, ORG_1213, ORG_1314, ORG_1415, ORG_1516.

Filters

AEKEY_FLAG	1
DIAG2_N	14 (Poisoning inc. overdose)

Fields:

FYEAR	Financial year
ARRIVALAGE	Arrival age (0-14, 14-44, 45+)
ARRIVALDATE	Arrival date (calendar months)
DIAG2_N	Primary A&E diagnosis
PROCODE5	Provider code
AEKEY_FLAG	Total attendances

Table Format

Table 1 (2007/08) to Table 9 (2015/16)

ProviderCode	ProviderName	Age_Band	Month	Year	Total_Attendances	Poison_Attendances
xxx	xxx	0-14	Jan	2007-08	SUM (AEKEY_FLAG)	SUM (AEKEY_FLAG) where DIAG2_NN = 14
			Feb	2007-08		
			Mar	2007-08		
			etc	etc		
		14-44	Jan	2007-08		
			Feb	2007-08		
			Mar	2007-08		
			etc	etc		
		45+	Jan	2007-08		
			Feb	2007-08		
			Mar	2007-08		
			etc	etc		

APC

Data Source

HES_APC_0708, HES_APC_0809, HES_APC_0910, HES_APC_1011, HES_APC_1112, HES_APC_1213, HES_APC_1314, HES_APC_1415, HES_APC_1516.

ORG_0708, ORG_0809, ORG_0910, ORG_1011, ORG_1112, ORG_1213, ORG_1314, ORG_1415, ORG_1516.

Filters

ADMIMETH	21, 22, 23, 24, 25, 28, 2A, 2B, 2C, 2D (emergency admissions)
DIAG_3_NN	T36-T50 (Drugs), T52-T60 & T63-T65 (Poisoning other causes)

Fields:

FYEAR	Financial year
STARTAGE_CALC	Arrival at start of episode (0-14, 14-44, 45+)
ADMIMETH	Method of admission
ADMIDATE	Date of admission (calendar months)
PROCODE5	Provider code
DIAG_3_NN	Primary diagnosis
FAE	Finished Admission Episode

Table Format

Table 1 (2007/08) to Table 9 (2015/16)

ProviderCode	ProviderName	Age_Band	Month	Year	Emergency Admissions	Drugs	Poisoning - other causes
xxx	xxx	0-14	Jan	2007-08	sum(FAE) where ADMIMETH like '2%'	sum(FAE) where ADMIMETH like '2%' and DIAG_3 = T36 to T50	sum(FAE) where ADMIMETH like '2%' and DIAG_3 = T52 to T60 & T63 to T65
			Feb	2007-08			
			Mar	2007-08			
			etc	etc			
		14-44	Jan	2007-08			
			Feb	2007-08			
			Mar	2007-08			
			etc	etc			
		45+	Jan	2007-08			
			Feb	2007-08			
			Mar	2007-08			
			etc				

NOTES:

Provide additional information to help the customer understand the data/assumptions/output etc

e.g.

- HES Analysis Guide disclosure control rules will be applied to this tabulation and some data may be suppressed.
- This breakdown is likely to produce a large amount of suppressed data which will therefore reduce the amount of useable rows.

The cost for the production and dissemination of the requested tabulation as specified in this document is £1800 plus VAT. Please provide a Purchase Order Number in addition to the acceptance of the Tabulation Specification.

Signed: ……..

Print Name: ……..

Organisation: ……

Date: ……

References

[1] Public Health England, "National poisons information service annual report 2014/15." `http://www.npis.org/NPISAnnualReport2014-15.pdf`, 2015.

[2] D. N. Bateman, R. Carroll, J. Pettie, T. Yamamoto, M. E. M. O. Elamin, L. Peart, M. Dow, J. Coyle, K. R. Cranfield, C. Hook, E. A. Sandilands, A. Veiraiah, D. Webb, A. Gray, P. I. Dargan, D. M. Wood, S. H. L. Thomas, J. W. Dear, and M. Eddleston, "Effect of the UK's revised paracetamol poisoning management guidelines on admissions, adverse reactions and costs of treatment," *British Journal of Clinical Pharmacology*, vol. 78, no. 3, pp. 610–618, 2014.

[3] K. Prescott, R. Stratton, A. Freyer, I. Hall, and I. Le Jeune, "Detailed analyses of self-poisoning episodes presenting to a large regional teaching hospital in the UK," *British Journal of Clinical Pharmacology*, vol. 68, no. 2, pp. 260–268, 2009.

