



# Development and Testing of Early Detection Systems for Seasonal and Pandemic Influenza

**Muqrin Abdulrahman Almuqrin**

Department of Mathematics and Statistics

University of Strathclyde

Glasgow, UK

2018

This thesis is submitted to the University of Strathclyde for the  
the degree of Doctor of Philosophy in the Faculty of Science

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

© : The copyright of this thesis belongs to the author under the terms of the United Kingdom Copyright Acts as qualified by the University of Strathclyde Regulation 3.50. Due acknowledgement must always be made of the use of any material contained in, or derived from, this thesis.

## *Dedicated to*

To whom was the first motivation and my greatest fan of my studies and research;  
to the solid bond and the greatest person of my life;  
to his spirit and the sudden death of his life in the middle of my preparation of this  
research on the 24th of July 2016;

To my father Dr. Abdulrahman Almuqrin -may Allah have mercy on him-.  
To my mother, Noura Alnasser, who has done more than she can to encourage me to  
continue and succeed.

To my uncle Sheikh. Ali Almuqrin;  
to the closest friend to my father Sheikh. Abdullah Almasa'ar;  
to all my brothers and sisters;  
to my wife Reham Alnasser who was a permanent assistant to me from the start; and  
to my children Abdulrahman, Hosam and Noura.

# Acknowledgements

First and foremost, I am extremely thankful to Almighty Allah for his blessings and providing me with the ability to carry out this research, without which none of my work would have been possible.

I would like to say thank you to the Department of Mathematics and Statistics of the University of Strathclyde for their continuous support, patience and understanding all these years, and also for making me feel comfortable from the first moment of my arrival.

I would also like to thank my supervisors Prof. Chris Robertson and Dr. Alison Gray, for their time and effort to provide me with all the important help and support, so that this project was made possible, and also for their patience and understanding.

To the sun and the moon of my life: my parents. To the hands that have never stopped praying for me. To your eyes, I am honoured to give you my effort for 5 years. I have done nothing without your unconditional love and support. Always thanks for hearts that never put me down, have believed in me, and who have also shared my moments of weakness. But more importantly, who have also shared my happiness and passion for success. I have always heard your strong voices declaring “you will do it!” My father, I wished you were alive, so you could rejoice in what I have achieved through my PhD journey, due to your constant encouragement.

To my sisters and brothers, thank you for your continuous support. I have never stopped being proud of you, my love for you all.

Most importantly, I would like to express my greatest appreciation to my wife, who has always supported me to pursue my dreams. My children have been the inspiration behind my passion to improve myself and accomplish more in life.

Many thanks go to Health Protection Scotland (HPS) for the providing of the influenza-like illness (ILI) and acute-respiratory illness (ARI) data used in the research of this project.

Many thanks go to Majmaah University and also Saudi Cultural Bureau in London who supported me before I left Saudi Arabia and also during my PhD journey.

I would also like to thank Mr. Ian Thurlbeck for all the invaluable IT support that I have needed throughout my PhD journey.

Many thanks to Dr. Jiafeng Pan for her important advice and support with using R/Rstudio software.

Many thanks to Dr. Omar Alshaimi for his appreciated advices and support with using L<sup>A</sup>T<sub>E</sub>X.

A big thank you for all of my friends who helped and supported me during my stay in Glasgow since 2014. Thank you all for all the good times we had outside the university.

# Publications and presentations arising from this thesis

On 4th September 2017, some of this thesis was presented as a talk in the Royal Statistical Society conference in Glasgow [5].

# Abstract

Medical and statistical studies and research have shown that respiratory viruses are important and dangerous to human life.

Daily data were studied for Acute Respiratory Illness (ARI) and influenza-like illness (ILI) consultations for surveillance for influenza from all 14 health boards (HBs) in Scotland. The National Health Service (NHS) provided the data from 2009 to 2014.

The weekly case ratio ( $WCR$ ) method, developed for pandemic detection, may be a useful way of detecting seasonal influenza. The target is to find a simple way to extend the  $WCR$  method and compare it to other well established systems. This method is based on two terms, the value of  $WCR$  defined as the total influenza rates reported to all GPs in week  $w$  divided by the total rates in week  $w - 1$  and  $N_{HB}$  defined as the number of HBs which have a  $WCR > 1$ . We use daily data for ILI consultations. The starting point is using Scotland data to investigate how effective the  $WCR$  method would be for Scotland data. We then extend this situation, through simulation based upon the Scottish data to have more than 14 HBs with the same structure as Scotland.

The next step is to develop the  $WCR$  algorithm for smaller spatial areas. We created another data structure using 30, 40 and 50 HBs from the original Scottish data, using different population size structures, then simulated more than 3,000,000 cases of ILI, considering the rate as flat during the year, then we got the joint distribution for  $WCR$  and  $N_{HB}$  in the case of no epidemic.

We attempted to identify a rejection region for a test using the null joint distribution between  $WCR$  and  $N_{HB}$ .

The modified  $WCR$  system exceeds performance of other systems in some circumstances.

# Abbreviations

ACIP : The Advisory Committee on Immunizations Practices.

ARI : Acute respiratory illness.

ARIMA : Autoregressive integrated moving average.

*AW* : Alert week.

*CD* : Correct detection.

CDC : The Centre for Disease Control and Prevention.

CDSC : Communicable Disease Surveillance Centre.

CI : Confidence intervals.

CUSUM : Cumulative summation.

FA : False alarm.

GAM : Generalised additive model.

GLM : Generalised linear model.

GP : General Practitioner.

HBs : Health boards.

HMM : Hidden Markov model.

HPAI : Highly Pathogenic Avian Influenza.

HPS : Health Protection Scotland.

$H_0$  : Null hypothesis.

$H_1$  : Alternative hypothesis.

IMPACT : IMmediate Possibility to ACcess oseltamivir Treatment.

ILI : Influenza-like illness.

LPAI : Low Pathogenicity Avian Influenza.

*MED* : Median detection time.

MEM : Moving epidemic method.

*n.area* : The number of locations (the same as HBs).

*ng* : The number of locations (the same as HBs and *n.area*).

NHS : The National Health Service.

$N_{HB}$  : The number of HBs which report more ILI or ARI cases in week  $w$  compared with week  $w - 1$ .

$n_i$  : Similar to  $N_{HB}$  but starting at  $\frac{ng}{2}$  and going up to the end of  $ng$ .

*n.i.midway* : The number of increases above midway in  $N_{HB}$ .

NPV : Negative predictive value.

OLS : Ordinary least squares.

$P_1$  : Scotland population structure.

$P_2$  : Equal population structure.

pdf : Probability density function.

pmf : Probability mass function.



PPV : Positive predictive value.

RSV : Respiratory syncytial virus.

SARI : Severe acute respiratory infection.

SARS : Severe acute respiratory syndrome.

SE : Southeast.

Sen : Sensitivity.

SERVIS : Scottish Enhanced Respiratory Virus Infection Surveillance.

Sp : Specificity.

var : Variance.

WHO : The World Health Organization.

*WCR* : Weekly Case Ratio.

$\mu$  : Mean.

$\sigma$  : Standard deviation.

$\sigma^2$  : Variance.

# Contents

<b>Acknowledgements</b>	<b>iii</b>
<b>Publications and presentations arising from this thesis</b>	<b>v</b>
<b>Abstract</b>	<b>vi</b>
<b>Abbreviations</b>	<b>vii</b>
<b>1 Background about influenza</b>	<b>1</b>
1.1 What is influenza? . . . . .	1
1.2 Transmission . . . . .	2
1.2.1 Classification of influenza . . . . .	3
1.3 Symptoms . . . . .	3
1.4 History of influenza . . . . .	4
1.5 Burden of influenza . . . . .	6
1.6 Pandemics, seasonal influenza . . . . .	7
1.6.1 Geographic pattern . . . . .	8
1.7 Vaccination . . . . .	9
1.8 Treatments . . . . .	10
1.9 Aims and objectives . . . . .	11
1.10 Conclusion . . . . .	12
<b>2 Data Description</b>	<b>13</b>
2.1 Scotland Data . . . . .	13
2.2 Methods . . . . .	15
2.3 Data . . . . .	15
2.4 Scotland Data analysis . . . . .	16
2.5 GAM model . . . . .	19
2.5.1 ARI Data for each (HB) with GAM models . . . . .	20
2.5.2 GAM models only for health boards . . . . .	30
2.6 Comments . . . . .	34

<b>3</b>	<b>Influenza detecting methods</b>	<b>35</b>
3.1	Threshold method . . . . .	36
3.2	Time series technique . . . . .	37
3.2.1	ARIMA method . . . . .	37
3.3	Moving Epidemic Method (MEM) . . . . .	38
3.4	Cumulative Summation (CUSUM) Method . . . . .	41
3.5	Hidden Markov Model . . . . .	43
3.6	Weekly Cases Ratio (WCR) method . . . . .	44
3.7	Farrington method . . . . .	45
3.8	Conclusion . . . . .	48
<b>4</b>	<b>Applying and modifying detection systems</b>	<b>50</b>
4.1	Applying the MEM method to the data for Scotland . . . . .	50
4.2	Applying the <i>WCR</i> method to the Scotland data . . . . .	66
4.2.1	Using the <i>WCR</i> method for summer . . . . .	66
4.2.2	Using the <i>WCR</i> method for winter . . . . .	72
4.2.3	The joint probability distribution of ( <i>WCR</i> , $N_{HB}$ ) . . . . .	78
<b>5</b>	<b>Overview of the system - Simulation</b>	<b>86</b>
5.1	Simulation with different spatial locations and structures . . . . .	86
5.2	Different spatial locations - Steady State . . . . .	88
5.2.1	Approaches to choosing different baselines for the simulation . . . . .	90
5.2.2	Choosing from Scottish ILI counts . . . . .	91
5.2.3	Fixed rate in all locations . . . . .	91
5.2.4	Varying rate in different locations . . . . .	92
5.3	Finding functions to relate <i>WCR</i> and $N_{HB}$ (1) . . . . .	92
5.3.1	New simulation for <i>WCR</i> and $N_{HB}$ . . . . .	99
5.4	Finding functions to relate <i>WCR</i> and $N_{HB}$ (2) . . . . .	101
5.4.1	Modelling of means . . . . .	102
5.4.2	Modelling of variance . . . . .	107
5.5	Testing coverage intervals (CI) for the link between <i>WCR</i> and $N_{HB}$ . . . . .	113
5.6	Conclusion . . . . .	116
<b>6</b>	<b>Detection system</b>	<b>118</b>
6.1	General view of the system . . . . .	119
6.1.1	Example . . . . .	122
6.2	Notes and modifications to generalize the system . . . . .	125
6.3	More general terms in the system . . . . .	126
6.3.1	Linear Function ( $f_1$ ) . . . . .	127

6.3.2	Ellipse Equation ( $f_2$ ) . . . . .	128
6.3.3	Concave Ellipse Equation ( $f_3$ ) . . . . .	130
6.3.4	Exponential function ( $f_4$ ) . . . . .	130
6.3.5	Example . . . . .	131
6.4	Making the system more general . . . . .	133
6.5	Conclusion . . . . .	139
<b>7</b>	<b>Simulation study</b>	<b>141</b>
7.1	Plan of the simulation study . . . . .	142
7.2	Different Shape Cases . . . . .	144
7.2.1	Simulate from a GAM from all 5 years historical data . . . . .	144
7.2.2	Simulate from a GAM from only the last year . . . . .	146
7.2.3	Repeating with an increasing trend . . . . .	147
7.2.4	Repeating with a decreasing trend . . . . .	148
7.2.5	Serfling model . . . . .	149
7.2.6	Linear Model . . . . .	151
7.3	Illustration . . . . .	154
7.4	Performance measurements . . . . .	154
7.5	Conclusion . . . . .	156
<b>8</b>	<b>Testing <math>WCR</math> with other systems (1)</b>	<b>157</b>
8.1	Simulation for different scenarios . . . . .	157
8.2	Testing the $WCR$ system (1) . . . . .	159
8.2.1	Testing $WCR_{v2}$ and $WCR_{v3}$ . . . . .	168
8.2.2	Testing $WCR_{v1,2,3}$ with different setting of the alert week . . . . .	170
8.3	Conclusion . . . . .	173
<b>9</b>	<b>Testing <math>WCR</math> with other systems (2)</b>	<b>176</b>
9.1	Symmetry deception . . . . .	177
9.2	Baseline of simulation (hypotheses) . . . . .	178
9.3	Modelling means and variance . . . . .	182
9.3.1	Testing of mean and variance equation . . . . .	184
9.4	Testing the $WCR$ system (2) . . . . .	188
9.4.1	Testing $WCR$ with different alert weeks . . . . .	191
9.5	Improving $WCR$ performance . . . . .	193
9.6	Conclusion . . . . .	200
<b>10</b>	<b>Conclusions and further work</b>	<b>201</b>
10.1	Conclusions . . . . .	202
10.2	Further work . . . . .	203

---

<b>Appendix</b>	<b>205</b>
<b>A Chapter 5: Overview of the system - Simulation</b>	<b>205</b>
A.1 Different spatial locations - Steady State . . . . .	205
A.1.1 Investigation of Normality of $WCR$ : histograms of $WCR$ . . . . .	205
A.1.2 Investigation of Normality of $WCR$ : normal QQ plots . . . . .	213
A.2 Finding functions to relate $WCR$ and $N_{HB}$ (1) . . . . .	221
A.3 Finding functions to relate $WCR$ and $N_{HB}$ (2) . . . . .	225
<b>B Chapter 6: Detection system</b>	<b>231</b>
B.1 Making the system more general . . . . .	231
<b>C Chapter 7: Simulation study</b>	<b>236</b>
C.1 Illustration . . . . .	236
<b>D Chapter 9: Testing <math>WCR</math> with other systems (2)</b>	<b>237</b>
D.1 Testing $WCR$ with other systems (2) . . . . .	237
D.2 Testing of mean and variance equation . . . . .	239

# List of Figures

1.1	Influenza map [182] . . . . .	8
2.1	Map of the territorial health boards in Scotland [66] . . . . .	14
2.2	ARI weekly consultations rates in 2009-2015 for Scotland, for all 14 health boards. . . . .	17
2.3	ILI weekly consultations rates in 2009-2015 for Scotland, for all 14 health boards. . . . .	18
2.4	ARI consultation rates in 2009-2010 for Scotland with fitted GAM models shown in red. . . . .	21
2.5	ARI consultation rates in 2010-2011 for Scotland with fitted GAM models shown in red. . . . .	23
2.6	ARI consultation rates in 2011-2012 for Scotland with fitted GAM models shown in red. . . . .	24
2.7	ARI consultation rates in 2012-2013 for Scotland with fitted GAM models shown in red. . . . .	26
2.8	ARI consultation rates in 2013-2014 for Scotland with fitted GAM models shown in red. . . . .	27
2.9	ARI consultation rates in 2014-2015 for Scotland with fitted GAM models shown in red. . . . .	29
2.10	GAM models for 2009-2010 . . . . .	30
2.11	GAM models for 2010-2011 . . . . .	31
2.12	GAM models for 2011-2012 . . . . .	32
2.13	GAM models for 2012-2013 . . . . .	32
2.14	GAM models for 2013-2014 . . . . .	33
2.15	GAM models for 2014-2015 . . . . .	34

4.1	Epidemic threshold, levels of intensity and modelled ARI season 2014-2015, by HBs. MEM plots include the pre-epidemic period (blue), epidemic period (red), and post-epidemic (green), while the triangle gives the alert week and the red line presents the pre-epidemic threshold, while the post-epidemic threshold is presented by the green line. A 5 year series (2009/10 to 2013/14) is used in MEM. . . . .	54
4.2	Epidemic threshold, levels of intensity and modelled ILI season 2014-2015, by HBs. MEM plots include the pre-epidemic period (blue), epidemic period (red), and post-epidemic (green), while the triangle gives the alert week and the red line presents the pre-epidemic threshold, while the post-epidemic threshold is presented by the green line. A 5 year series (2009/10 to 2013/14) is used in MEM. . . . .	55
4.3	MEM figures for ARI data using 5, 4, 3 and 2 years historical data to detect a pandemic in winter 2014/15. MEM plots include the pre-epidemic period (blue), epidemic period (red), and post-epidemic (green), while the triangle gives the alert week and the red line presents the pre-epidemic threshold, while the post-epidemic threshold is presented by the green line. We used the default setting of <code>mem</code> package [160] in R [129] which is the threshold parameter ( $\delta$ ) calculated from the arithmetic mean and point 95% confidence interval (standard deviations). . . . .	57
4.4	Sensitivity and specificity using MEM for ARI data with 5, 4, 3 and 2 years historical data to detect a pandemic in winter 2014/15. . . . .	58
4.5	MEM figures for ILI data using 5, 4, 3 and 2 years historical data to detect a pandemic in winter 2014/15. MEM plots include the pre-epidemic period (blue), epidemic period (red), and post-epidemic (green), while the triangle gives the alert week and the red line presents the pre-epidemic threshold, while the post-epidemic threshold is presented by the green line. We used the default setting of <code>mem</code> package [160] in R [129] which is the threshold parameter ( $\delta$ ) calculated from the arithmetic mean and point 95% confidence interval (standard deviations). . . . .	59
4.6	Sensitivity and specificity using MEM for ILI data with 5, 4, 3 and 2 years historical data to detect a pandemic in winter 2014/15. . . . .	60
4.7	Sensitivity and specificity using MEM for ARI data with 2 years historical data to detect a pandemic in the following winter. . . . .	61
4.8	Sensitivity and specificity using MEM with ILI data with 2 years historical data to detect a pandemic in the following winter. . . . .	62
4.9	Sensitivity and specificity plots using MEM for ARI data with 3 years historical data to detect a pandemic in the following winter. . . . .	63

4.10	Sensitivity and specificity plots using MEM for ILI data with 3 years historical data to detect a pandemic in the following winter. . . . .	64
4.11	Sensitivity and specificity plots using MEM for ARI data with 4 years historical data to detect a pandemic in the following winter. . . . .	65
4.12	Sensitivity and specificity plots using MEM for ILI data with 4 years historical data to detect a pandemic in the following winter. . . . .	65
4.13	$WCR$ against weeks in 2010 (April to September). . . . .	67
4.14	$WCR$ against weeks in 2011 (April to September). . . . .	68
4.15	$WCR$ against weeks in 2012 (April to September). . . . .	70
4.16	$WCR$ against weeks in 2013 (April to September). . . . .	71
4.17	$WCR$ against weeks in 2014 (April to September). . . . .	72
4.18	$WCR$ against weeks in 2010/11 (October to March). . . . .	74
4.19	$WCR$ against weeks in 2011/12 (October to March). . . . .	75
4.20	$WCR$ against weeks in 2012/13 (October to March). . . . .	76
4.21	$WCR$ against weeks in 2013/14 (October to March). . . . .	77
4.22	The joint frequency distribution of $(WCR, N_{HB})$ over all health boards in all years, for summer and winter combined. Individual points represent different weeks within any one year. . . . .	79
4.23	The joint frequency distribution of $(WCR, N_{HB})$ over all health boards in all years, for winter. Individual points represent different weeks within any one year. . . . .	79
4.24	The joint frequency distribution of $(WCR, N_{HB})$ over all health boards in all years, for summer. Individual points represent different weeks within any one year. . . . .	80
5.1	$WCR$ against $N_{HB}$ for flat rates with 14 HBs and population of 5 million. The black gradient reflects the intensity of the number of simulations. This uses 10,000 cases each week for 6 years. . . . .	89
5.2	$WCR$ against $N_{HB}$ for flat rates with 30 HBs and population of 25 million. The black gradient reflects the intensity of the number of simulations. This uses 10,000 cases each week for 6 years. . . . .	89
5.3	$WCR$ against $N_{HB}$ for flat rates with 50 HBs and population of 50 million. The black gradient reflects the intensity of the number of simulations. This uses 10,000 cases each week for 6 years. . . . .	90
5.4	The relation between $N_{HB}$ and $\mu$ ( $mu$ ) in the cases of 14, 30 and 50 HBs (left to right). The red lines are straight lines fitted to the mean of $WCR$ in each HB using models $m_{14}$ , $m_{30}$ , $m_{50}$ in Appendix A.2. The size of circles represents the number of simulations at each $N_{HB}$ . . . . .	94