[4] P. C. Burcham, *An Introduction to Toxicology*. Springer, 2014.

[5] D. R. Camidge, R. J. Wood, and D. N. Bateman, "The epidemiology of self-poisoning in the UK," *British Journal of Clinical Pharmacology*, vol. 56, no. 6, pp. 613–619, 2003.

[6] D. L. Eaton and S. G. Gilbert, *Casarett & Doulls Toxicology: The Basic Science of Poisons, 6th ed.* Taylor & Francis, 2001.

[7] E. A. Sandilands and D. N. Bateman, "The epidemiology of poisoning," *Medicine*, vol. 44, no. 2, pp. 76 – 79, 2016.

[8] K. Hawton, H. Bergen, D. Casey, S. Simkin, B. Palmer, J. Cooper, N. Kapur, J. Horrocks, A. House, R. Lilley, R. Noble, and D. Owens, "Self-harm in England: a tale of three cities," *Social Psychiatry and Psychiatric Epidemiology*, vol. 42, no. 7, pp. 513–521, 2007.

[9] E. G. Tyrrell, E. Orton, K. Sayal, R. Baker, and D. Kendrick, "Differing patterns in intentional and unintentional poisonings among young people in England, 19982014: a population-based cohort study," *Journal of Public Health*, vol. 39, no. 2, pp. e1–e9, 2017.

[10] R. A. Payne, J. J. Oliver, M. Bain, A. Elders, and D. N. Bateman, "Patterns and predictors of re-admission to hospital with self-poisoning in scotland," *Public Health*, vol. 123, no. 2, pp. 134–137, 2009.

[11] C. N. A. Mbeledogu, E. V. Cecil, C. Millett, and S. Saxena, "Hospital admissions for unintentional poisoning in preschool children in england; 2000–2011," *Archives of Disease in Childhood*, vol. 100, no. 2, pp. 180–182, 2015.

[12] "Scottish index of multiple deprivation." `http://www.gov.scot/Topics/Statistics/SIMD`, February 2017.

[13] "English indices of deprivation 2015." `https://www.gov.uk/government/statistics/english-indices-of-deprivation-2015`, September 2015.

[14] P. Townsend, P. Phillimore, and A. Beattie, *Health and Deprivation: Inequality and the North*. Croon Helm, 1988.

[15] V. Carstairs and R. Morris, "Deprivation and health in scotland," *Health Bull (Edib)*, vol. 48, no. 4, pp. 162–175, 1990.

[16] S. B. Rafnsson, J. J. Oliver, R. A. Elton, and D. N. Bateman, "Poisons admissions in edinburgh 1981-2001: agent trends and predictors of hospital readmissions," *Human & Experimental Toxicology*, vol. 26, no. 1, pp. 49–57, 2007.

[17] R. Cook, R. Allcock, and M. Johnston, "Self-poisoning: current trends and practice in a UK teaching hospital," *Clinical Medicine*, vol. 8, no. 1, pp. 37–40, 2008.

[18] A. I. C. Teo and J. G. Cooper, "The epidemiology and management of adult poisonings admitted to the short-stay ward of a large scottish emergency department," *Scottish Medical Journal*, vol. 58, no. 3, pp. 149–153, 2013.

[19] L. Groom, D. Kendrick, C. Coupland, B. Patel, and J. Hippisley-Cox, "Inequalities in hospital admission rates for unintentional poisoning in young children," *Injury Prevention*, vol. 12, no. 3, pp. 166–170, 2006.

[20] K. Hawton, "Psychiatric assessment and management of deliberate self-poisoning patients," *Medicine*, vol. 44, no. 2, pp. 103–105, 2016.

[21] S. Darke, L. Degenhardt, and R. Mattick, *Mortality amongst illicit drug users: epidemiology, causes and intervention*. Cambridge University Press, 2006.

[22] "Drug related death in Scotland 2015." https://www.nrscotland.gov.uk/files//statistics/drug-related-deaths/15/drugs-related-deaths-2015.pdf, 2016.

[23] V. Fearn, "Deaths related to drug poisoning in england and wales: 2015 registrations." https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/bulletins/deathsrelatedtodrugpoisoninginenglandandwales/2015registrations, 2016.

[24] L. Barnsdale and J. Trgosz, "Drug related hospital statistics Scotland 2015/16." https://www.isdscotland.org/Health-Topics/Drugs-and-Alcohol-Misuse/Publications/2016-09-27/2016-09-27-DrugHospitalStatistics-Report.pdf, 2016.

[25] A. D. Shah, D. M. Wood, and P. I. Dargan, "Survey of ICD-10 coding of hospital admissions in the UK due to recreational drug toxicity," *QJM: An International Journal of Medicine*, vol. 104, no. 9, pp. 779–784, 2011.

[26] J. B. Florentine and C. Crane, "Suicide prevention by limiting access to methods: a review of theory and practice," *Social science & medicine*, vol. 70, no. 10, pp. 1626–1632, 2010.

[27] P. Molloy, R. Chambers, and T. Cork, "How well are national guidelines relating to the general sales of aspirin and paracetamol, adhered to by retail stores: a mystery shopper study," *BMJ open*, vol. 6, no. 1, p. e010081, 2016.

[28] K. Hawton, H. Bergen, S. Simkin, S. Dodd, P. Pocock, W. Bernal, D. Gunnell, and N. Kapur, "Long term effect of reduced pack sizes of paracetamol on poisoning deaths and liver transplant activity in england and wales: interrupted time series analyses," *Bmj*, vol. 346, p. f403, 2013.

[29] D. N. Bateman, D. Gorman, M. Bain, J. Inglis, F. House, and D. Murphy, "Legislation restricting paracetamol sales and patterns of self-harm and death from paracetamol-containing preparations in Scotland," *British journal of clinical pharmacology*, vol. 62, no. 5, pp. 573–581, 2006.

[30] P. N. Newsome, A. J. Bathgate, N. C. Henderson, A. J. MacGilchrist, J. N. Plevris, G. Masterton, O. J. Garden, A. Lee, P. C. Hayes, and K. J. Simpson, "Referral patterns and social deprivation in paracetamol-induced liver injury in Scotland," *The Lancet*, vol. 358, no. 9293, pp. 1612–1613, 2001.

[31] C. Wilkins, "Recent developments with the establishment of a regulated legal market for new psychoactive substances (legal highs) in New Zealand," *Drug and alcohol review*, vol. 33, no. 6, pp. 678–680, 2014.

[32] S. L. Hill, S. C. Harbon, J. Coulson, G. A. Cooper, G. Jackson, D. J. Lupton, J. A. Vale, and S. H. Thomas, "Methoxetamine toxicity reported to the national poisons information service: clinical characteristics and patterns of enquiries (including the period of the introduction of the UK's first temporary class drug order)," *Emerg Med J*, pp. emermed–2012, 2013.

[33] Public Health England, "National poisons information service annual report 2012/13."
`http://www.npis.org/NPISAnnualReport2012-13.pdf`, 2013.

[34] UK Legislation, "Psychoactive substances act (2016)." http://www.legislation.gov.uk/ukpga/2016/2/contents/enacted.

[35] UK Legislation, "Misuse of drugs act (1971)." https://www.legislation.gov.uk/ukpga/1971/38.

[36] G. E. Appelbe and J. Wingfield, *Dale and Appelbe's Pharmacy and Medicines Law*. Pharmaceutical Press, 2013.

[37] UK Legislation, "Medicines act (1968)." https://www.legislation.gov.uk/ukpga/1968/67.

[38] UK Legislation, "Poisons act (1972)." https://www.legislation.gov.uk/ukpga/1972/66.

[39] UK Legislation, "Deregulation act (2015)." https://www.legislation.gov.uk/ukpga/2015/20.

[40] Council of European Union, "Council directive (EU) 67/548/eec," 1967. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31967L0548:EN:HTML.

[41] Council of European Union, "Classification, packaging and labelling of dangerous preparations." http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=URISERV%3Al21273, 2008.

[42] Council of European Union, "Council regulation (EU) no 1272/2008," 2008. http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32008R1272.

[43] E. Commission, "Contact list of appointed bodies for the submission of the information relating to emergency health response, according to art. 45(1) reg. (EC) no 1272/2008." `http://ec.europa.eu/DocsRoom/documents/5219/attachments/1/translations`, 2014.

[44] D. Bateman, A. Good, C. Kelly, and W. Laing, "Poisons information in Scotland: delivery of TOXBASE on the internet," *Health Informatics Journal*, vol. 8, no. 2, pp. 67–72, 2002.

[45] "The National Institute for Health Care Excellence: Poisoning or overdose - management." `https://cks.nice.org.uk/poisoning-or-overdose#`, 2012.

[46] D. N. Bateman and A. M. Good, "Five years of poisons information on the internet: the UK experience of TOXBASE," *Emergency medicine journal*, vol. 23, no. 8, pp. 614–617, 2006.