5.5	Predicted mean of $WCR$ for the case of 14 HBs (red) and predicted mean $WCR$ from models based on the cases of 30 (blue) and 50 (black) HBs from figure 5.4. . . . .	95
5.6	The relation between $N_{HB}$ and $\sigma$ ( $sd$ ) in the cases of 14, 30 and 50 HBs (left to right). The red lines are straight lines fitted to the standard deviations of $WCR$ in each HB using models <code>sd14</code> , <code>sd30</code> , <code>sd50</code> in Appendix A.2. The size of circles represents the number of simulations at each $N_{HB}$ . . . . .	96
5.7	The suitable power ( $\lambda$ ) of $\sigma$ in the cases of 14, 30 and 50 HBs using the Box-Cox technique, shown as the maximum point of the log-likelihood curve. 98	98
5.8	The relation between $N_{HB}$ and $\sigma$ in the cases of 14, 30 and 50 HBs. The red lines present the models with the power of 14, -1 and 3 in the cases of 14, 30 and 50 HBs respectively using models <code>sd14BC</code> , <code>sd30BC</code> , <code>sd50BC</code> in Appendix A.2. The size of circles represents the number of simulations in each $N_{HB}$ . . . . .	98
5.9	Predicted standard deviation of $WCR$ for the case of 14 HBs with the power of 14 (red) and predicted standard deviations of $WCR$ from models based on the cases of 30 HBs with the power of -1 (blue), and 50 HBs with the power of 3 (black) from figure 5.8. . . . .	99
5.10	$WCR$ against $N_{HB}$ for flat rates with 40 HBs and population of 50 million. The black gradient reflects the intensity of the number of simulations. . . . .	100
5.11	$WCR$ against $N_{HB}$ for flat rates with 40 HBs and population of 50 million. The black gradient reflects the intensity of the number of simulations. This uses 10,000 cases each week for 6 years. Red squares show the upper and lower 95% confidence intervals for $WCR$ from the $\mu$ and $\sigma$ from the model from the case of 14 HBs, and purple circles represent the same for HBs =30, and green triangles represent the same for HBs =50. . . . .	101
5.12	Means ( $m$ ) of $WCR$ against $n.i.midway$ for $ng=$ 14, 30, 40, 50 represented as circled points. Dark points represent the predicted values from model ( <code>z3</code> ) table 5.9. The size of circles represents the number of simulations for each $ni$ . . . . .	105
5.13	Slopes of the model for the mean $WCR$ against $ng=$ 14, 20, 30, 40, 50. The blue line represents the predicted slope for each $ng$ (from the model ( <code>z4</code> ) in table 5.11). . . . .	106
5.14	Variances ( $var$ ) of $WCR$ against $n.i.midway$ for $ng=$ 14, 30, 40, 50 represented as circles. The size of circles represents the number of simulations for each $ni$ . . . . .	108
5.15	Histogram of variance of $WCR$ over all $N_{HB}$ in all $ng$ , which clearly shows the variances are not normally distributed. . . . .	109

5.16	Box plots of variance of $WCR$ for each $ng$ , to show the distribution. . . . .	109
5.17	Variances against <i>n.i.midway</i> for $ng= 14, 30, 40, 50$ represented as circled points. Dark points represent the predicted values from model (z6) in table 5.13. The size of circles represents the number of simulations for each $N_{HB}$ . . . . .	111
5.18	Intercepts of the model (z6) for variance of $WCR$ (Coefficient ( $ng$ )) against $ng=14, 20, 30, 40, 50$ . The blue line represents the predicted intercept for each $ng$ from model (z7) in table 5.15. . . . .	112
5.19	$WCR$ against $N_{HB}$ for 3,110,000 cases with $ng=25$ . The red squares represent the 95% CI and the green triangles represent the 99% CI using the general equations to get the mean and standard deviation of $WCR$ . The black gradient reflects the intensity of the number of simulations. This uses 10,000 cases each week for 6 years. . . . .	114
5.20	$WCR$ against $N_{HB}$ for 3,110,000 cases with $ng=34$ . The red squares represent the 95% CI and the green triangles represent the 99% CI using the general equations to get the mean and standard deviation of $WCR$ . The black gradient reflects the intensity of the number of simulations. This uses 10,000 cases each week for 6 years. . . . .	115
5.21	$WCR$ against $N_{HB}$ for 3,110,000 cases with $ng=44$ . The red squares represent the 95% CI and the green triangles represent the 99% CI using the general equations to get the mean and standard deviation of $WCR$ . The black gradient reflects the intensity of the number of simulations. This uses 10,000 cases each week for 6 years. . . . .	116
6.1	General approach to identifying a rejection region from the distribution of $WCR$ and $N_{HB}$ , which can apply to any case of $N_{HB}$ . . . . .	120
6.2	Size of Binomial probabilities for a given $ng$ , with success probability 0.5, because we assumed that $WCR$ are symmetric in section 5.2. . . . .	122
6.3	$WCR$ against $N_{HB}$ for flat rates with 40 HBs and population of 50 million. The black gradient reflects the intensity of the number of simulations. This uses 10,000 cases each week for 6 years. Red steps show the cut-offs from the $\mu$ and $\sigma$ calculated from the case of 14 HBs and purple steps represent the cut-offs from HBs= 30 and green steps represent the cut-offs from HBs= 50 after scaling them from the midway to the end of HBs to be equivalent to this case with HBs= 40. . . . .	124

- 6.4 Quantile probabilities ( $k$ ) at each  $N_{HB}$  are shown by the red squares for  $f_1$  (section 6.3.1), and by purple triangles for  $f_2$  (section 6.3.2). Blue triangles represent  $f_3$  (section 6.3.3) and the orange diamonds represent  $f_4$  (section 6.3.4). Dark green circles represent the manual values ( $m$ ) of quantile probabilities (section 6.1). This is based on our example in the following section 6.3.5 which considers HBs= 40. Here 1 in the  $X$ - axis represents  $N_{HB}^{**}$ , and, 7 represents  $N_{HB}^*$ . . . . . 129
- 6.5 The general shape of an ellipse centered on (0,0). . . . . 130
- 6.6 Quantiles (cut-offs) from  $N_{HB}^{**}$  to  $N_{HB}^* - 1$  are represented by the red squares for  $f_1$  in section 6.3.1, while purple triangles represent  $f_2$  in section 6.3.2. Blue triangles represent  $f_3$  in section 6.3.3, and the orange diamonds represent the quantile from  $f_4$  in section 6.3.4. Dark green circles show the manual values to calculate quantiles ( $m$ ) in section 6.1. . . . . 132
- 6.7  $WCR$  against  $N_{HB}$  for steady rates using  $r = 5$  with 14 HBs and a 5 million population.  $N_{HB}^*$  is presented in the dotted vertical orange line while the dotted vertical purple line represents  $N_{HB}^{**}$ . The red shaped area columns came from calculating the quantile (cut-offs) using  $\mu$  and  $\sigma$  from many simulated cases with a flat rate. They represent around 5% of the most extreme observations under the steady state. The green shaped area columns came from calculating  $qnorm$  (cut-offs) using  $\mu$  and  $\sigma$  from equations 5.4.7 and 5.4.10. They represent around 9% of the most extreme observations under the steady state. Darkness of the black colour reflects the number of simulated cases. . . . . 138
- 7.1 Simulated weekly cases from the historical ILI data (2009-2015). . . . . 145
- 7.2 MEM and CUSUM methods' performance with simulated cases from the historical ILI and ARI data; MEM plots (left) include the pre-epidemic period (blue), epidemic period (red), and post-epidemic (green), while the triangle gives the alert week and the red line presents the pre-epidemic threshold, while the post-epidemic threshold is presented by the green line. A 5 year series (2009/10 to 2013/14) is used in MEM. CUSUM plots (right) include the normal weeks (black) and the epidemic weeks (red), while the dotted line presents the threshold based on the fact that 95% of the area of a normal distribution is within 1.96 standard deviations of the mean. . . . . 146
- 7.3 Simulated weekly cases from the historical data (repeating 2013/14 scenario 5 times). . . . . 147
- 7.4 Simulated weekly cases from the historical data (repeating 2013/14 scenario 5 times with an increasing trend). . . . . 148

- 7.5 Simulated weekly cases from the historical data (repeating the 2009/10 scenario 5 times with a decreasing trend. There is a discrete jump at the end of each year because we selected year 2009/10 and this year has an end rate higher than its start rate (as the idea is to get a decreasing trend). . . 149
- 7.6 Simulated weekly cases from the Serfling model (single peak in winter),  $\alpha = -9, \beta_1 = 1, \beta_2 = 1$ . . . . . 150
- 7.7 Simulated weekly cases from the Serfling model (double peaks with main peak in winter),  $\alpha = -9, \beta_1 = 1, \beta_2 = 0.5, \beta_3 = 0, \beta_4 = 1$ . . . . . 150
- 7.8 Simulated weekly cases from the Serfling model (peak then increasing trend),  $\alpha = -9, \beta_1 = -0.1, \beta_2 = -0.1, \beta_3 = 0.4, \beta_4 = 1.2$ . . . . . 151
- 7.9 Simulated weekly cases from a linear shape (stability at a low level, then sudden stability at a high level, before going back suddenly to a low level). 152
- 7.10 Simulated weekly cases from a linear shape (stability at a low level, then a gradual increase to a high level before going back gradually to a low level). 153
- 7.11 Simulated weekly cases from a linear shape (stability at a low level, then a gradual deviation to a peak level, then a sudden decrease to a low level). . . 153
- 7.12 Example of specifying the start of rates increasing, and calculating  $SN$ ,  $SP$  and  $MED$ . More details are given in the text above. A1, A2 and A3 represent alarms 1, 2 and 3 respectively. . . . . 156
- 8.1  $WCR$  against  $N_{HB}$  for historical rates with 14 HBs and a 5 million population.  $N_{HB}^*$  is presented in the dotted vertical orange line while the dotted vertical purple line represents  $N_{HB}^{**}$ . The blue shaped area columns came from calculating the quantile using  $\mu$  and  $\sigma$  from the flat rate. They represent the 5% most extreme observations under the steady state. Darkness of the black colour reflects the number of simulated cases. . . . . 159
- 8.2 The top figure shows a comparison of detection of alert weeks between  $WCR$  (blue triangles), CUSUM (green dots) and MEM (red crosses). The bottom figure shows the scenario used to clarify the location of the jump in rates; it shows the mean of the 100 simulated cases. . . . . 161
- 8.3 The top figure shows a comparison of detection of alert weeks between  $WCR$  (blue triangles), CUSUM (green dots) and MEM (red crosses). The bottom figure shows the scenario used to clarify the location of the increase in rates; it shows one of the 100 simulated cases. . . . . 162
- 8.4 The beginning of the system ( $N_{HB}^{**}$ ) when the cumulative distribution of the Binomial function is close to 80%, and the ending of the system ( $N_{HB}^*$ ) when the cumulative distribution of the Binomial function is close to 99.99%. 163

- 8.5 Comparison of detection of alert weeks between  $WCR_{v1}$  (blue triangles), CUSUM (green dots). The bottom figure shows the scenario used to clarify the location of the increase in rates (this scenario is based on historical data from 2009 to 2014 and the target season is the winter of 2014-2015 (Sep 2014 - April 2015)); it shows the mean of the 100 simulated cases. . . . . 165
- 8.6 Comparison of false alarms (FA) in weeks between  $WCR_{v1}$  (blue triangles), CUSUM (green dots) and MEM (red crosses); in the flat rate case with 100 simulations. . . . . 167
- 9.1 Binomial probability mass function with probability = 0.5 shown by the green columns and the actual Binomial pmf with probability = 0.4159 shown by the orange column, for 34 HBs. . . . . 178
- 9.2 Means ( $m$ ) of  $WCR$  against  $N_{HB}$  for fixed rates (labelled 1, 2, 3) and varying rates (labelled 4, 5) in table 9.3 represented as circles based on 3,110,000 simulations of ILI consultations using the steady state case. The size of circles represents the number of simulations for each  $ni$ . . . . . 180
- 9.3 Means ( $m$ ) of  $WCR$  against  $N_{HB}$  for fixed rates (labelled 6, 7, 8) and varying rates (labelled 9, 10) in table 9.3 represented as circles based on 3,110,000 simulations of ILI consultations using the steady state case. The size of circles represents the number of simulations for each  $ni$ . . . . . 181
- 9.4 Variances ( $Var$ ) of  $WCR$  against  $N_{HB}$  for fixed rates (labelled 6, 7, 8) and varying rates (labelled 9, 10) in table 9.3 represented as circles based on simulation of 3,110,000 ILI consultations using the steady state case. The size of circles represents the number of simulations for each  $ni$ . . . . . 181
- 9.5 Means ( $m$ ) of  $WCR$  against  $N_{HB}$  for fixed rates (labelled 6, 7, 8) and varying rates (labelled 9, 10) in table 9.3 represented as open circles. Coloured solid circles represent the predicted values from (z7) model. The size of circles represents the number of simulations for each  $ni$ . Despite trying more complex models with interactions and transformations, (z7) model did not fit very well as there is under prediction with high rates like 10 per 100,000. . . . . 183
- 9.6 Variances ( $Var$ ) of  $WCR$  against  $N_{HB}$  for fixed rates (labelled 6, 7, 8) and varying rates (labelled 9, 10) in table 9.3 represented as open circles. Coloured solid circles represent the predicted values from (z8) model. The size of circles represents the number of simulations for each  $ni$ . . . . . 183

- 9.7 *WCR* against  $N_{HB}$  for steady rates using fixed rate = 1 per 100,000 population with 14 HBs.  $N_{HB}^*$  is presented in the dotted vertical orange line while the dotted vertical purple line represents  $N_{HB}^{**}$ . The red shaped columns came from calculating the quantiles (cut-offs) using  $\mu$  and  $\sigma$  from many simulated cases (3,110,000) in the flat rate situation. They represent around 5% most extreme observations under the steady state. The green columns came from using `qnorm` to calculate cut-offs using  $\mu$  and  $\sigma$  from equations 9.3.1 and 9.3.2. They represent around 6% most extreme observations under the steady state. Darkness of the black colour reflects the number of simulated cases. . . . . 185
- 9.8 *WCR* against  $N_{HB}$  for steady rates using fixed rate = 1 per 100,000 population with 14 HBs.  $N_{HB}^*$  is presented in the dotted vertical orange line while the dotted vertical purple line represents  $N_{HB}^{**}$ . The red shaped columns present cut-offs using 3,110,000 simulations from the fixed rates situation, while the light blue columns present cut-offs using simulations from varying rates. . . . . 187
- 9.9 *WCR* against  $N_{HB}$  for steady rates using fixed rate = 1 per 100,000 population with 34 HBs.  $N_{HB}^*$  is presented in the dotted vertical orange line while the dotted vertical purple line represents  $N_{HB}^{**}$ . The red shaped columns present cut-offs using 3,110,000 simulated cases from the fixed rates situation, while the light blue columns present cut-offs using simulated cases from varying rates. . . . . 188
- 9.10 The top figure shows a comparison of detection of alert weeks between *WCR* (blue triangles), CUSUM (green dots) and MEM (red crosses). The bottom figure shows the scenario used to clarify the location of the jump in rates; it shows the mean of the 100 simulated cases. Here  $ng=14$ . . . . . 194
- 9.11 The top figure shows a comparison of detection of alert weeks between *WCR* (blue triangles), CUSUM (green dots) and MEM (red crosses). The bottom figure shows the scenario used to clarify the location of the gradual increase in rates; it shows the mean of the 100 simulated cases. Here  $ng=34$ . Green dotted lines represent the start and the end sensitivity area according to MEM setting while orange dotted lines represent the alert week (week of increase) as the first week when the rate is at least 10% higher than in week 1. 195
- A.1 Histograms of *WCR* for each  $N_{HB}$  in the case of HBs=14 with best fitting normal pdf. 3,110,000 simulated cases from the steady state are used here. . 205
- A.2 Histograms of *WCR* for each  $N_{HB}$  in the case of HBs=14 with best fitting normal pdf. 3,110,000 simulated cases from the steady state are used here. . 206

- A.3 Histograms of  $WCR$  for each  $N_{HB}$  in the case of HBs=30 with best fitting normal pdf. 3,110,000 simulated cases from the steady state are used here.  $WCR$  does not start from zero of  $N_{HB}$  because the joint distributions of  $WCR$  and  $N_{HB}$  in this cases do not start from zero, as shown in figure 5.2. . . . . 207
- A.4 Histograms of  $WCR$  for each  $N_{HB}$  in the case of HBs=30 with best fitting normal pdf. 3,110,000 simulated cases from the steady state are used here. . . . . 208
- A.5 Histograms of  $WCR$  for each  $N_{HB}$  in the case of HBs=30 with best fitting normal pdf. 3,110,000 simulated cases from the steady state are used here.  $WCR$  does not extend to the end of  $N_{HB}$  because the joint distributions of  $WCR$  and  $N_{HB}$  in this cases do not extend to the end of  $N_{HB}$ , as shown in figure 5.2. . . . . 209
- A.6 Histograms of  $WCR$  for each  $N_{HB}$  in the case of HBs=50 with best fitting normal pdf. 3,110,000 simulated cases from the steady state are used here.  $WCR$  does not start from zero of  $N_{HB}$  because the joint distributions of  $WCR$  and  $N_{HB}$  in these cases do not start from zero, as shown in figure 5.3. . . . . 210
- A.7 Histograms of  $WCR$  for each  $N_{HB}$  in the case of HBs=50 with best fitting normal pdf. 3,110,000 simulated cases from the steady state are used here. . . . . 211
- A.8 Histograms of  $WCR$  for each  $N_{HB}$  in the case of HBs=50 with best fitting normal pdf. 3,110,000 simulated cases from the steady state are used here.  $WCR$  does not extend to the end of  $N_{HB}$  because the joint distributions of  $WCR$  and  $N_{HB}$  in these cases do not extend to the end of  $N_{HB}$ , as shown in figure 5.3. . . . . 212
- A.9 Normal QQ plot of  $WCR$  for each  $N_{HB}$  in the case of HBs=14. . . . . 213
- A.10 Normal QQ plot of  $WCR$  for each  $N_{HB}$  in the case of HBs=14. . . . . 214
- A.11 Normal QQ plot of  $WCR$  for each  $N_{HB}$  in the case of HBs=30.  $WCR$  does not start from zero of  $N_{HB}$  because the joint distributions of  $WCR$  and  $N_{HB}$  in these cases do not start from zero, as shown in figure 5.2. . . . . 215
- A.12 Normal QQ plot of  $WCR$  for each  $N_{HB}$  in the case of HBs=30. . . . . 216
- A.13 Normal QQ plot of  $WCR$  for each  $N_{HB}$  in the case of HBs=30.  $WCR$  does not extend to the end of  $N_{HB}$  because the joint distributions of  $WCR$  and  $N_{HB}$  in these cases do not extend to the end of  $N_{HB}$ , as shown in figure 5.2. . . . . 217
- A.14 Normal QQ plot of  $WCR$  for each  $N_{HB}$  in the case of HBs=50.  $WCR$  does not start from zero of  $N_{HB}$  because the joint distributions of  $WCR$  and  $N_{HB}$  in these cases do not start from zero, as shown in figure 5.3. . . . . 218
- A.15 Normal QQ plot of  $WCR$  for each  $N_{HB}$  in the case of HBs=50. . . . . 219
- A.16 Normal QQ plot of  $WCR$  for each  $N_{HB}$  in the case of HBs=50.  $WCR$  does not extend to the end of  $N_{HB}$  because the joint distributions of  $WCR$  and  $N_{HB}$  in these cases do not extend to the end of  $N_{HB}$ , as shown in figure 5.3. . . . . 220

- B.1 *WCR* against  $N_{HB}$  for steady rates using  $r = 1$  with 14 HBs and a 5 million population.  $N_{HB}^*$  is presented in the dotted vertical orange line while the dotted vertical purple line represents  $N_{HB}^{**}$ . The red shaped area columns came from calculating the quantile (cut-offs) using  $\mu$  and  $\sigma$  from many simulated cases with a flat rate. They represent around 5% of the most extreme observations under the steady state. The green shaped area columns came from calculating **qnorm** (cut-offs) using  $\mu$  and  $\sigma$  from equations 5.4.7 and 5.4.10. They represent around 5% of the most extreme observations under the steady state. Darkness of the black colour reflects the number of simulated cases. . . . . 231
- B.2 *WCR* against  $N_{HB}$  for steady rates using  $r = 10$  with 14 HBs and a 5 million population.  $N_{HB}^*$  is presented in the dotted vertical orange line while the dotted vertical purple line represents  $N_{HB}^{**}$ . The red shaped area columns came from calculating the quantile (cut-offs) using  $\mu$  and  $\sigma$  from many simulated cases with a flat rate. They represent around 5% of the most extreme observations under the steady state. The green shaped area columns came from calculating **qnorm** (cut-offs) using  $\mu$  and  $\sigma$  from equations 5.4.7 and 5.4.10. They represent around 10% of the most extreme observations under the steady state. Darkness of the black colour reflects the number of simulated cases. . . . . 232
- B.3 *WCR* against  $N_{HB}$  for steady rates using  $r = 1$  with 34 HBs and a 5 million population.  $N_{HB}^*$  is presented in the dotted vertical orange line while the dotted vertical purple line represents  $N_{HB}^{**}$ . The red shaped area columns came from calculating the quantile (cut-offs) using  $\mu$  and  $\sigma$  from many simulated cases with a flat rate. They represent around 5% of the most extreme observations under the steady state. The green shaped area columns came from calculating **qnorm** (cut-offs) using  $\mu$  and  $\sigma$  from equations 5.4.7 and 5.4.10. They represent around 4% of the most extreme observations under the steady state. Darkness of the black colour reflects the number of simulated cases. . . . . 233



B.4 *WCR* against  $N_{HB}$  for steady rates using  $r = 5$  with 34 HBs and a 5 million population.  $N_{HB}^*$  is presented in the dotted vertical orange line while the dotted vertical purple line represents  $N_{HB}^{**}$ . The red shaped area columns came from calculating the quantile (cut-offs) using  $\mu$  and  $\sigma$  from many simulated cases with a flat rate. They represent around 5% of the most extreme observations under the steady state. The green shaped area columns came from calculating **qnorm** (cut-offs) using  $\mu$  and  $\sigma$  from equations 5.4.7 and 5.4.10. They represent around 16% of the most extreme observations under the steady state. Darkness of the black colour reflects the number of simulated cases. . . . . 234

B.5 *WCR* against  $N_{HB}$  for steady rates using  $r = 10$  with 34 HBs and a 5 million population.  $N_{HB}^*$  is presented in the dotted vertical orange line while the dotted vertical purple line represents  $N_{HB}^{**}$ . The red shaped area columns came from calculating the quantile (cut-offs) using  $\mu$  and  $\sigma$  from many simulated cases with a flat rate. They represent around 5% of the most extreme observations under the steady state. The green shaped area columns came from calculating **qnorm** (cut-offs) using  $\mu$  and  $\sigma$  from equations 5.4.7 and 5.4.10. They represent around 17% of the most extreme observations under the steady state. Darkness of the black colour reflects the number of simulated cases. . . . . 235

D.1 *WCR* against  $N_{HB}$  for steady rates using fixed rate = 3 per 100,000 population with 14 HBs.  $N_{HB}^*$  is presented in the dotted vertical orange line while the dotted vertical purple line represents  $N_{HB}^{**}$ . The red shaped columns came from calculating the quantiles (cut-offs) using  $\mu$  and  $\sigma$  from many simulated cases (3,110,000) in the flat rate situation. They represent around 5% most extreme observations under the steady state. The green columns came from using **qnorm** to calculate cut-offs using  $\mu$  and  $\sigma$  from equations 9.3.1 and 9.3.2. They represent around 3% most extreme observations under the steady state. Darkness of the black colour reflects the number of simulated cases. . . . . 239

- D.2 *WCR* against  $N_{HB}$  for steady rates using fixed rate = 5 per 100,000 population with 14 HBs.  $N_{HB}^*$  is presented in the dotted vertical orange line while the dotted vertical purple line represents  $N_{HB}^{**}$ . The red shaped columns came from calculating the quantiles (cut-offs) using  $\mu$  and  $\sigma$  from many simulated cases (3,110,000) in the flat rate situation. They represent around 5% most extreme observations under the steady state. The green columns came from using `qnorm` to calculate cut-offs using  $\mu$  and  $\sigma$  from equations 9.3.1 and 9.3.2. They represent around 4% most extreme observations under the steady state. Darkness of the black colour reflects the number of simulated cases. . . . . 240
- D.3 *WCR* against  $N_{HB}$  for steady rates using fixed rate = 1 per 100,000 population with 34 HBs.  $N_{HB}^*$  is presented in the dotted vertical orange line while the dotted vertical purple line represents  $N_{HB}^{**}$ . The red shaped columns came from calculating the quantiles (cut-offs) using  $\mu$  and  $\sigma$  from many simulated cases (3,110,000) in the flat rate situation. They represent around 5% most extreme observations under the steady state. The green columns came from using `qnorm` to calculate cut-offs using  $\mu$  and  $\sigma$  from equations 9.3.1 and 9.3.2. They represent around 6% most extreme observations under the steady state. Darkness of the black colour reflects the number of simulated cases. . . . . 241
- D.4 *WCR* against  $N_{HB}$  for steady rates using fixed rate = 3 per 100,000 population with 34 HBs.  $N_{HB}^*$  is presented in the dotted vertical orange line while the dotted vertical purple line represents  $N_{HB}^{**}$ . The red shaped columns came from calculating the quantiles (cut-offs) using  $\mu$  and  $\sigma$  from many simulated cases (3,110,000) in the flat rate situation. They represent around 5% most extreme observations under the steady state. The green columns came from using `qnorm` to calculate cut-offs using  $\mu$  and  $\sigma$  from equations 9.3.1 and 9.3.2. They represent around 7% most extreme observations under the steady state. Darkness of the black colour reflects the number of simulated cases. . . . . 242

- D.5 *WCR* against  $N_{HB}$  for steady rates using fixed rate = 5 per 100,000 population with 34 HBs.  $N_{HB}^*$  is presented in the dotted vertical orange line while the dotted vertical purple line represents  $N_{HB}^{**}$ . The red shaped columns came from calculating the quantiles (cut-offs) using  $\mu$  and  $\sigma$  from many simulated cases (3,110,000) in the flat rate situation. They represent around 5% most extreme observations under the steady state. The green columns came from using `qnorm` to calculate cut-offs using  $\mu$  and  $\sigma$  from equations 9.3.1 and 9.3.2. They represent around 0.7% most extreme observations under the steady state. Darkness of the black colour reflects the number of simulated cases. . . . . 243

# List of Tables

1.1	Occurrences of dangerous influenza subtypes H5 and H7 (HPAI) in different countries from 1959 [3]. . . . .	5
2.1	Abbreviations of names of Scottish health boards. . . . .	15
2.2	Deviance explained for each GAM of ARI rates in Scotland, for each health board and year. . . . .	28
3.1	Short list of advantages, disadvantages and examples of when and where the detection systems are used. . . . .	47
4.1	Pre-epidemic and post-epidemic thresholds based on historical data (ARI) until 2013-2014, and comparison with the observed 2014–2015 season; $n$ denotes number of cases and $w$ are weeks. . . . .	53
4.2	Pre-epidemic and post-epidemic thresholds based on historical data (ILI) until 2013-2014, and comparison with the observed 2014–2015 season; $n$ denotes number of cases and $w$ are weeks. . . . .	53
4.3	Means of $WCR$ in 2010 (April to September) for each health board; higher values are shown in bold. . . . .	68
4.4	Means of $WCR$ in 2011 (April to September) for each health board; higher values are shown in bold. . . . .	69
4.5	Means of $WCR$ in 2012 (April to September) for each health board; higher values are shown in bold. . . . .	69
4.6	Means of $WCR$ in 2013 (April to September) for each health board; higher values are shown in bold. . . . .	70
4.7	Means of $WCR$ in 2014 (April to September) for each health board; higher values are shown in bold. . . . .	73
4.8	Means of $WCR$ in 2010/11 (October to March) for each health board; higher values are shown in bold. . . . .	73
4.9	Means of $WCR$ in 2011/12 (October to March) for each health board; higher values are shown in bold. . . . .	74

4.10	Means of $WCR$ in 2012/13 (October to March) for each health board; higher values are shown in bold. . . . .	77
4.11	Means of $WCR$ in 2013/14 (October to March) for each health board; higher values are shown in bold. . . . .	78
4.12	The Kolmogorov-Smirnov test $p$ -values for tests of $WCR$ in each summer period against the corresponding winter period. . . . .	81
4.13	The Anderson-Darling $p$ -values for the normality test of $WCR$ in each summer period and winter period. . . . .	82
4.14	The Kolmogorov-Smirnov test $p$ -values for tests of $N_{HB}$ in each summer period against the corresponding winter period. . . . .	82
5.1	Population structures considered. . . . .	87
5.2	Approaches to choosing different baselines for the simulation. . . . .	92
5.3	Short summary of models $m14$ , $m30$ and $m50$ for the mean of $WCR$ . . . . .	93
5.4	Short summary of model $sd14$ , $sd30$ and $sd50$ for the standard deviation of $WCR$ . . . . .	96
5.5	Short summary of models $sdBC14$ , $sdBC30$ and $sdBC50$ for the transformed standard deviation of $WCR$ . . . . .	97
5.6	General framework used for modelling $\mu$ and $\sigma^2$ for each HBs = 14, 30, 40, 50. The reason that $\mu$ and $\sigma^2$ for HBs= 30, 40, 50 do not continue to the end of HBs is because the joint distributions of $WCR$ and $N_{HB}$ in those cases do not continue to the end of HBs, as shown in figures 5.2, 5.3, 5.10 and 5.11. . . . .	103
5.7	Short summary of model (z1) for the mean of $WCR$ . . . . .	103
5.8	Short summary of model (z2) for the mean of $WCR$ . . . . .	104
5.9	Short summary of model (z3) for the mean of $WCR$ . . . . .	104
5.10	Slopes in the model (z3) for the mean $WCR$ for each $ng$ . . . . .	105
5.11	Short summary of model (z4). . . . .	106
5.12	Short summary of model (z5) where $var$ represents the variance of $WCR$ in each $N_{HB}$ and $n.i.midway$ represents $N_{HB}$ starting from the mid-point to the end of HBs. Weights reflects the number of simulated cases in each $N_{HB}$ and $ng$ reflects HBs (14, 20, 30, 40, 50). . . . .	110
5.13	Short summary of model (z6) for the variance of $WCR$ . . . . .	110
5.14	Intercepts of the model (z6) of the variance of $WCR$ for each value of $ng$ . . . . .	111
5.15	Short summary of model (z7). . . . .	112

6.1	Percentage of simulations above the cut-offs for $WCR$ , using different HBs, from the total number of simulations. Here, limits are calculated from 3,110,000 simulated cases. The bold text corresponds to the actual number of HBs (40). . . . .	125
6.2	Percentage of coverage in the rejection region, using different approaches and the total number of simulations (3,110,000). . . . .	133
6.3	Details of key values for 1 to 14 HBs weekly total and their populations (more details in the text). . . . .	135
6.4	Example to show how the ratio $WCR$ is distributed in 10,000 simulations from a Poisson ( $\theta$ ) pmf. . . . .	137
6.5	Percentage of coverage of the rejection region using the total number of simulations (3,110,000), where $r$ is the scaling factor used. "simulation" means the cut-offs were based on the actual simulated cases, while "equation" means that the cut-offs were calculated based on mean and variance equations 5.4.7 and 5.4.10. . . . .	139
8.1	Comparing performance of $WCR_{v1}$ with CUSUM and MEM in different scenarios, Sen. represents sensitivity, Sp. represents specificity, $MED$ represents median time to detection, F means that the method failed to raise an alarm. . . . .	166
8.2	Comparing performance of $WCR_{v2}$ and $WCR_{v3}$ in different scenarios, Sen. represents sensitivity, Sp. represents specificity, $MED$ represents median time to detection. . . . .	170
8.3	Comparing performance of $WCR_{v1}$ with CUSUM and MEM in different scenarios, Sen. represents sensitivity, Sp. represents specificity, $MED$ represents median time to detection, F means that the method failed to raise an alarm. . . . .	172
8.4	Comparing performance of $WCR_{v2}$ and $WCR_{v3}$ in different scenarios, Sen. represents sensitivity, Sp. represents specificity, $MED$ represents median time to detection. . . . .	173
9.1	Example for the number of cases in each $N_{HB}$ with $ng=14$ . The total number of simulated cases is 3,110,000. This is based on steady state simulation.	177
9.2	Actual mean of $N_{HB}$ divided by $ng$ when $ng=14$ and 34. These came from 3,110,000 simulations of ILI cases in the steady state case. . . . .	177
9.3	Illustration of parameters of each approach of simulation. . . . .	179
9.4	Short summary of model (z7). . . . .	182
9.5	Short summary of model (z8). . . . .	182

- 9.6 Percentage of coverage of the rejection region using the total number of simulations (3,110,000), where  $fr$  means fixed rate, “simulation” means the cut-offs are based on the actual simulated cases (3,110,000), while “equation” means that the cut-offs are calculated based on mean and variance equations 9.3.1 and 9.3.2. The simulated cases did to calculate the cut-off and then these are repeats with further simulations using the same cut-off. . . . . 186
- 9.7 Percentage of coverage of the rejection region using the total number of simulations (3,110,000), where  $fr$  means fixed rate, “simulation” means the cut-offs are based on the actual simulated cases (3,110,000), while “equation” means that the cut-offs are calculated based on mean and variance equations 9.3.1 and 9.3.2. The first set of simulated cases were used to calculate the cut-off and then these were repeated with further simulations using the same cut-off. . . . . 186
- 9.8 Comparing performance of  $WCR$ , when  $ng=14, 34$  and  $fr = 1, 3, 5$  with  $\alpha = 5\%$ , with CUSUM and MEM in different scenarios, using 100 simulations. Sen. means sensitivity, Sp. means specificity,  $MED$  means median time to detection, and E.S. (epidemic start) defines the approach of using the alert week (here the alert week is considered as the first week when the rate is at least 10% higher than in week 1). . . . . 190
- 9.9 Comparing performance of  $WCR$  when  $ng=14, 34$  and  $fr = 1, 3, 5$  with  $\alpha = 5\%$ , with CUSUM and MEM in different scenarios, using 100 simulations. Sen. means sensitivity, Sp. means specificity,  $MED$  means median time to detection, and E.S. defines the approach of using the alert week (here the alert week is considered as the same as in the MEM setting). . . . . 192
- 9.10 Comparing performance of  $WCR$ , when  $ng=14, 34$  and  $fr = 1, 3, 5$  with  $\alpha = 1\%$ , with CUSUM and MEM in different scenarios, using 100 simulations. Sen. means sensitivity, Sp. means specificity,  $MED$  means median time to detection, and E.S. defines the approach of using the alert week (here the alert week is considered as the first week when the rate is at least 10% higher than in week 1). . . . . 196
- 9.11 Comparing performance of  $WCR$  when  $ng=14, 34$  and  $fr = 1, 3, 5$  with  $\alpha = 1\%$ , with CUSUM and MEM in different scenarios, using 100 simulations. Sen. means sensitivity, Sp. means specificity,  $MED$  means median time to detection, and E.S. defines the approach of using the alert week (here the alert week is considered as the same as in the MEM setting). . . . . 199
- C.1 Week locations and numbers during the year, where  $W$ = week number (1-52), and  $WL$ : week location in the year (31-52) and (1-30), and  $WB$ : week begin date. . . . . 236

# Chapter 1

## Background about influenza

This chapter contains background information about influenza, while the next chapter 2 is about description of data. In chapter 3, statistical methods for the detection of outbreaks are reviewed. Applying and modifying some well-known detection systems for influenza is presented in chapter 4. Chapters 5 and 6 give an overview of the system. Different scenarios for the simulation study are reviewed in chapter 7. Results of tests of the detection system and comparison with other systems are presented in chapters 8 and 9. Chapter 10 draws conclusions.

In this chapter, in sections 1.1, 1.2 and 1.2.1, there are a definition and description of influenza, transmission and classification of influenza, while in section 1.3 the symptoms are discussed. During a pandemic or seasonal influenza case, people should take precautions regarding transmission, and this is reported in section 1.2. Section 1.4, reports history of influenza, while the burden of influenza is discussed in section 1.5. Sometimes a virus becomes out of control and is reported as a pandemic or seasonal influenza, respectively discussed in section 1.6. Section 1.7 discusses vaccination against influenza and treatment after infection is reported in section 1.8.