[47] F. Fox, P. Stallard, and G. Cooney, "GP's role identifying young people who self-harm: a mixed methods study," *Family practice*, vol. 32, no. 4, pp. 415–419, 2015.

[48] K. Houston, C. Haw, E. Townsend, and K. Hawton, "General practitioner contacts with patients before an after deliberate self harm," *British Journal of General Practice*, vol. 53, pp. 365–370, 2003.

[49] S. Greene, P. Dargan, and A. Jones, "Acute poisoning: understanding 90% of cases in a nutshell," *Postgraduate Medical Journal*, vol. 81, pp. 204–216, 2005.

[50] A. M. Elliott, A. McAteer, D. Heaney, L. D. Ritchie, and P. C. Hannaford, "Examining the role of Scotland's telephone advice service (NHS 24) for managing health in the community: analysis of routinely collected NHS 24 data," *BMJ open*, vol. 5, no. 8, p. e007293, 2015.

[51] J. Munro, F. Sampson, and J. Nicholl, "The impact of NHS Direct on the demand for out-of-hours primary and emergency care," *Br J Gen Pract*, vol. 55, no. 519, pp. 790–792, 2005.

[52] "Hospital accident and emergency activity, 2014-15." `http://webarchive.nationalarchives.gov.uk/20180307210307/http://digital.NHS.uk/catalogue/PUB23070`, 2015.

[53] "Who attends emergency departments." `https://www.isdscotland.org/Health-Topics/Emergency-Care/Publications/2015-09-29/2015-09-29-EmergencyCare-Report.pdf`, 2015.

[54] "NHS Wales statistics." `http://researchbriefings.parliament.uk/ResearchBriefing/Summary/SN06994`, 2016.

[55] K. Pyper, M. Eddleston, D. N. Bateman, D. Lupton, S. Bradberry, E. Sandilands, S. H. Thomas, J. P. Thompson, and C. Robertson, "Hospital usage of TOXBASE in great britain: Temporal trends in accesses 2008 to 2015," *Human & experimental toxicology*, 2018.

[56] B. Lynch, "Spatial surveillance." Unpublished report for Health Protection Scotland, 2018.

[57] "Scottish neighbourhood statistics." `https://statistics.gov.scot/home`. Accessed: 28 Jan 2019.

[58] "Scottish government urban rural classification." `https://www2.gov.scot/Topics/Statistics/About/Methodology/UrbanRuralClassification`. Accessed: 28 Jan 2019.

[59] S. N. Wood, *Generalized Additive Models: An Introduction with R*. Chapman and Hall/CRC, 2 ed., 2017.

[60] S. N. Wood, "Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models," *Journal of the Royal Statistical Society (B)*, vol. 73, no. 1, pp. 3–36, 2011.

[61] S. N. Wood, N. Pya, and B. Säfken, "Smoothing parameter and model selection for general smooth models (with discussion)," *Journal of the American Statistical Association*, vol. 111, pp. 1548–1575, 2016.

[62] S. N. Wood, "Stable and efficient multiple smoothing parameter estimation for generalized additive models," *Journal of the American Statistical Association*, vol. 99, no. 467, pp. 673–686, 2004.

[63] R Core Team, *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2018.

[64] D. Lee and T. Neocleous, "Bayesian quantile regression for count data with application to environmental epidemiology," *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, vol. 59, no. 5, pp. 905–920, 2010.

[65] P. A. Moran, "Notes on continuous stochastic phenomena," *Biometrika*, vol. 37, no. 1/2, pp. 17–23, 1950.

[66] G. Shaddick and J. V. Zidek, *Spatio-Temporal Methods in Environmental Epidemiology*. CRC Press, Taylor & Francis Group, 2016.

[67] J. B. Mowry, D. A. Spyker, D. E. Brooks, A. Zimmerman, and J. L. Schauben, "2015 annual report of the American Association of Poison Control Centers National Poison Data System (NPDS): 33rd annual report," *Clinical Toxicology*, vol. 54, no. 10, pp. 924–1109, 2016.

[68] A. Downing and R. Wilson, "Temporal and demographic variations in attendance at accident and emergency departments," *Emergency Medicine*, vol. 19, no. 6, pp. 531–535, 2002.

[69] E. Dilkes-Frayne, "Drugs at the campsite: Socio-spatial relations and drug use at music festivals," *International journal of drug policy*, vol. 33, pp. 27–35, 2016.