### 1.1 What is influenza?

Today, there are a lot of disasters that threaten humans in their lives. Wars are one of the obstacles of life. There is a kind of war that does not kill people by using bullets and guns, but by the spread of viruses in the air. It is “viral diseases”.

According to Whimbey et al. [175], respiratory viruses are more active in winter than other seasons. Also influenza accounts for nearly 50% of virus diseases. The first ap-



pearance of an influenza virus was in 1878 in South Asia, but there is no detailed report regarding it [3,148,171]. It is an indication of the seriousness of influenza that pandemics ravaged human beings in the past for example, 1918, 1957–1958, 1968–1969 and 2009 pandemics [153]. Because influenza transmits quickly between infected and healthy humans (direct infection) as well as between infected animals and humans (non-direct infection), it is important to take advantage of science and research methods that help to discover influenza before an epidemic or pandemic occurs, to be prepared for taking protective vaccines before widespread infection. Early detection of the emergence of influenza each year is important for planning and treatment.

Humans are still trying to study treatments and preventive vaccinations for viruses. Diseases caused by viruses have been known as famous diseases from the past. Respiratory viruses are viruses which are more active in winter than other seasons [77,97,175]. Also, influenza accounts nearly for 50% of the viral diseases [152,154]. Influenza is one of the well-known illnesses which infect humans through influenza viruses. The first appearance of an influenza virus was in 1878 in South Asia, but there is no detailed report regarding it. It was for a young girl who lived near to poultry and she was infected, then she died. After that, her mother was infected. Actually, there is evidence for the virus moving from poultry to humans (non-direct transmission) and also from human to human (direct transmission) [3,24,87]. Another kind of source of a virus non-directly transmitted from animals is the virus reported to infect humans through pigs [171].

## 1.2 Transmission

A first step to combat spread of the disease when it occurs in the community is full knowledge and awareness about the modes of transmission of the disease. The influenza virus is able to continue in the population because it is capable of continuing evolution in mutations, and there are significant changes in the composition of the virus, leading to epidemics that can cause morbidity and mortality in a global catastrophe [18,31]. People infected with influenza in the context of health care and hospitals pose a threat to workers in hospitals, as well as inpatients, in what can lead to the rapid spread of the virus and its outbreak in a community. The opinion of many health workers is that the solution to this is to isolate a patient with the virus. A humid environment is considered capable of transmitting influenza viruses and of ensuring the survival of the virus alive if there is not enough ventilation. It is also important to mention that desiccation of the virus does not eliminate infectivity [172]. Non-pharmaceutical interventions, including facemasks, improved hand hygiene, cough etiquette, isolation of the sick and quarantine of exposed individuals, social distancing measures, and travel restrictions, have been used

by many countries to control an influenza pandemic, at its beginning especially. However for pandemic or non-pandemic influenza, concerning modes of influenza virus transmission among people, there is no strong evidence of the validity of simple personal protective measures, such as facemasks and hand hygiene [36, 149]. Doctors can clearly diagnose influenza in patients by recording signs and symptoms soon after fever starts and using a clinical case definition and knowledge to know that the influenza virus is spreading within the community [8]. Also the age of humans is an important factor in infection. One study showed that 1% and 26% of people who are older than 18 and younger than 64 are infected with influenza annually [23]. A community should know the ways that human influenza is transmitted, to protect itself, families and their children. Human influenza can infect people by direct contact through inhalation of infectious droplets. In 1997, one infection was reported for influenza A (H5N1) during daily contact between a mother and her child. During the time that the influenza virus is alive in the environment, there are many ways for more infection to spread. From the first thinking about and understanding of influenza, it has been known that infected (contaminated) water can keep the influenza virus alive and transmit the infection directly during swimming or by breathing spray from infected water [151, 181]. In hospitals, diagnosing influenza virus infection is more conservative, and less defensive detection methods such as antigen detection and virus isolation are used, with less accuracy [158]. The best solution to avoid the disease influenza is by taking preventive vaccinations against the virus, as well as to stay away from infected people and crowded places without ventilation.

### 1.2.1 Classification of influenza

In 1995, influenza viruses were classified to HA and NA groups. HA means “haemagglutinin” and NA means “neuraminidase”. Only the influenza A virus was known from birds, but also viruses of all 15 (H1, H2, . . . , H15)(HA) and 9 (N1, N2, . . . , N9) (NA) influenza subtypes have been taken from avian species. Classification of influenza viruses gives us individual characteristics for some influenza subtype viruses. Two subtypes H5 and H7 are challenging because of their near 100% chance of causing “Fowl Plague”, “highly pathogenic avian influenza” (HPAI). Other viruses are much less virulent than H5 and H7 for respiratory disease, denoted “low pathogenicity avian influenza” (LPAI) [3, 159].

## 1.3 Symptoms

Symptoms or signs of any disease or illness can be defined as indicators which can be diagnosed by a doctor as evidence of illness or infection. Some cases need specific analysis to confirm illness or infection. The duration between infection and manifestation of disease

differs according to some factors like kind of illness, age and patient immunity. Usually, the first point of diagnosis of any disease is with the general practitioner (GP), as most people are registered with a GP [101]. Many studies in the past and at the current time collect their data from GP records, as the diseases in the world are preceded by symptoms and signs which give direct or non-direct evidence of a specific ailment. Influenza has a fixed cycle of indicator symptoms. The combination of fever and cough has been shown to be an accurate symptom of influenza infection in the influenza season. In one paper about influenza symptoms, patients were monitored daily by nurses for the presence of a cough (more than usual or acute onset) and elevated body temperature (37.8 C and over). Other symptoms that might be associated with influenza are sore throat, rigours or chills, myalgia, headache, obstructed nose or inflammation of the membrane of the nose, acute onset of symptoms, and malaise or weakness. Cough is one of the influenza symptoms with the highest sensitivity compared with the other symptoms, because influenza symptoms are not all equally common [96,158]. Monitoring GP records can indicate the presence of influenza in a community from the frequency of relevant symptoms in people which are highly indicative of influenza, i.e. serious cough and fever (body temperature 37.8 C or over). Also cough and fever are the best two predictors of a laboratory-confirmed better diagnosis of influenza. Many studies show that cough and fever are a better indicator of influenza infection than either syndrome alone, e.g. [20,96]. From those studies, there is no difference in influenza symptoms between different subtypes of influenza viruses. Influenza disrupts the daily activities of individuals because most people are incapacitated by this disease, which be a considerable burden to society, e.g. causing many days of lost work, or death in more vulnerable people.

## 1.4 History of influenza

Some influenza viruses have been known for some time, such as H1, H2 and H3, but this does not mean there are only three viruses which can infect humans. In 1878, one disease known as “Fowl Plague” caused mortality for humans and in 1901, the main cause for this disease was known but not in detail until 1995, because in this year influenza viruses were classified to HA and NA groups. In 1959, a first case regarding HPAI was observed in poultry and reported 17 times for five turkeys and twelve chickens, caused by the H5 influenza subtype [3,155].

Influenza A subtype H7 was reported in England in 1991. The rate of virus spread was different between virus locations, as the reports show that the rate in England in 1991 was much less than others in the USA in 1983, Mexico in 1994 and Pakistan in 1995. Because the virus was reported in different geographical locations, the disease become widespread,

causing large number of birds to die or be killed as part of a control policy [3].

Table 1.1 shows some occurrences of different strains of influenza since 1959.

Reported HPAI isolates from poultry since 1959				
No	Region	Kind of poultry	Kind of Influenza	Subtype of virus
1	Scotland	Chicken	A	H5N1
2	England	Turkey	A	H7N3
3	Ontario	Turkey	A	H5N9
4	Victoria	Chicken	A	H7N7
5	Germany	Chicken	A	H7N7
6	England	Turkey	A	H7N7
7	Pennsylvania	Chicken	A	H5N2
8	Ireland	Turkey	A	H5N8
9	Victoria	Chicken	A	H7N7
10	England	Turkey	A	H5N1
11	Victoria	Chicken	A	H7N3
12	Queensland	Chicken	A	H7N3
13	Mexico	Chicken	A	H5N2
14	Pakistan	Chicken	A	H7N3
15	NSW	Chicken	A	H7N4
16	Hong Kong	Chicken	A	H5N1
17	Italy	Chicken	A	H5N2

Table 1.1: Occurrences of dangerous influenza subtypes H5 and H7 (HPAI) in different countries from 1959 [3].

In 1997, the influenza virus H5 killed 6 of 18 infected humans. Some kinds of influenza virus infect people non-directly, like influenza virus which comes from animals in the human environment, such as birds or pigs [144]. People believed that avian influenza could not move outside birds and infect humans. This event clearly led scientists to two main causes for avian influenza, i.e. live poultry markets as breeding grounds for the influenza virus, and also an Asian source of the H5 influenza virus apart from Hong Kong severe acute respiratory syndrome (SARS). Influenza viruses like H1, . . . , H15 have subtypes. H5N1 virus was found in duck meat imported from China between 1997 and 2003. A complication in this problem was that China did not have official reporting for the H5N1 virus. During such events, countries must show the public that they are taking steps to contain the virus [118, 171]. Another kind of influenza is H7, which needs lab analysis to create an antivirus. One serious problem in this situation is that H5 and H7

are very fast spreading viruses which lead to quick infection between chickens and turkeys before moving to infect humans. H7 and H9 subtypes can also cause influenza in humans. In the past ten years from 2003, influenza arrived in humans because both H6 and H9 influenza viruses had spread from a wild aquatic bird reservoir. One subtype of influenza virus called H9N2 infected both humans and pigs [118,171].

One avian influenza virus called H5N1 infected humans through birds. In 2009, H5N1 killed humans in Southeast Asia, and the first human death was in Indonesia [151]. In Asia, one kind of influenza virus is widespread, called influenza A subtype (H9N2) [119]. In 2002, the global spread of SARS caused the public to become keenly aware of emerging infectious disease. In 2003, highly pathogenic strains of avian influenza virus, also H5N1 and H7N7 subtypes, were included among those strains which can pass to humans. Before 1997, direct avian to human influenza transmission was unknown. H5N1 (avian influenza) can move to infect humans and this happened in February 2003 when a young girl from Hong Kong died of an unidentified respiratory illness after she visited mainland China [42]. After that the fatal illness passed to her father, showing that this virus can also move from human to human. In March 2003, another alarming situation with a highly pathogenic H7N7 avian influenza outbreak occurred at the other side of the world, in the poultry industry of the Netherlands. Many veterinarians died of respiratory infection. Fortunately, the worst case scenarios did not come about in either of these 2003 avian influenza virus outbreaks.

Doctors and researchers are now better equipped with technologies and reagents to rapidly identify and respond to pandemic influenza threats, since 1997. This development increases the responsibility of scientific communities to use those improvements to stop hearsay and give public and nations accurate knowledge regarding influenza, means of infection and also ways to avoid infection. The World Health Organization (WHO) started the development of a Global Agenda for influenza surveillance and control in 2001. In May 2002, a document written by WHO was adopted after a proposal was made and public comment was invited. This document advocates developing and improving methods and reagents that can be used to rapidly identify all influenza virus subtypes, as part of integrated influenza surveillance in humans and animals. The WHO with its global influenza network works in international cooperation for the rapid exchange of viruses, reagents and information to improve early rapid surveillance for all influenza virus subtypes [171].

## 1.5 Burden of influenza

From the results of our literature search we conclude that the influenza virus occupies a large number of scientific studies, as well as being a major subject of study in the statistical work of health organizations in the world. Winter is the season when preparations must be made before the spread of the virus is greatest. Several studies have discussed the

issue of global influenza virus, confirming that influenza virus is an annual major concern in public health and is responsible for an average of 36,000 excess deaths annually. Also 114,000 hospitalization cases per year are caused by the influenza virus [84, 85, 108, 172].

## 1.6 Pandemics, seasonal influenza

There is a famous and straightforward definition for a pandemic, accepted internationally, written in the Dictionary of Epidemiology: “an epidemic occurring worldwide, or over a very wide area, crossing international boundaries and usually affecting a large number of people” [171]. Influenza outbreaks differ between geographical locations, meaning an outbreak can happen in one country over a limited time period. In 1995, community surveillance by the Scottish Centre for Infection and Environmental Health monitored outbreaks of influenza A and B starting in week 6 and 1 respectively [126]. Sometimes pandemic influenza happens as outbreaks like avian influenza A (subtype H5N1), which infected poultry and moved to humans in Southeast (SE) Asia. At first, infection is caused by contact between birds and humans with limited spread from human to human, but then infection spreads between humans after a human is infected, which can lead to an outbreak then spread worldwide via the global transportations network. This happens more quickly than an adequate supply of the vaccine matching the new version can be produced. This is the main reason for building models for rapid detection by researchers and scientists [9, 55, 86, 143], to create more time to manage the outbreak. Pandemic influenza occurs unpredictably, such as the H7N7 virus in European poultry and humans, the H5N1 virus which infected Asian poultry and humans, and also swine flu which infected pigs in the United States [171]. The main challenge during pandemic influenza is the time factor, because manufacturing and developing a vaccine to match the new identified influenza takes 6-8 weeks [85]. Infections also depend on some factors like time, age, and location, because many studies observed that a pandemic during the winter season has the highest level of infection. Also during non-pandemic influenza years, people who are younger than 65 years have 10% less influenza illness according to GP consultations [23].

To conclude, at a time when we see influenza come and infect humans, then kill them, unexpectedly, part of a solution should be storing treatments and generic drugs for use at the first attack of influenza. One treatment that it is advised to store is “Oseltamivir” [151].

Another kind of influenza infection in the community occurs in the winter, which is called “seasonal influenza”. The influenza virus spreads around the world in winter, and many studies confirm that the major season for influenza infection is in the winter. The strains are usually different from one year to the other. It is possible to get infected by the influenza when it is not influenza season, but this is less likely. An average of 10% of the population suffers from annual influenza in the world during the weeks of an influenza

outbreak [8].

One of the signs that many cases of influenza occur when weather temperature decreases is that numbers of recorded deaths rise in winter. This sign leads scientists and health organizations in the world to take care and prepare to face this season to avoid poorer health scenarios [95]. There is clearly oscillation in occurrence of influenza, as recorded by many studies. One reason reported by the American Journal of Clinical Nutrition, is acute decrease of vitamin D, because the main source for vitamin D is sun exposure; thus, season can affect serum vitamin D concentrations [157].

### 1.6.1 Geographic pattern

Often the spread of viral diseases varies from one country to another depending on the nature of the atmosphere and weather in those states. According to WHO, global seasonal influenza activity is different depending on geographical location. It is more active and continuing in the southern hemisphere, while elsewhere influenza activity stays low.

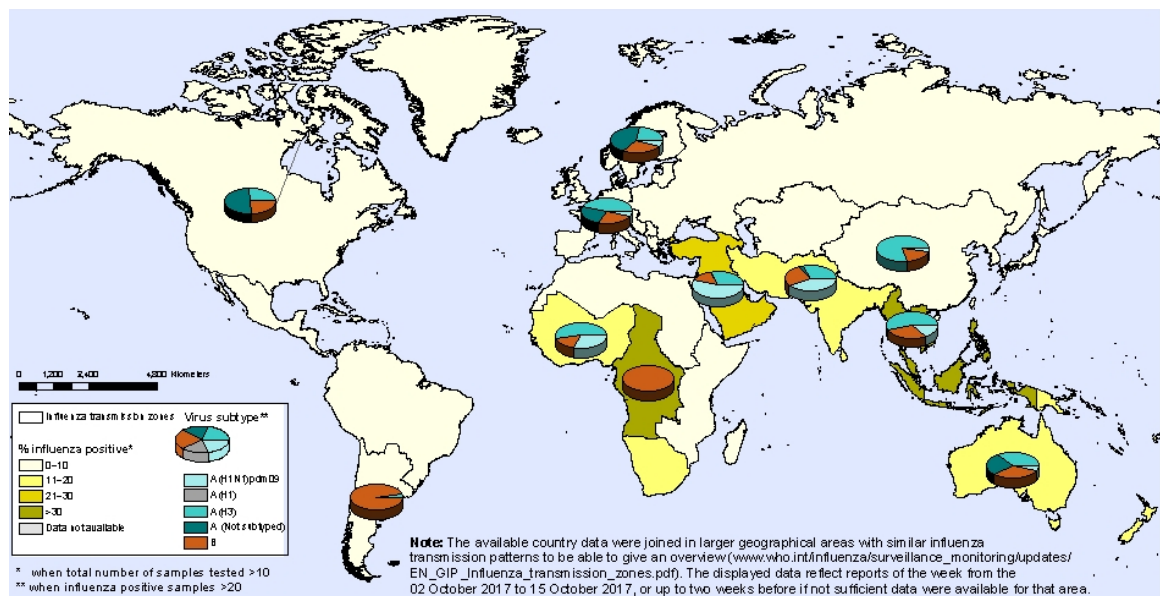


Figure 1.1: Influenza map [182]

The map in figure 1.1 comes from WHO, and shows the last update on 20th October 2017 Update number 301 [182] notes some points, which are that over the study period:

- Influenza and respiratory syncytial virus (RSV) activity took a downward trend throughout most of the sub-regions in conservative South America.
- Influenza activity continued to decrease, with influenza B viruses most frequently detected in Southern Africa.
- In Oceania, seasonal influenza activity continued to drop, with influenza A(H3N2)

triumphant, followed by B viruses.

- Influenza and RSV activity remained at low levels overall in the tropical countries of South America.
- Respiratory illness indicators in the Caribbean and Central American countries and influenza activity stayed low in general, but RSV activity remained high in several countries.
- Influenza activity remained low in general in Southern Asia. In India, influenza A(H1N1)pdm09 and A(H3N2) virus detections continued to be reported.
- In Western Asia, there was increasing influenza activity in Oman, with influenza A(H1N1)pdm09 virus predominantly detected, followed by a small proportion of A(H3N2) and B viruses.
- Influenza activity remained low in general in East Asia.
- In Western and Middle Africa, influenza detections continued to be reported, with all seasonal influenza subtypes present in the region. In Eastern Africa, little influenza activity was reported, with the exception of Réunion Island.
- Little to no influenza virus detections were reported in North Africa.
- In Central Asia, Influenza-like illness (ILI) and severe acute respiratory infection (SARI) indicators appeared to increase in Kazakhstan, Tajikistan and Uzbekistan, with few influenza detections.
- Influenza activity remained low in Europe.
- Overall, influenza virus activity remained low, with detections of predominantly influenza A(H3N2) and B viruses in North America.

## 1.7 Vaccination

First, it is usual to face challenges for development of any vaccines against any virus such as the influenza virus. This because influenza viruses have many subtypes, and developing any antiviruses requires study of all specifications and life cycle for this virus subtype. On the other hand, we also need to study and know all impacts, effectiveness, positive and negative effects and duration of effectiveness for humans who are infected with the virus. Those processes must be done as soon as possible from the first information about the virus, in order to face it and also to avoid virus spread. In the history of vaccines since the early 1980s, developed countries established large-scale influenza vaccination programmes, now starting in developing countries [31, 63]. Regarding the influenza virus,



some challenges can be faced by early knowledge about the subtypes, from the time of first emergence. If the challenge is time, e.g. the next influenza pandemic were to begin tomorrow, using a vaccine will confine the disease, which means the possibility for the infection to infect people will be reduced. Now, there are tissue culture – based and live debilitates vaccines authorized and allowable in some countries, which could complement the supply of inactive vaccines [171]. People in some age groups need more health care, like older people or small children. Persons of 65 years old or more are advised by the Advisory Committee on Immunization Practices (ACIP) of the Public Health Service to receive the influenza vaccine annually.

More recent studies have shown reduced levels of influenza infection and death of people who receive vaccine. Results in the study by Nichol et al. [110] found that elderly people living in the community had less infection than people in hospitals during the influenza seasons, while they all took the vaccine against influenza [14, 99]. That is the reason for doctors advising people to take the vaccine against influenza before they have any symptoms. The current recommendation from ACIP to each person with increased risk for complications of influenza or who would like to avoid influenza is to have annual vaccination against influenza [109]. Also, the Centre for Disease Control and Prevention (CDC) in the United States and the Department of Health in the United Kingdom (UK) strongly advise the influenza vaccine for people who have chronic illnesses or who are in long-term care homes. It is a serious issue because comparison of elderly people who receive the vaccine with those who refuse the vaccine shows a reduction in infection and mortality for vaccinated people [126]. Developments in medical devices and advanced technology make it possible for doctors and people who are interested in public health to do more work and research to give humanity successful solutions for influenza disease. One third of people who are infected require hospitalization, and one tenth die from influenza. Treatment and prophylaxis with antiviral drugs during influenza outbreaks in healthcare institutions are recommended by ACIP. Rapid diagnostic tests and efficient monitoring systems are very important because the first introduction of control measures depends on them [158]. To treat influenza effectively requires taking antivirals directly with no waiting for traditional laboratory diagnosis [96]. However valuable advice for the public is that influenza vaccine is recommended, especially for persons of a certain age [23].

## 1.8 Treatments

There is clearly a need for effective treatments that can reduce influenza impact on society and individuals. During influenza, using antivirals as early as possible can help to control the influenza infection and to prevent the infection of uninfected people. Amantadine and

Rimantadine are two currently available antiviral agents for the treatment of influenza, but are not widely used because of limited availability in some countries, and there is also concern about side effects of those treatments. They are also effective against influenza subtype A alone, which leads to problems when a mixed outbreak occurs, so it is clearly important to look for antiviral agents able to face more subtypes, which can be used with different cases. An antiviral agent to face the previous challenge, the treatment of influenza A and B, is called Zanamivir. The advantage of using Zanamivir is that it reduces the time over which symptoms appear by about 3 days, and studies also dramatically illustrate a clear response from the second day when using this treatment [95]. Another oral prodrug of oseltamivir carboxylate can treat influenza A and B, called oseltamivir, and it also has fast response from symptom onset. Taking oral oseltamivir as early as possible after the onset of influenza symptoms can start the effectiveness after 24 hours from receiving the drug. This is called the IMPACT (IMmediate Possibility to ACcess oseltamivir Treatment). Educating the community about the symptoms of flu that everyone can observe is very important to start disbursement procedures for appropriate treatment early on. Early intervention reduces the total duration of the disease by half or more compared to waiting and treating at a later time, resulting in faster recovery and resumption of normal activities [8]. Occurrence of a primary infection in children is associated with higher rates of viral replication, owing to a lack of previous immunity, so there is a difference in resistance rates between children and adults [10, 13, 41].

## 1.9 Aims and objectives

This thesis concerns the development and testing of statistical methods for detecting outbreaks of influenza.

The main aim of this research is trying to develop the *WCR* method for use to detect seasonal flu by establishing a simpler method of calculating the critical values (cut-offs). We will start by reviewing and studying various well-known methods used for early warning systems and control of infectious diseases, and to try to improve on these.

Also the aim is to investigate trends in the GP consultation rates over time and to compare the patterns in the different health boards, to develop an extended *WCR* method to detect the next influenza epidemic and to study *WCR* performance compared to CUSUM and MEM under situations of an increase in different circumstances. Such an extended method would start to detect the beginning of the seasonal flu epidemic each year and to give as much advance warning as possible.

The focus in this project is finding a simple way to extend the *WCR* method and compare it to other well established systems.

## 1.10 Conclusion

This first chapter of this thesis has discussed the definition, transmission, symptoms, history and burden of influenza, as well as pandemics and seasonal influenza, vaccination and treatments of influenza.

Chapter 2 now examines ILI and ARI data for Scotland, which will be used as a basis for the work in the rest of the thesis.

## Chapter 2

# Data Description

As influenza is such a serious respiratory virus, most world countries have surveillance systems to try and detect when the virus is circulating in their communities. This is often based upon consultations for respiratory symptoms at GP practices. Respiratory viruses like influenza infections are the most widespread causes of primary care consultation and appear to be a serious economic burden worldwide [35, 92]. The annual attack rate of influenza is estimated at 5–10% in adults and 20–30% in children in a typical season, although not all cases request medical care and are captured by the surveillance systems [105, 183]. It is widely admitted that surveillance of influenza should address the following objectives: monitoring the circulating virus strains in the early stages, the timing, strength and riskiness of the epidemic patterns, providing information about the implicit risk conditions, as well as provisioning epidemiological and virological support for a pandemic at an early stage [180].

In this chapter, we investigate daily data for Acute Respiratory Illness (ARI) and influenza like illness (ILI) consultations for surveillance for influenza in Scotland, as they give signs or evidence about influenza infection [55, 106, 140, 184], and are used as a basis for the later work in this thesis. This chapter uses data description and modelling for data familiarisation.

### 2.1 Scotland Data

We are using data collected from all health boards in Scotland. There are fourteen health boards which are spatially located. See figure 2.1 for these administrative areas. The health boards are responsible for protecting and improving the health of the population, as well as the provision of health care services such as the National Health Service (NHS) [102].

The NHS provided the data from 2000 to 2014. The Department of Mathematics and Statistics in the University of Strathclyde has formal collaboration which gives access to these data and they do not require ethical approve as the data are aggregated from GPs and these data are anonymised.

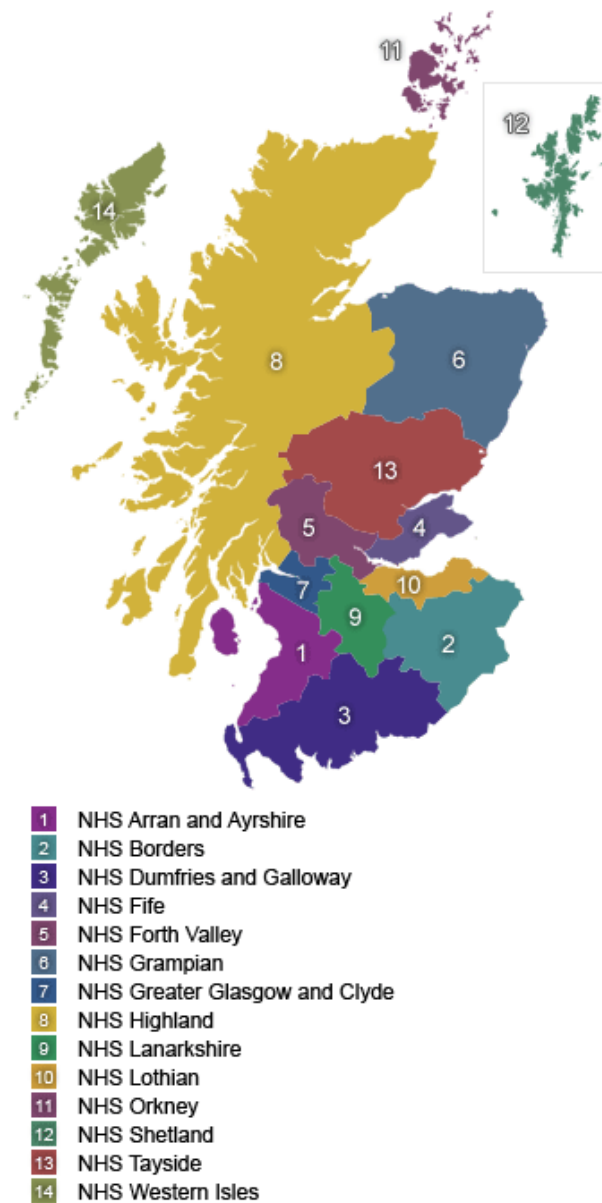


Figure 2.1: Map of the territorial health boards in Scotland [66]

Table 2.1 shows a list of HBs and their abbreviations used in this thesis. The National Health Service (NHS) provided the data on the ILI and ARI consultations from the mid-year of 2009 to the mid-year of 2015. The aim of this analysis is to investigate the trends in the GP consultation rates over time and to compare the patterns in the different health boards.

Health Board	Abbreviation
Ayrshire and Arran	AA
Borders	BR
Dumfries and Galloway	DG
Fife	FF
Forth Valley	FV
Grampian	GG
Greater Glasgow and Clyde	GGC
Highland	HI
Lanarkshire	LN
Lothian	LO
Orkney	OR
Shetland	SH
Tayside	TY
Western Isles	WI

Table 2.1: Abbreviations of names of Scottish health boards.

## 2.2 Methods

We use the software R [129] to analyse NHS consultation rates for ILI and ARI. We use Poisson regression to model the data as the numbers of consultations are considered as count data and because the counts are small compared to the total population size in each HB. We also use a generalised additive model (GAM) model with spline time functions each year to describe the seasonal trends within a year. The second method used to summarise the seasonal trend is to use a fourteen day moving average model. Finally, we will compare our GAM models for each health board every year to investigate the similarity of the seasonal trend in each health board [194].

## 2.3 Data

We treat the data for each health board separately, then run a 12 months' time series analysis for each of them. Number of GP consultations for ILI and ARI by the population are available on a daily basis for complete years starting from August 2009 until the end of July 2010 and so on until the end of July 2015. There are 6 years of complete data, 2009-10, 2010-11, 2011-12, 2012-13, 2013-14, 2014-15 where each year runs from August to July.

## 2.4 Scotland Data analysis

Swine flu has been reported as outbreaks in many countries in the world in 2009, e.g. Mexico [25, 121], America in New York [83, 127], Australia and New Zealand [7, 19] and in many other countries in the world. Scotland is one of those countries which suffered from influenza in 2009 [1]. So we expect high rates when we examine either ARI or ILI consultations in Scotland. It is worth studying because of this.

We will start by presenting ARI GP consultation rates in the whole of Scotland from the mid-year of 2009 to the mid-year of 2015. Figure 2.2 shows the ARI rates during 2009 to 2015, and clearly the rate exceeds the normal range for ARI in the winter seasons in general. As there was an outbreak in 2009, we see the rates for ARI in winter of 2009 reported as more than 800 consultations per 100,000 population and similarly in winter of 2010. In the winter of 2011, 2012, 2013 and 2014, we see decreases in reported rates to under 600 consultations per 100,000 population.

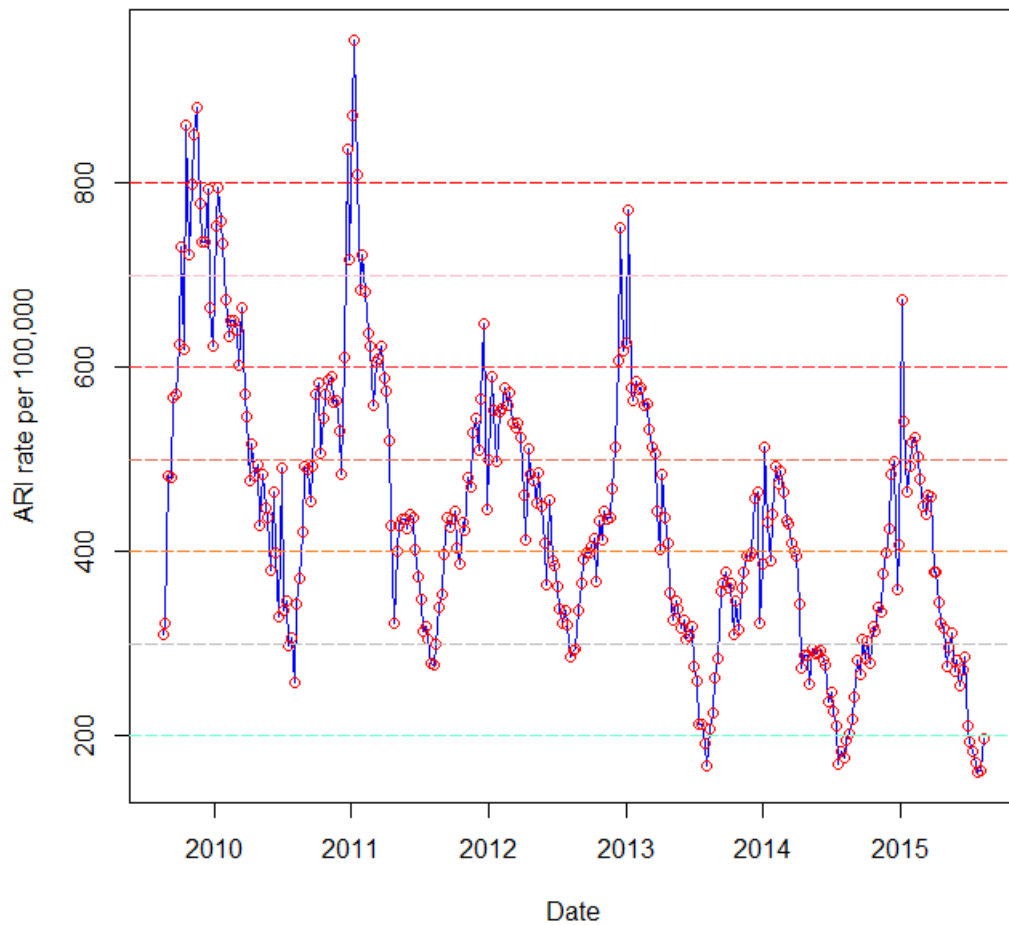


Figure 2.2: ARI weekly consultations rates in 2009-2015 for Scotland, for all 14 health boards.

Figure 2.3 presents the ILI GP consultation rates for the whole of Scotland from mid 2009 to mid 2015. We can see the rates in the winter of 2009 exceed the normal threshold (50 consultations per 100,000 population [56,64,142]). The ILI rates exceed 200 consultations per 100,000 population in winter 2009 and 150 consultations in winter of 2010 and decrease to fewer than 25 consultations per 100,000 population reported in the winter of 2014. In the winters of 2012 and 2014 the reported rates were lower than those reported in the winter of 2011 and 2013 but still under the normal fixed threshold.



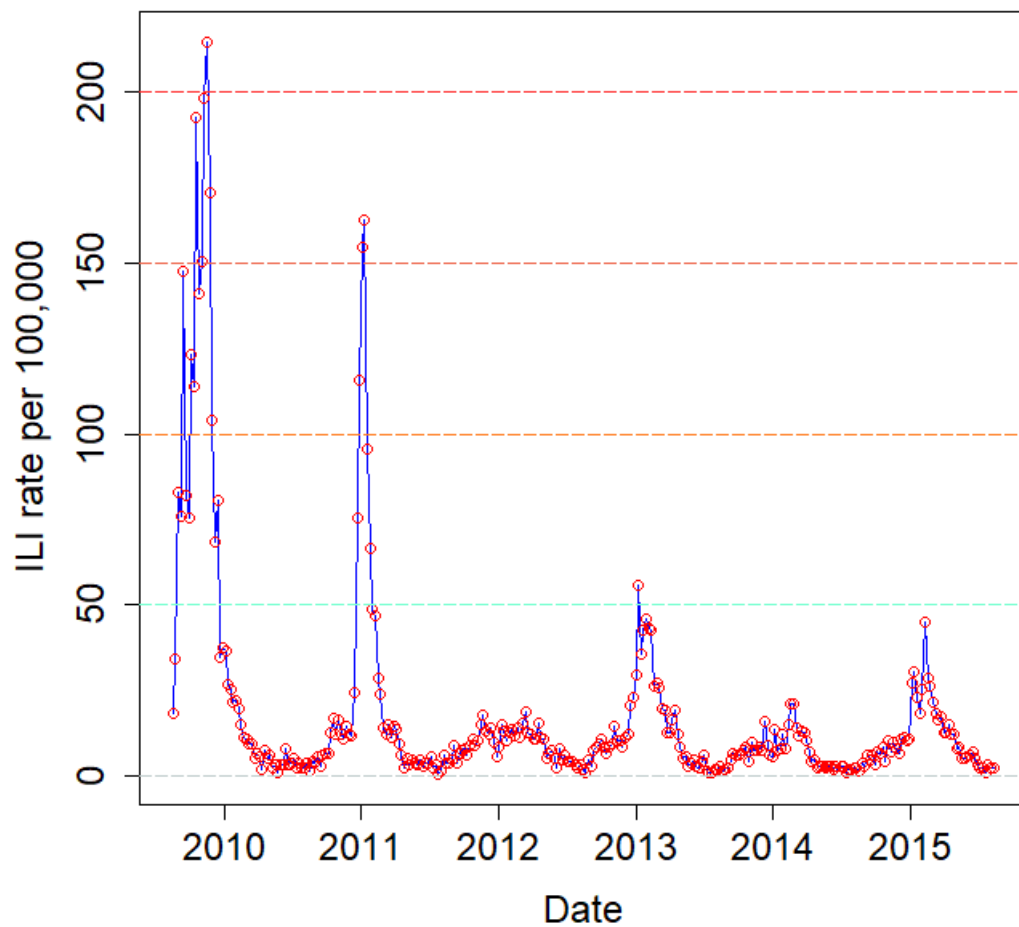


Figure 2.3: ILI weekly consultations rates in 2009-2015 for Scotland, for all 14 health boards.

When we see figure 2.3, there are lots of zeros in the ILI consultations rates in 2009-2015 for Scotland, so we will find many zero rates when we present the individual rates by dividing Scotland into the 14 health boards. So we choose now to focus on the ARI data for investigating the data. Next we will present the individual ARI rates by dividing Scotland into the 14 health boards.

As the ILI data is more specific to influenza, we used ILI data later when developing the detection system.

## 2.5 GAM model

Statistical tools can now help with understanding the trend in real data and also can be useful to generate other simulated observations. Generalized Additive Models (GAMs) have become very important methods for modelling data in recent years. We chose a GAM model in this data analysis because it is more flexible than linear regression, as GAMs can model non-linear trends. Consultation rates tend to be seasonal but a parametric model with sine and cosine terms does not fully capture the seasonal pattern, and so using Generalized Linear Models (GLMs) with sine and cosine terms to model the seasonal component is not the best solution. Tools for GAMs (and generalized ridge regression) are provided by the R `mgcv` package [178]. Penalized regression splines are used to represent smooth terms in the package `mgcv`, which can be used to fit a generalized additive model to data. The aim of this package is to provide the convenience of GAM modelling in R, and also model selection methodology was much improved from an earlier version [57,134,179]: the smooth functions were rewritten by using an acceptably chosen set of basic functions. Each of those functions has a joint penalty which enables its (effective) degrees of freedom to be controlled through a single smoothing parameter [179]. There is a large family of splines which can be used with GAMs in `mgcv`, such as cubic splines, B-Splines, natural splines, thin-plate splines and smoothing splines. `S(Time)` is one of the smoothing splines which is used in GAMs in this analysis.