[70] "NHS 24 winter plan." `https://www.NHS24.scot/data/uploads/PDF/NHS_board/Plans/corporate/NHS-24-winter-plan-2018-19.pdf`. Accessed: 24 Jan 2019.

[71] "2015 population estimates analysis tool." `https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/datasets/populationestimatesanalysistool`, 2015. Accessed: 6 Feb 2019.

[72] C. J. Curtis and G. L. Simpson, "Trends in bulk deposition of acidity in the uk, 19882007, assessed using additive models," *Ecological Indicators*, vol. 37, pp. 274 – 286, 2014. Threats to upland waters.

[73] G. L. Simpson, "Identifying periods of change in time series with gams." `http://www.fromthebottomoftheheap.net/2014/05/15/identifying-periods-of-change-with-gams/`, 2014. Accessed: 11 September 2017.

[74] Public Health England, "National poisons information service annual report 2005/06."
`http://www.npis.org/NPISAnnualReport2005-06.pdf`, 2015.

[75] "Statistics on drug misuse."
http://content.digital.NHS.uk/catalogue/PUB21159/drug-misu-eng-2016-rep.pdf,
2016.

[76] G. Jessen, B. Jensen, E. Arensman, U. Bib-Brahe, P. Crepet, D. D. Leo, K. Hawton, C. Haring, H. Hjelmeland, K. Michel, *et al.*, "Attempted suicide and major public holidays in europe: findings from the who/euro multicentre study on parasuicide," *Acta Psychiatrica Scandinavica*, vol. 99, no. 6, pp. 412–418, 1999.

[77] "Seasonal affective disorder: Sad symptoms." `https://www.mind.org.uk/information-support/types-of-mental-health-problems/seasonal-affective-disorder-sad/sad-symptoms/#.WaU4qrKGOUk`, 2016. Accessed: 26 Oct 2017.

[78] H. Gabennesch, "When promises fail a theory of temporal fluctuations in suicide," *Social Forces*, vol. 67, no. 1, pp. 129–145, 1988.

[79] "Swedish poisons information centre annual report."
`https://giftinformation.se/globalassets/publikationer/gic-annual-report-2016.pdf`, 2016.

[80] M. Rocchi, P. Miotto, and A. Preti, "Seasonal variation in suicides and in deaths by unintentional illicit acute drug intoxications," *Addiction biology*, vol. 9, no. 3-4, pp. 255–263, 2004.

[81] C. Haw and K. Hawton, "Suicide and self-harm by drowning: a review of the literature," *Archives of suicide research*, vol. 20, no. 2, pp. 95–112, 2016.

[82] S. Simkin, K. Hawton, P. S. Yip, and C. H. Yam, "Seasonality in suicide: a study of farming suicides in england and wales.," *Crisis: The Journal of Crisis Intervention and Suicide Prevention*, vol. 24, no. 3, p. 93, 2003.

[83] B. Cavanagh, S. Ibrahim, A. Roscoe, H. Bickley, D. While, K. Windfuhr, L. Appleby, and N. Kapur, "The timing of general population and patient suicide in england, 1997–2012," *Journal of affective disorders*, vol. 197, pp. 175–181, 2016.

[84] J. Canner, K. Giuliano, S. Selvarajah, E. Hammond, and E. Schneider, "Emergency department visits for attempted suicide and self harm in the USA: 2006–2013," *Epidemiology and psychiatric sciences*, pp. 1–9, 2016.

[85] P. Casey, I. Gemmell, U. Hiroeh, and C. Fulwood, "Seasonal and socio-demographic predictors of suicide in ireland: a 22 year study," *Journal of affective disorders*, vol. 136, no. 3, pp. 862–867, 2012.

[86] D. Rock, D. M. Greenberg, and J. F. Hallmayer, "Increasing seasonality of suicide in Australia 1970–1999," *Psychiatry research*, vol. 120, no. 1, pp. 43–51, 2003.

[87] D. Madah-Amiri, T. Clausen, L. Myrmel, G. Brattebø, and P. Lobmaier, "Circumstances surrounding non-fatal opioid overdoses attended by ambulance services," *Drug and alcohol review*, vol. 36, no. 3, pp. 288–294, 2017.

[88] E. Griffin, C. B. Dillon, G. ORegan, P. Corcoran, I. J. Perry, and E. Arensman, "The paradox of public holidays: Hospital-treated self-harm and associated factors," *Journal of affective disorders*, vol. 218, pp. 30–34, 2017.

[89] H. Valtonen, K. Suominen, T. Partonen, A. Ostamo, and J. Lönnqvist, "Time patterns of attempted suicide," *Journal of affective disorders*, vol. 90, no. 2, pp. 201–207, 2006.