We consider the smoother given by the function  $f(X_i)$  in:

$$Y = \alpha + f(X_i) + \varepsilon_i \quad (2.5.1)$$

where  $\varepsilon_i \sim N(0, \sigma^2)$ . Now, an expression for function  $f(X_i)$  can be written like a linear regression model, which means that  $f(X_i)$  is built up in basic constituents, called the basis functions  $b_j(X_i)$  as:  $f(X_i) = \sum_{j=1}^p \beta_j \times b_j(X_i)$ .

Consider  $p = 4$ , giving:  $f(X_i) = \beta_1 b_1(X_i) + \beta_2 b_2(X_i) + \beta_3 b_3(X_i) + \beta_4 b_4(X_i)$ .

Now assume that  $b_1(X_i) = 1, b_2(X_i) = X_i, b_3(X_i) = X_i^2$  and  $b_4(X_i) = X_i^3$ .

So,

$$f(X_i) = \beta_1 + \beta_2 X_i + \beta_3 X_i^2 + \beta_4 X_i^3. \quad (2.5.2)$$

This is a cubic polynomial and can produce a wide range of shapes, depending on the values of  $\beta_1, \beta_2, \beta_3$  and  $\beta_4$  and the value of function  $f$  using the expression in equation 2.5.1, but the problem is that in practice we do not know the values of  $\beta_1, \beta_2, \beta_3$  and  $\beta_4$ .

Also the shapes of the function  $f$  are not flexible enough to model more complicated patterns. We do know the value of  $X_i$  as it is a measured explanatory variable. The usual approach to solve those two problems is that the explanatory variables  $X_i$  are already known, so consider dividing the dependent variable  $Y = f(X_i)$  into four segments and for each segment fit the model in the last equation 2.5.2 using ordinary least squares (OLS), so we can find  $\beta_1, \beta_2, \beta_3$  and  $\beta_4$  per segment, and by using multiple segments we can fit more complicated patterns [194].

Many smoothing techniques can be used to do this GAM curve fitting. We used the R package `mgcv` [194]. A GAM in `mgcv` uses splines identified by `s(Time)` where time here is  $X$ , and we used `s(Time)` as the ARI or ILI cases are not linear with changing time. The advantage of using GAMs in `mgcv` is that we can do cross-validation to fit the model, especially for health boards which have a large geographical area with low population, and we could generalise the models to include terms for age group, gender and spatial location (i.e. health board), for example. We choose explained deviance as a measure of the fit of the predictive model.

### 2.5.1 ARI Data for each (HB) with GAM models

Here we are presenting the real ARI cases in each year for each health board (HB) individually and also we fitted GAMs to see the general pattern for each HB, after smoothing of the data. Figure 2.4 shows the first year analysis from 01/08/2009 to 31/07/2010 for ARI consultations. We see that the majority of health boards have almost the same pattern in the GAM model. The rates start from under 100 consultations per 100,000 population and increase in December and January to rates around 150-200 per 100,000 population. There is clearly a different pattern in health boards Tayside (TY) and Lanarkshire (LN), which start from rates of over 150 consultations per 100,000 population before September. Some of the models in this 12 month period did not fit the data very well because the deviance explained ranges from 7% in Orkney (OR) to 55-57% in Lothian (LO) and Highland (HG). In most of the health boards, the best fitting GAM is based upon about 9 degrees of freedom and the model captures the general seasonal pattern throughout the year. This model does not capture the daily variation associated with weekday. See table 2.2 for a summary of model fit for each HB and year.

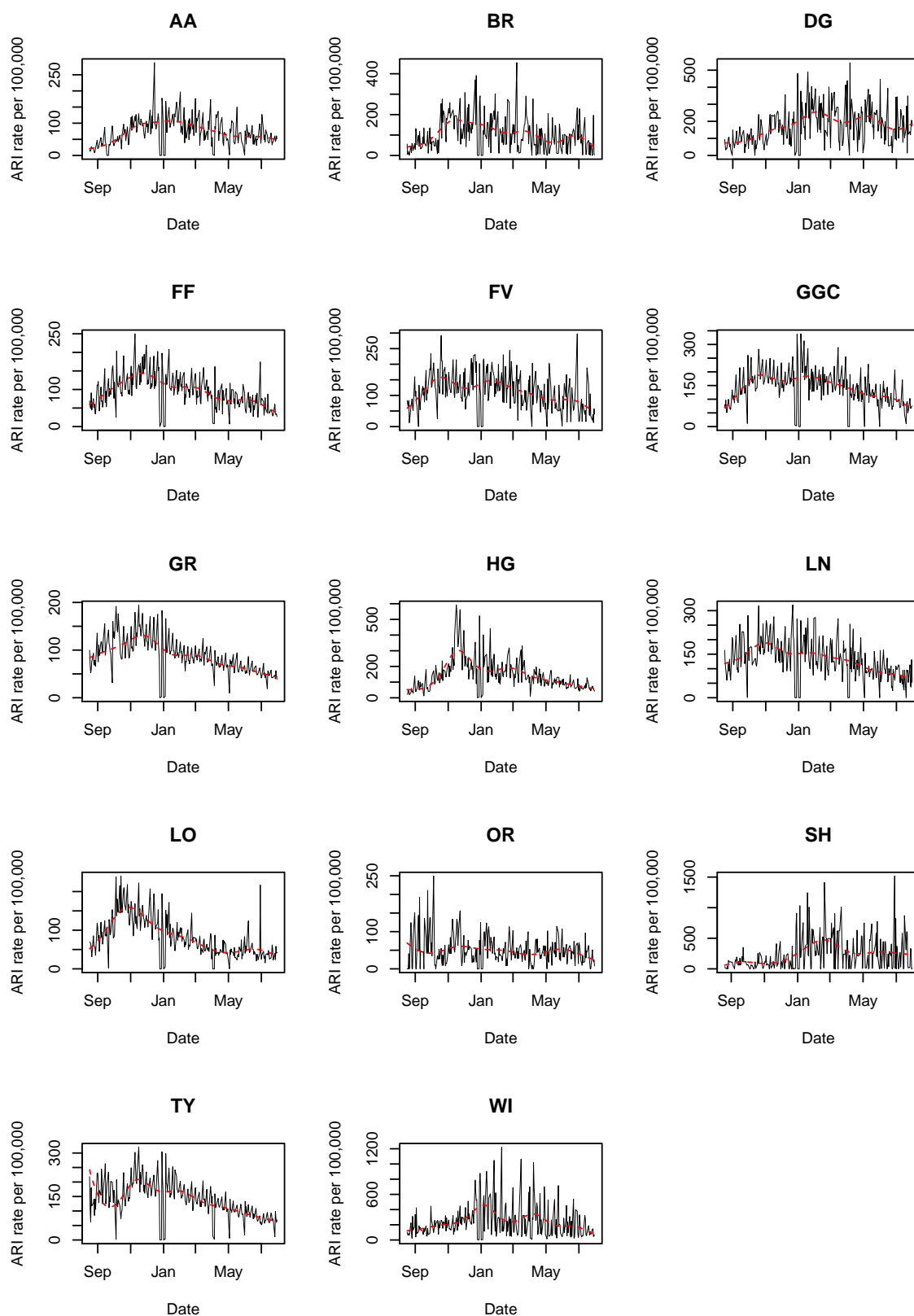


Figure 2.4: ARI consultation rates in 2009-2010 for Scotland with fitted GAM models shown in red.

Figure 2.5 shows the second year analysis from 01/08/2010 to 31/07/2011. We see that the majority of health boards have almost the same pattern in the GAM model. The rates are not much different from the previous year and start from under 100 consultations per 100,000 population and increase in December and January to rates around 200-300 per 100,000 population, except for Shetland (SH) which shows under 100 consultations per 100,000 population, and Borders (BR), which is clearly different and increases from 100-400 consultations per 100,000 population. There is clearly a different degree of variation in health boards Western Isles (WI) and Shetland (SH), which have very marked daily variation. Some of the GAM models in these 12 months did not fit the data very well because the deviance explained ranges from 4% in Shetland (SH) to 44% and 46% in Grampian (GR) and Lanarkshire (LN) respectively. See table 2.2 for a summary of model fit for each HB and year.

Figure 2.6 shows the third year analysis from 01/08/2011 to 31/07/2012, and we see that the majority of health boards have almost the same pattern in the GAM model. The rates are not much different from the previous years and start from under 100 consultations per 100,000 population and increase in December and January to rates around 150-250 per 100,000 population, except for Highland (HG), which shows under 150 consultations per 100,000. The peak winter rate is lower in this year compared to previous years. Highland (HG) shows a very unusual increase on 18/04/2012 in the consultations per 100,000 population. There is clearly a difference in variation for health boards Orkney (OR) and Western Isles (WI) which have very acute daily variation during this year. Some of the GAM models in these 12 months did not fit the data very well because the deviance explained ranges from 15% in Lanarkshire (LN) to 33% in Grampian (GR). See table 2.2 for a summary of model fit for each HB and year.

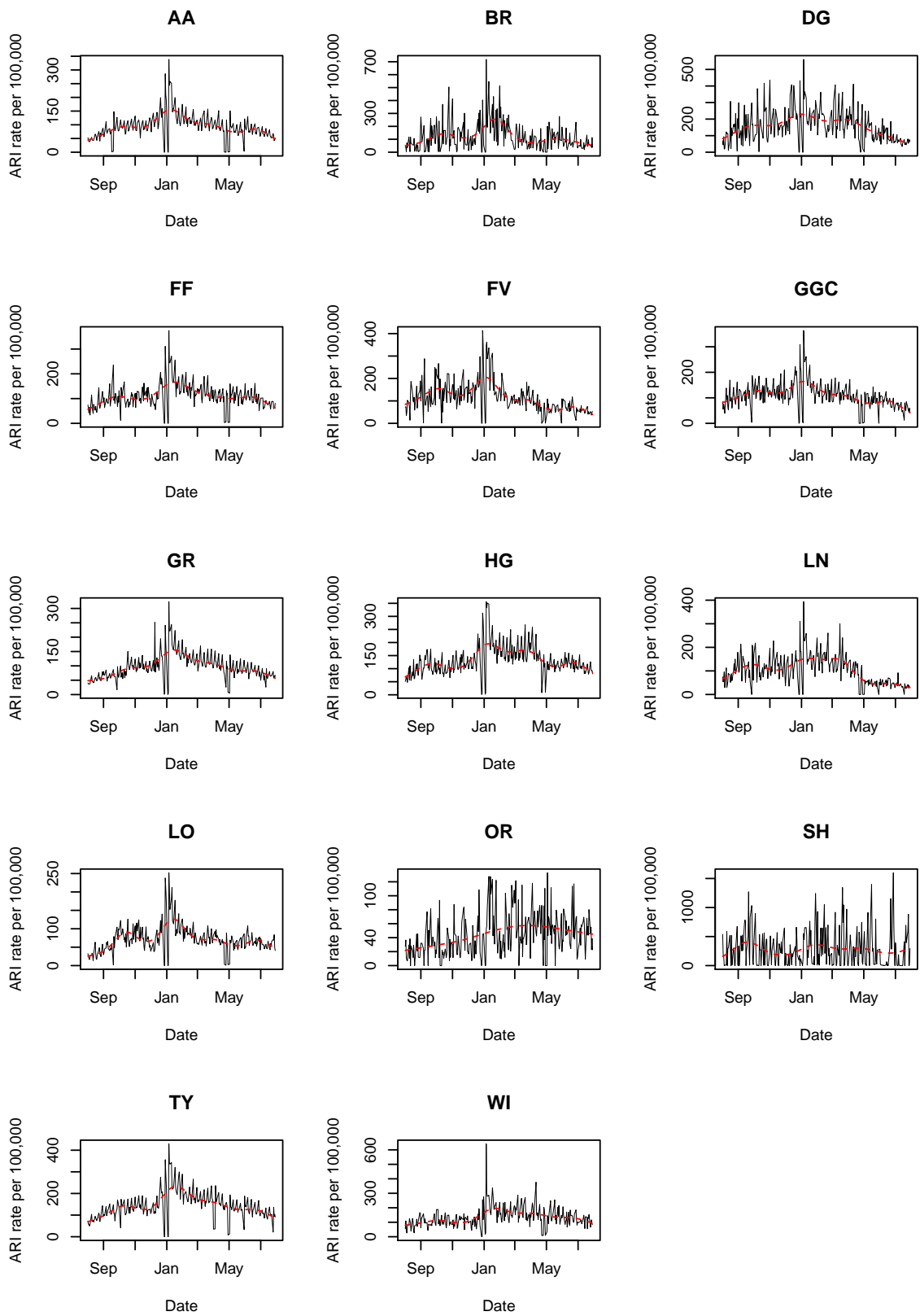


Figure 2.5: ARI consultation rates in 2010-2011 for Scotland with fitted GAM models shown in red.

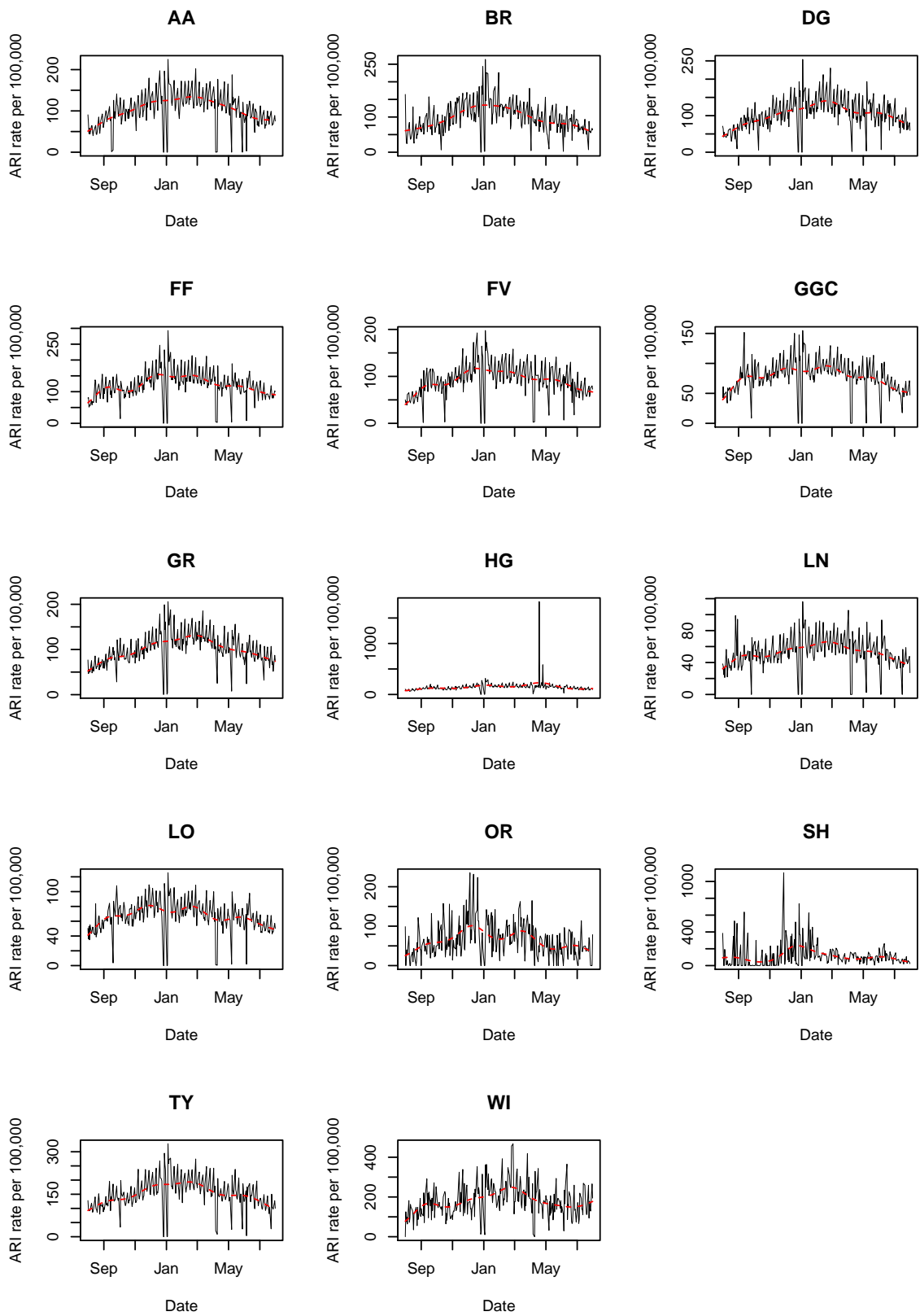


Figure 2.6: ARI consultation rates in 2011-2012 for Scotland with fitted GAM models shown in red.

Figure 2.7 shows the fourth year analysis from 01/08/2012 to 31/07/2013, and we see that the majority of health boards have almost the same pattern of the GAM model. The rates are not much different from the previous years, which start from under 100 consultations per 100,000 population and increase in December and January to rates around 200-250 per 100,000 populations. Highland (HG) again in this year shows a very unusual increase on 25/06/2013 by to 1078.87 consultations per 100,000 populations. There is clearly a different variation in health boards Orkney (OR) and Shetland (SH) which have very acute daily variation during this year. Some of the GAM models in these 12 months did not fit the data very well because the deviance explained ranges from 23-24% in Shetland (SH) and Western Isles (WI) to 49% in Grampian (GR). See table 2.2 for a summary of model fit for each HB and year.

Figure 2.8 shows the fifth year analysis from 01/08/2013 to 31/07/2014, and we see that the majority of health boards have almost the same pattern of the GAM model. The rates are low when compared with the previous years, starting from under 80 consultations per 100,000 population and increasing in December and January to rates around 100-150 per 100,000 population. There is clearly a different variation in health board Orkney (OR), which has very acute daily variation during this year. Some of the GAM models in these 12 months did not fit the data very well because the deviance explained ranges from 25% in Orkney (OR) to 42-43% in Tayside (TY), Dumfries and Galloway (DG) and Grampian (GR). In most of the health boards, the best fitting GAM model is based upon about 9 degrees of freedom and the model captures the general seasonal pattern throughout the year. This model does not capture the daily variation associated with weekday. See table 2.2 for a summary of model fit for each HB and year.