[90] H. Parker, L. Williams, and J. Aldridge, "The normalization of sensiblerecreational drug use: Further evidence from the north west england longitudinal study," *Sociology*, vol. 36, no. 4, pp. 941–964, 2002.

[91] S. Thomas, L. Bevan, S. Bhattacharyya, M. Bramble, K. Chew, J. Connolly, B. Dorani, K. Han, J. Horner, A. Rodgers, *et al.*, "Presentation of poisoned patients to accident and emergency departments in the north of england," *Human & experimental toxicology*, vol. 15, no. 6, pp. 466–470, 1996.

[92] "Poisons Information Centre of Ireland annual report 2014." `http://www.poisons.ie/docs/Final%20Report%202014.pdf`, 2015.

[93] J. L. Schafer and J. W. Graham, "Missing data: our view of the state of the art.," *Psychological methods*, vol. 7, no. 2, p. 147, 2002.

[94] R. R. Andridge and R. J. Little, "A review of hot deck imputation for survey non-response," *International statistical review*, vol. 78, no. 1, pp. 40–64, 2010.

[95] C. Saha and M. P. Jones, "Bias in the last observation carried forward method under informative dropout," *Journal of Statistical Planning and Inference*, vol. 139, no. 2, pp. 246–255, 2009.

[96] D. B. Rubin, *Multiple imputation for nonresponse in surveys*, vol. 81. John Wiley & Sons, 2004.

[97] D. A. Van Dyk and X.-L. Meng, "The art of data augmentation," *Journal of Computational and Graphical Statistics*, vol. 10, no. 1, pp. 1–50, 2001.

[98] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the em algorithm," *Journal of the royal statistical society. Series B (methodological)*, pp. 1–38, 1977.

[99] M. A. Tanner and W. H. Wong, "The calculation of posterior distributions by data augmentation," *Journal of the American statistical Association*, vol. 82, no. 398, pp. 528–540, 1987.

[100] J. Tobin, "Estimation of relationships for limited dependent variables," *Econometrica: journal of the Econometric Society*, pp. 24–36, 1958.

[101] "ICD10." `https://icd.who.int/browse10/2015/en#/T51-T65`, 2015. Accessed: 2 May 2016.

[102] "NHS Wales data dictionary." `http://www.datadictionary.wales.nhs.uk/#!WordDocuments/accidentandemergencydiagnosistypes.htm`. Accessed: 6 June 2019.

[103] "HES Data Dictionary: Accident and Emergency." `https://digital.nhs.uk/binaries/content/assets/legacy/pdf/3/l/hes_data_dictionary_-_accident_and_emergency.pdf`. Accessed: 6 June 2019.

[104] "Equity and excellence: Liberating the NHS." https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/213823/dh_117794.pdf, 2010. Accessed: 16 June 2019.

[105] L. Ewbank, T. J, and M. H, "NHS hospital bed numbers: past, present, future." https://www.kingsfund.org.uk/publications/NHS-hospital-bed-numbers, 2017. Accessed: 19 June 2019.

[106] D. J. Spigelhalter, "Funnel plots for comparing institutional performance," *Statistics in Medicine*, vol. 24, pp. 1185–1202, 2005.

[107] D. Eayres, "APHO technical briefing." https://fingertips.phe.org.uk/documents/APHO%20Tech%20Briefing%203%20Common%20PH%20Stats%20and%20CIs.pdf, 2008. Accessed: 24 May 2019.

[108] "Summary hospital-level mortality index." https://digital.NHS.uk/data-and-information/publications/ci-hub/summary-hospital-level-mortality-indicator-shmi. Accessed: 23 Jan 2019.

[109] "Care Quality Commission." https://www.cqc.org.uk/. Accessed: 23 Jan 2019.

[110] "WHO: Information sheet on opioid overdose." https://www.who.int/substance_abuse/information-sheet/en/. Accessed: 7 Feb 2019.

[111] R. Thanacoody, "Opioid overdose: Treatment algorithm." `https://bestpractice.bmj.com/topics/en-gb/339/treatment-algorithm`. Accessed: 3 April 2019.

[112] M. Elamin, D. James, P. Holmes, G. Jackson, J. Thompson, E. Sandilands, S. Bradberry, and S. Thomas, "Reductions in emergency department visits after primary healthcare use of the UK National Poisons Information Service," *Clinical Toxicology*, vol. 56, no. 5, pp. 342–347, 2018.