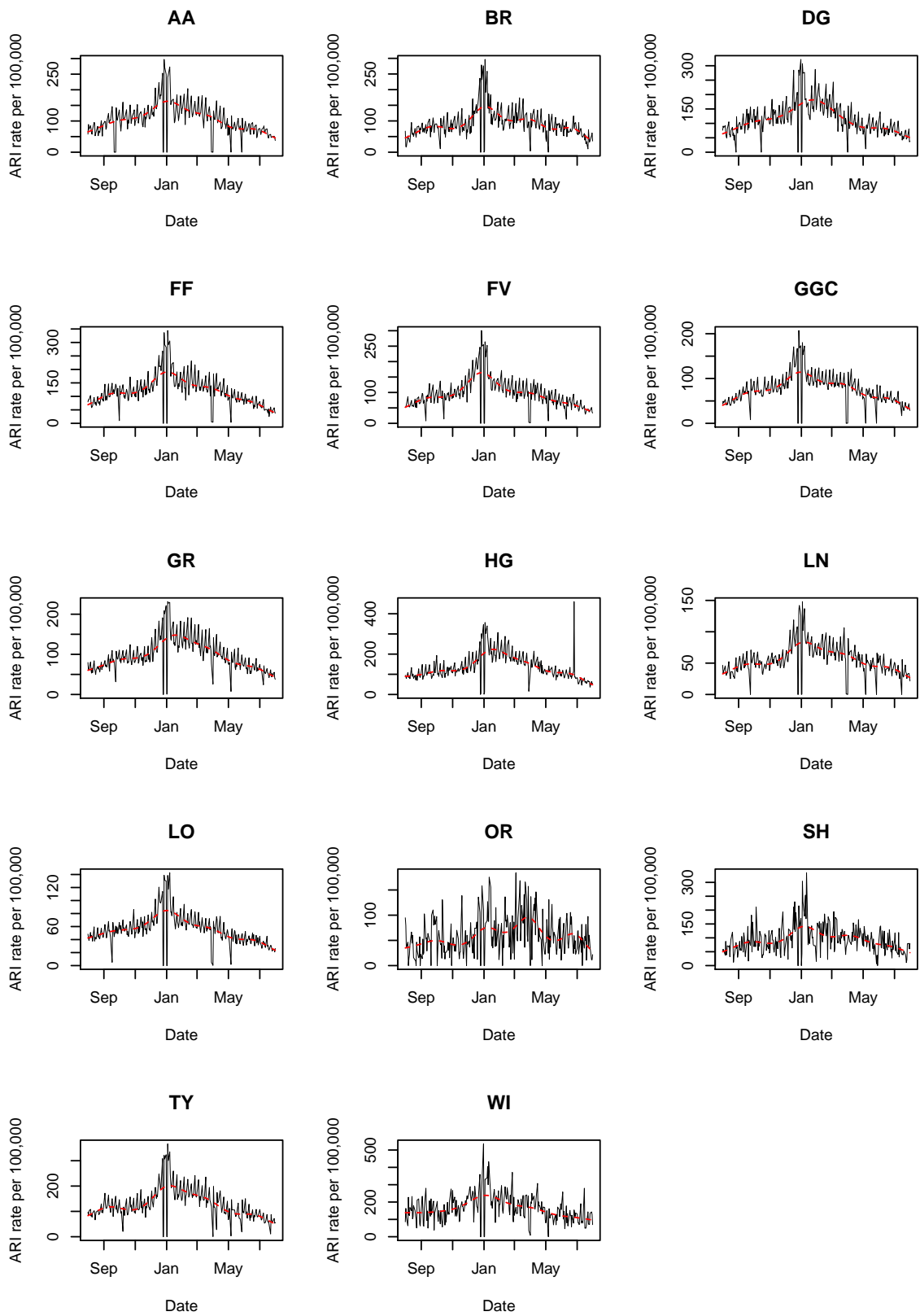


Figure 2.7: ARI consultation rates in 2012-2013 for Scotland with fitted GAM models shown in red.

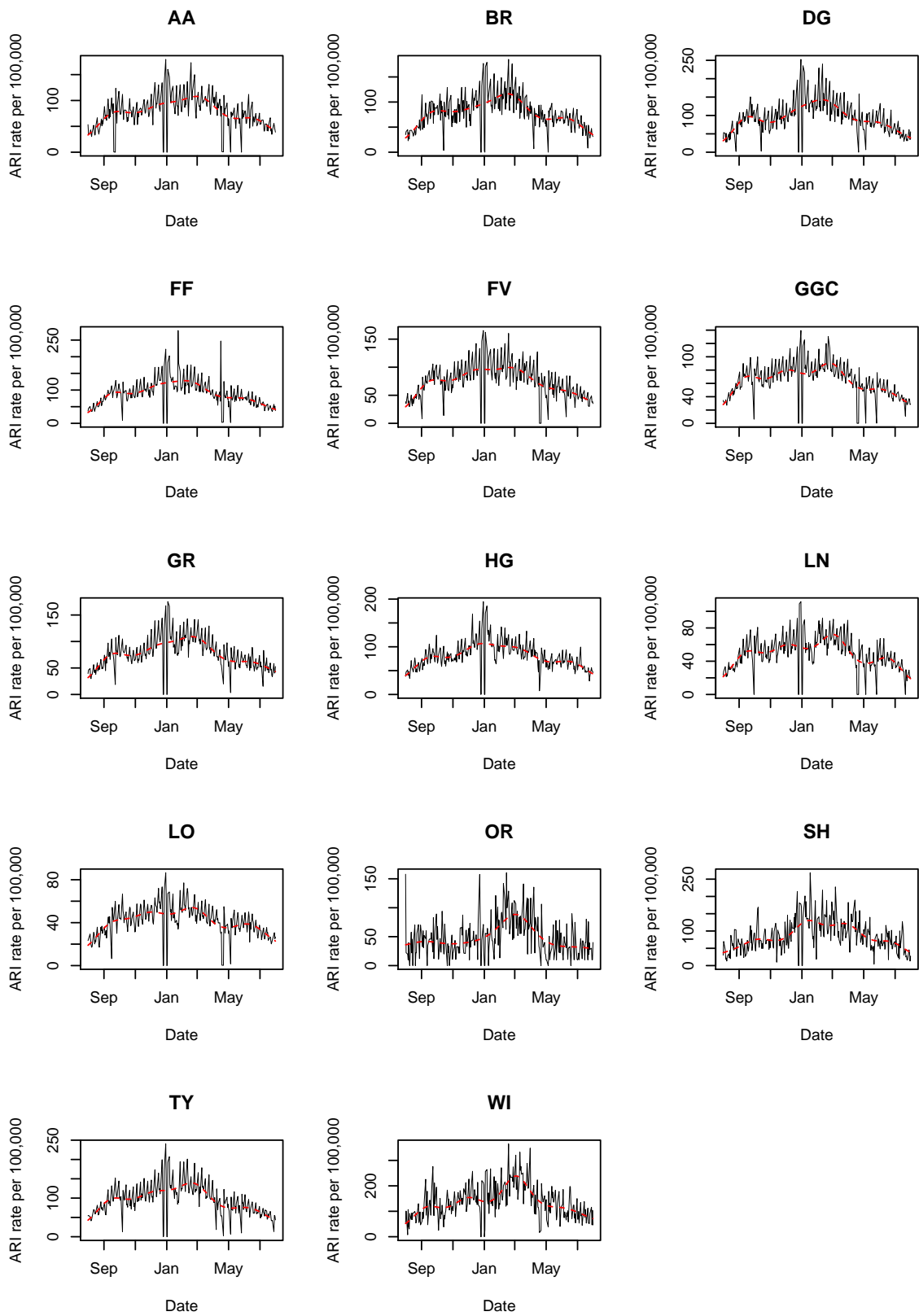


Figure 2.8: ARI consultation rates in 2013-2014 for Scotland with fitted GAM models shown in red.

The period from 01/08/2014 to 31/07/2015 is the last period we want to analyse in this section. It seems to be very similar to the period of 2013/14 as presented in figure 2.9 except for some decreasing rates in the health board of Greater Glasgow and Clyde (GGC) and some increasing rates in the health boards of Grampian (GR) and Lothian (LO). In most of the health boards, the best fitting GAM model is based upon about 9 degrees of freedom and the model captures the general seasonal pattern throughout the year. This model does not capture the daily variation associated with weekday. The majority of the GAM models in these 12 months did not fit the data very well because the deviance explained ranges from 24% in Western Isles (WI) to 54% in Highland (HG). See table 2.2 for a summary of model fit for each HB and year.

In general the models explain the data less well in the smaller HBs where the fluctuations are more erratic (table 2.2).

HB	2009-2010	2010-2011	2011-2012	2012-2013	2013-2014	2014-2015
AA	36.60%	35.70%	27.50%	38.40%	32.70%	39.40%
BR	23.30%	25.50%	29.50%	35.20%	41.40%	46.20%
DG	23.80%	27.90%	30.30%	42.90%	42.80%	40%
FF	37.60%	26.20%	23.80%	46%	41.80%	48.20%
FV	23.10%	39.20%	25.10%	46.20%	38%	41.70%
GGC	31.20%	30.20%	22.50%	40.80%	41.40%	47.10%
GR	49.80%	44%	32.70%	48.90%	42.60%	51.70%
HG	55%	35.90%	23.40%	50.60%	37.30%	54.30%
LN	26.80%	45.50%	15%	34.90%	36.30%	50.80%
LO	56.70%	39.50%	17.10%	41.10%	32.50%	39.30%
OR	7.05%	15.90%	20.40%	16.80%	24.50%	29.10%
SH	22.70%	3.95%	18.30%	22.80%	34.60%	27.90%
TY	46.30%	40.10%	25.40%	44.50%	42.20%	50.80%
WI	20.80%	24.80%	21.50%	23.90%	39.50%	24.10%

Table 2.2: Deviance explained for each GAM of ARI rates in Scotland, for each health board and year.

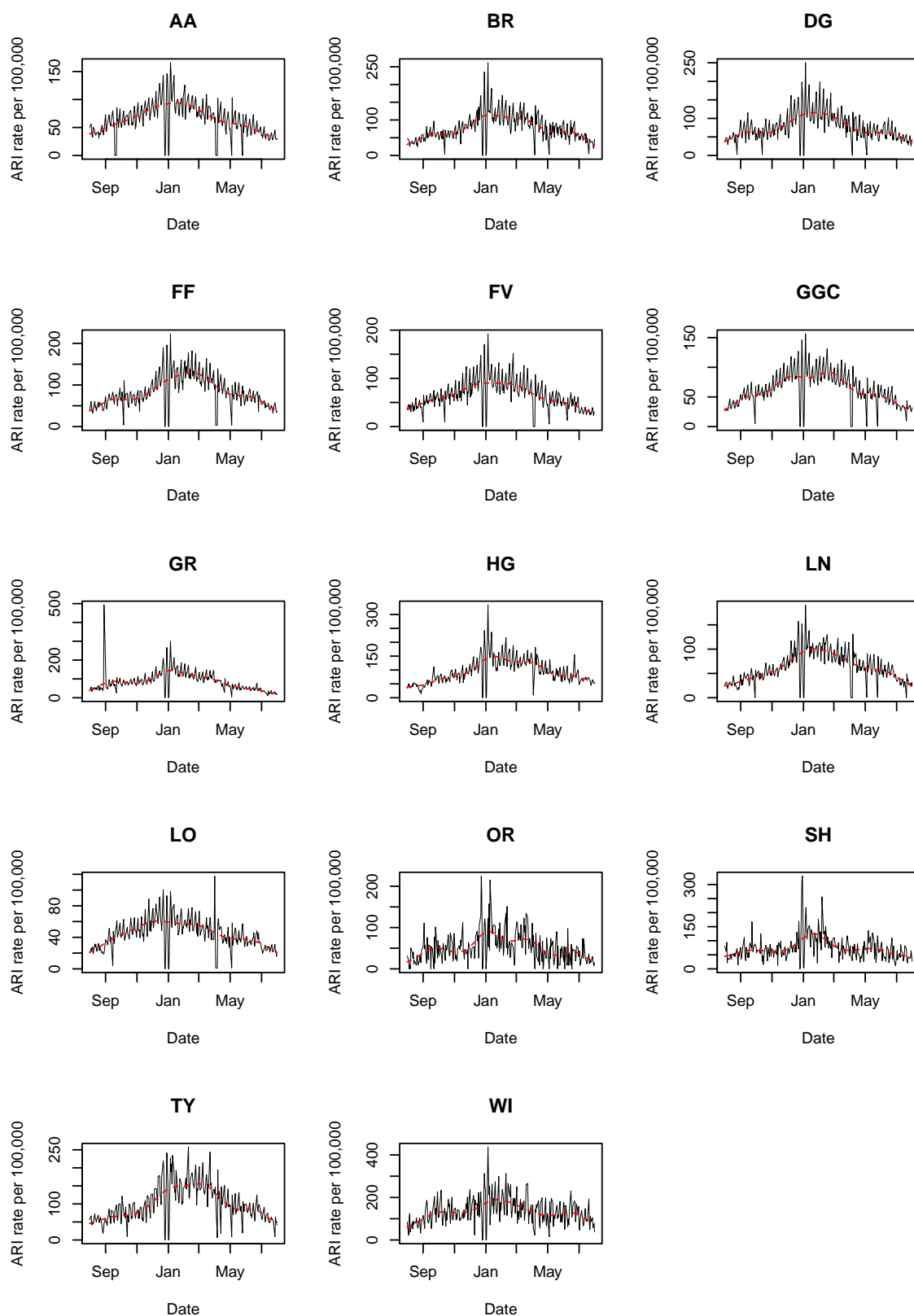


Figure 2.9: ARI consultation rates in 2014-2015 for Scotland with fitted GAM models shown in red.

### 2.5.2 GAM models only for health boards

In the next six graphs, we consider the GAMs' seasonal trends for the different health boards in each year separately without the original data. It will be easier to compare and check the patterns with different health boards. Smaller HBs make the picture confusing as they show unusual trends, which are very variable and often do not fit in with the general pattern, so we will drop them (SH, OR, WI and DG) from the plots. We also plot all HBs on the same plot for comparison.

Figures 2.10 to 2.15 show the fitted trend from the GAM models for each health board for the periods 2009-10, 2010-11, 2011-12, 2013-13, 2013-14 and 2014-15 respectively. The general pattern in each is that there is a clear peak around the period December to February (winter). In 01/08/2009 to 31/07/2010 (Figure 2.10) the peak for the majority of the boards is earlier in November and the range of rates over the larger boards is 80-200 consultations per 100,000 population in November. Of the larger regions, HG has a slightly different pattern, with a much higher rate in December. In March, the health boards show gradual decreasing to achieve finally around 90 consultations per 100,000 populations in July.

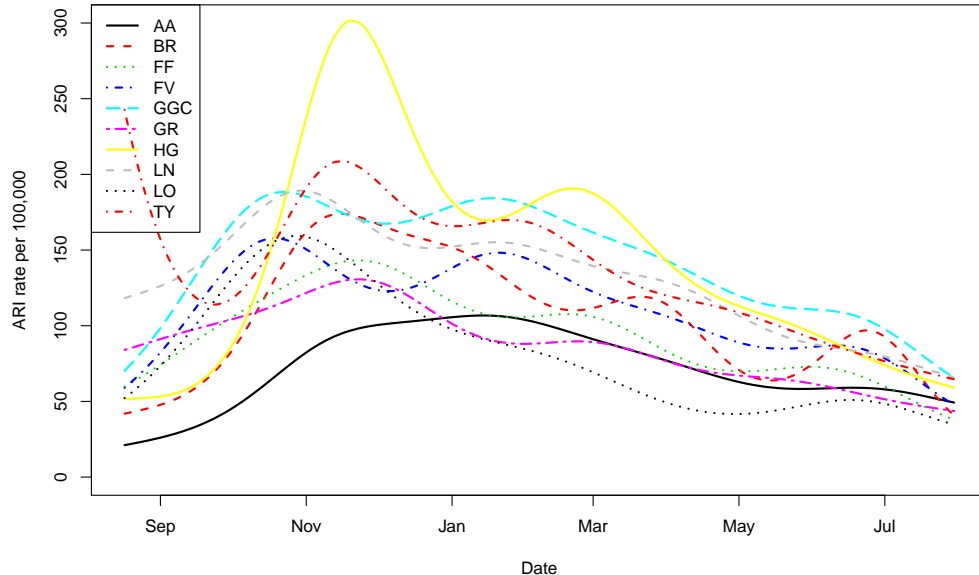


Figure 2.10: GAM models for 2009-2010

In figure 2.11, which covers the period of 01/08/2010 to 31/07/2011, there is a suggestion of 2 peaks, one in October and the main one in January. We see clearly that the different health boards always have the same patterns, starting from around 50-90 consultations in September then decreasing before going again to the annual increase in January to achieve

around 150 consultations. In March, the health boards show gradual decreasing to achieve around 100 consultations in July. Borders (BR) shows the highest variation during this year.

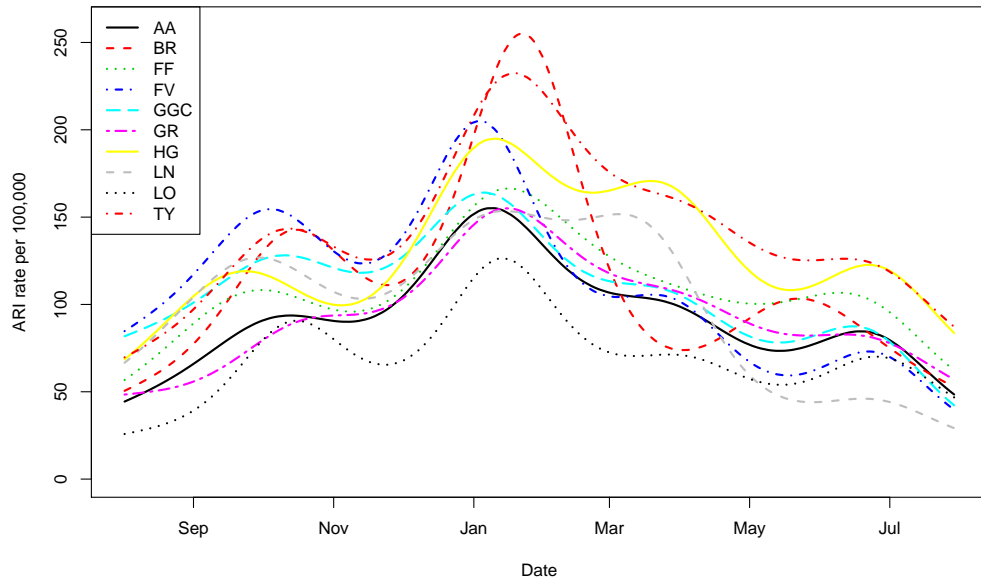


Figure 2.11: GAM models for 2010-2011

GAM models for different health boards from 01/08/2011 to 31/07/2012 are presented in Figure 2.12. We clearly see the majority of health boards (apart from the long yellow line – HG) always have the same pattern, keeping in a limited range starting with a median of 70 consultations until the beginning of November, then they increase to achieve around 100 consultations, until the beginning of May before they go back to 70 consultations in July. The health board of HG shows a similar pattern to the majority, but it leaves them from March to reach a peak at 240 consultations in May, then it goes back to the majority trend at the beginning of July. Also GR is high.

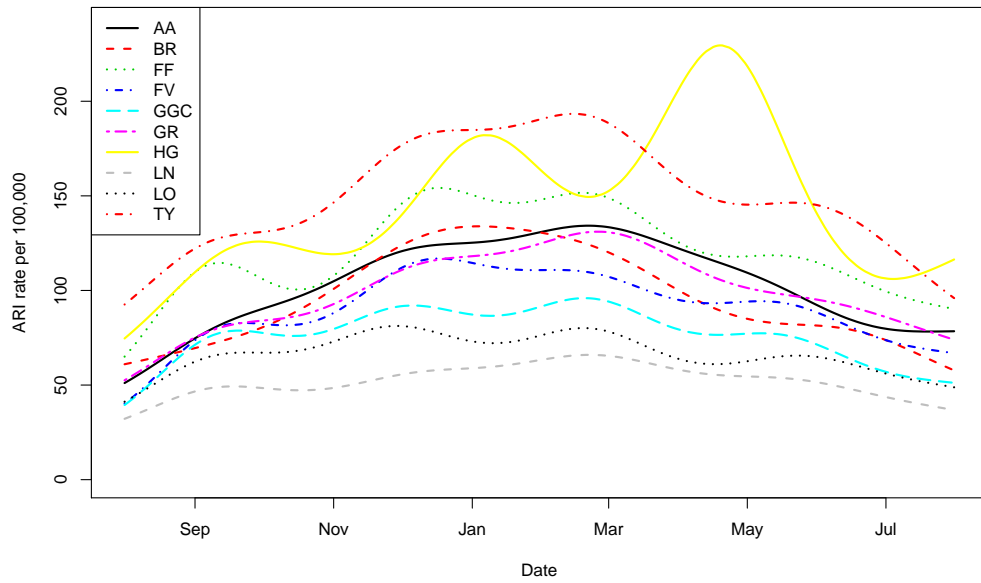


Figure 2.12: GAM models for 2011-2012

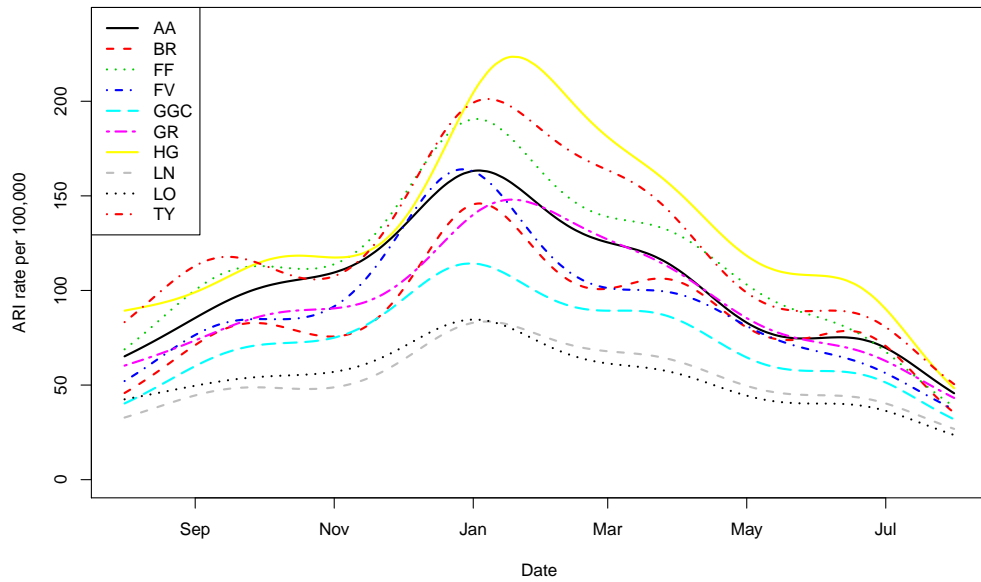


Figure 2.13: GAM models for 2012-2013

In 01/08/2012 to 31/07/2013 (Figure 2.13) there is a suggestion of 2 peaks, one in September and the main one in January. We see clearly that the different health boards always have the same patterns, starting from around 40-120 consultations in September, then they show very little decrease before they go back again to the annual increase in January

to achieve the median of 150 consultations. In March, the health boards show a gradual decrease to achieve around 70 consultations in July. HG again showed a peak, in February 2013.

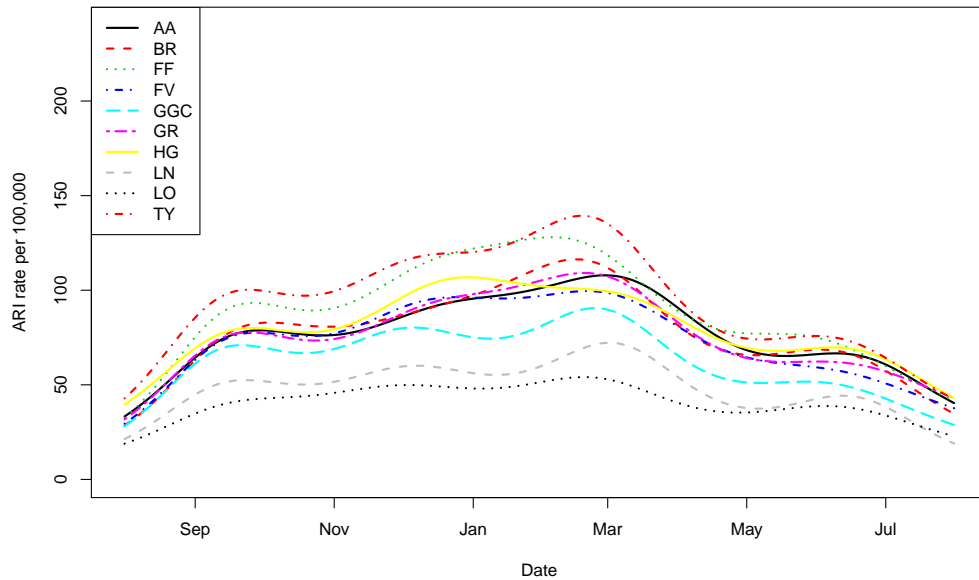


Figure 2.14: GAM models for 2013-2014

This year has shown a marked difference in trend compared to previous years. A similar pattern appears for all health boards. The greatest rise in this year was in March, which did not happen in the previous years. Figure 2.14 shows GAM models for different health boards from 01/08/2013 to 31/07/2014. We clearly see that the majority of health boards always have the same pattern. All health boards are keeping in a limited range, starting with a median of 60 consultations until the beginning of November, then they show a sharp jump to achieve around 100 consultations, until the arrival of the top level this year at the beginning of March, before decreasing to 50 consultations in July. Health boards show homogeneity, with very little disagreement in their patterns. Figure 2.15 shows GAM models for different health boards from 01/08/2014 to 31/07/2015, which show very similar patterns to the last year.



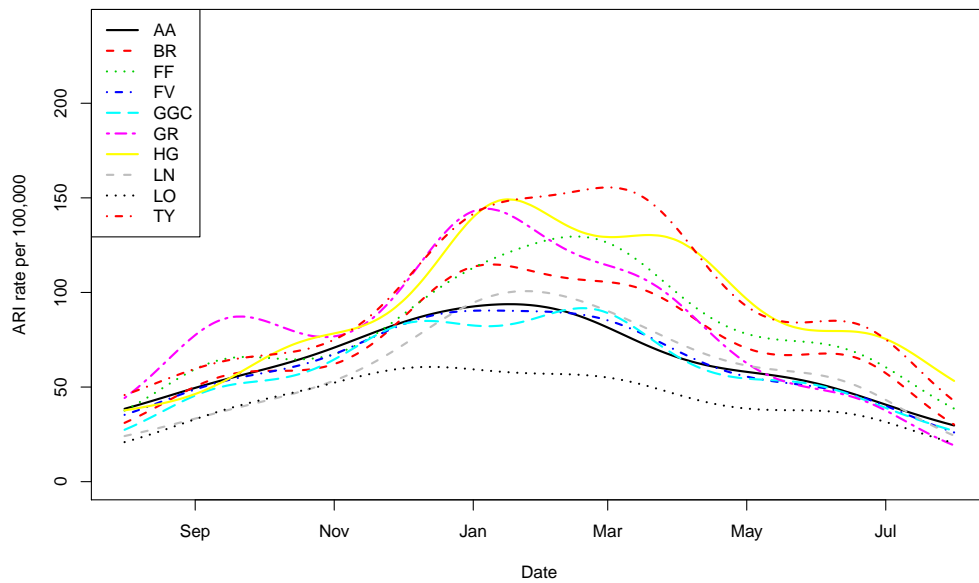


Figure 2.15: GAM models for 2014-2015

## 2.6 Comments

Given the graphs of Acute Respiratory Illness (ARI) consultations for different years and health boards, we note that there are differences between years for each health board, but it should be noted that mostly the trends (curves) are similar each year. We see a decrease in Acute Respiratory Illness (ARI) consultation rates in the summer in general, and a corresponding rise in the rates in the period of winter and low temperature can be seen clearly in the months December, January, February, March, then the level begins to decline and return to the normal level. A few health boards have clear differences in trend from the other health boards in some years. One published study in November 2014 (which dealt with the analysis of infectious diseases in England and Wales) has also indicated that there is a remarkable increase in numbers of hospital records during the winter period and low temperatures. It also showed differences in the records for influenza for different health sectors in England and Wales [75].

This chapter has investigated data trends in ARI consultation rates for all the health boards in Scotland over the period 2009-10 to 2014-15. It is important to be able to predict when the winter rise in the consultation rates will begin, in order to take any necessary action (as discussed in chapter 1). Chapter 3 now outlines the main methods used for routine monitoring of diseases in order to detect the start of an epidemic, some of which will be used in later chapters.

## Chapter 3

# Influenza detecting methods

In this chapter, and before starting the review of the research methods, here we present an overview of some of the methods used for the control of infectious diseases, as well as some early warning systems for these diseases. There are three important criteria to compare any early warning system: sensitivity (Does the system detect an increase when it occurs?), specificity (Does the system avoid an alarm when no increase occurs?) and finally the average time to detect an increase once it happens. The main measure from those is sensitivity, as it is a more important characteristic for influenza epidemic detection than specificity, since the announcement of the start of an influenza epidemic could encourage media attention to raise the level of preparedness of the people to benefit from the prevention of influenza and allow control measures such as the use of vaccination campaigns or the implementation of non-pharmacological interventions [162]. The general idea for creating a system to monitor the number of those infected by a particular disease and the establishment of an early warning system is not a new idea but a well established one. However with the development of science came additions, amendments and proposals to increase the efficiency of these methods. Surveillance systems are different depending on geographical location and the possibility of collection of relevant data used in each area.

Here is a search and evaluation of some available literature in early detection methods of infectious diseases, especially influenza. Google scholar is used here in this literature search.

There are many detection systems to monitor infectious diseases, such as detecting influenza epidemics using search engine query data [55], MassTag Polymerase-Chain-Reaction [79], rapid antigen tests [82] and many other systems depending on the available data which can be used. The applied work developing early warning systems raises fundamental questions such as: Which data are useful for early detection? What are the timeliness requirements for outbreaks caused by different agents? How should we measure timeliness

of a detection system for a specific type of outbreak? [164].

Below some methods using influenza data are presented in more detail.

### 3.1 Threshold method

The idea of the establishment of disease surveillance control systems goes back to the late 19th century, to Dr. Selwyn Collins, who worked as a chief of the Statistics Section in the Epidemiology Branch in the Communicable Disease Center in Atlanta. He was the first to think of the creation for monitoring systems of death due to influenza. The general idea is that the system revolves around three themes: firstly, estimation of expected deaths, rather than analysing the epidemic after mortality. Secondly, replacing deviations from expectancy with a graphic representation of “expected seasonal mortality” and actual deaths, and finally, use of total number of deaths instead of rates [139]. This method is considered as a basis for all methods of early prediction of influenza that set off an alarm after a certain point. There are multiple ways to get to the warning alarm, that combine speed, accuracy and sensitivity.

A large group of cities in the United States reported cases of deaths caused by influenza in September, 1918 [32]. One of the first studies that focused on early detection of influenza is the study conducted by Collins, collecting evidence in interviews about respiratory diseases since 1917 and until 1931. Collins noted that “for every period in which there is a definite peak of excess mortality from influenza and pneumonia, there is a corresponding peak of excess mortality credited to causes other than influenza and pneumonia”. For a given week, using the median (weekly) number of deaths during non-epidemic years, he defined the number of expected deaths from influenza and pneumonia, or from all causes. The difference between the observed number of deaths and the expected number of deaths gives the number of excess deaths from pneumonia and influenza, or from all causes. In the period 1956-1960, Eickhoff et al. [45] studied the weekly pneumonia and influenza (P and I) deaths which covered 108 cities and the obtained measure of the extent of an influenza epidemic was “the number of deaths due to P and I in excess of the usual expectancy”. In [45] deaths over successive four-week periods were used to fit a regression equation, then the data for the influenza epidemic weeks were omitted. The number predicted from the regression equation was used to define the expected number of deaths for a given period, and finally, the difference between the observed and expected numbers defined the excess deaths due to an influenza epidemic.

One method of detection uses the number of expected cases in week  $w$  and defines the cut-off point so an alarm will signal if this cut-off point is exceeded for two weeks consecutively. This method is called the threshold method [56, 142], and the dynamic linear model [80,

174]. In Scotland for example the baseline for influenza activities is between 0 to 50 consultations per 100,000 population, 50 to 600 consultations per 100,000 population for normal seasonal activity, while 600 to 1000 consultations per 100,000 population is higher than expected, and more than 1,000 consultations per 100,000 population is exceptional [56, 64, 142].

One of the methods predicting using a threshold technique is prediction (nowcasting) of influenza epidemics over years, using the total population, for example, in Östergötland County, Sweden [145].

## 3.2 Time series technique

This method is simple and free from complicated steps because it does not require specialized statistical software and can be applied directly on the raw data. Let  $n$  be the number in the series of observations. The proposed model for  $n$  observations  $y_t$  is described by the equations:

$$y_t = \theta_t + v_t$$

where  $v_t \sim N(0, V)$  and

$$\theta_t = \theta_{t-1} + w_t$$

where  $w_t \sim N(0, W)$  and  $\theta_t$  denotes the chain of unobserved system parameters which describe the correlation between sequential weeks, while the errors  $v_t$  and  $w_t$  are inwardly and reciprocally independent. From the data, variance  $V$  can be estimated. The parameter  $W$  must be prespecified and represents the assumed smoothness of influenza prevalence changes from week to week [37].

### 3.2.1 ARIMA method

In this section, we describe analyzing time series data using a class of stochastic models, called an autoregressive integrated moving average (ARIMA) process. These provide different classes of models, stationary or not, which can be applied to many time series datasets in epidemiology.

Let  $(Z_t, t = 1, 2, \dots)$  be a progression of dependent random variables called a time series or stochastic process. At time  $t$ ,  $Z_t$  might correspond to the observation of an experiment. Discrete empirical time series come from collections of dependent (not necessarily independent) observations of an experiment, taken at regular time intervals. Different classes of mathematical models in an ARIMA process describe a kind of dependence through the observations in an experiential time series. Stationary stochastic processes cannot describe many experiential time series (e.g., weekly pneumonia and influenza P & I), but some suitably differenced series may be stationary [29].

Processes for forecasting by ARIMA models start from fitting an ARIMA model to the empirical time series. Box and Jenkins [22] developed a four stage iterative procedure for selecting a model and forecasting: i) exploratory specification (or identification); ii) estimate parameters ; iii) check diagnostics; and iv) forecasting. Exploratory or preliminary specification of an ARIMA model is based on examination of two terms: 1. the autocorrelation function ( $r_k, k = 1, 2, \dots$ ) and partial autocorrelation function ( $\rho_k, k = 1, 2, \dots$ ) of a given series. For the time series  $Z_t$ ,  $r_k$  is the autocorrelation coefficient of order  $k$  between  $Z_t$  and  $Z_{t+k}$ , while the partial correlation coefficient of  $Z_t$  on  $Z_{t+k}$  after removing the effects of  $Z_{t+1}, Z_{t+2}, \dots, Z_{t+k-1}$  is the partial autocorrelation coefficient of order  $k$ ,  $\rho_k$ , where  $r_k$  and  $\rho_k$  are independent of  $t$  because of the stationarity. The nonlinear least squares method is able to estimate the parameters in the model when an ARIMA model is tentatively specified. Now, after the parameters are estimated, the “Q-statistic” (analogous to the Pearson’s goodness-of-fit Chi-square statistic) and inspection of the estimated noise series  $\alpha_t; t = 1, 2, \dots$ , can measure the goodness-of-fit of the model. If the model fits the data adequately, the model can be used to forecast the future observations of the series. If it does not fit, going back to the preliminary identification step, then doing the following stages can make forecasting more accurate.

Choi and Thacker [29] presented this method of detection based on the time series analysis of historical data and using the seasonal autoregressive integrated moving average (ARIMA) techniques in 1981. Their method showed something new, taking into account the time-dependence of the data. Held and others [68] used a stochastic model approach through analysing simulated data and real notification data taken from the German infectious disease surveillance system, administered by the Robert Koch Institute in Berlin.

Some recent publications have been using ARIMA methods to detect influenza outbreak in different health institutions in different countries such as Egypt [76], the United States [50, 116], China [150, 165, 166], Cameroon [93] and Abidjan [107].

### 3.3 Moving Epidemic Method (MEM)

One of the ways recently used to monitor influenza and to launch early warning to minimize human losses from disease is the Moving Epidemic Method (MEM) [161, 162]. The aim from this method is to model influenza epidemics by using historical data from different locations. The modus operandi can be summarized this way in three steps, which are: firstly, division of the time from the beginning to the end of the epidemic into three periods, as follows: the period before the epidemic (pre-epidemic), the epidemic period, and the period after the epidemic (post-epidemic); secondly, calculating a baseline level of activity and an epidemic threshold using data from the pre-epidemic and post-epidemic periods from historical data; and finally, computing thresholds for different levels of intensity

within the epidemic period.

MEM searches for the optimum epidemic point by starting with one week and seeing how many cases represent the total epidemic. The one-week period is chosen by maximizing the number of cases that a period can contain, that is, getting the peak. The one-week period which contains the most cases is the period that contains the peak. If the one-week period has, say, 50 cases and the total number of cases in that season is 500, then the one-week period represents 10% of the total (50/500). Then we add one more week to form the two-week period and repeat the process: to find the two-weeks period that contains the highest number of cases (probably the peak and a neighbour). If that period has, say 100 cases, then it represents 20% of cases. So adding one week (from one to two) we have gained 10% (from 10% to 20%).

For the first step, estimate the length of the epidemic period as the minimum number of sequential weeks with the maximum accumulated rates percentage (MAP), for each season individually. The maximum accumulated rate for a period of length  $r$ , expressed as a proportion  $p_j^r$  of the total rate for the  $j$ th season, gives the curve of the MAP. In each season  $j = 1, \dots, N$ :

$$t_j^r = \max_{k=1, \dots, S-r+1} \left\{ \sum_{i=k}^{k+r-1} t_{i,j} \right\}, \forall r = 1, \dots, S \quad \text{and} \quad p_j^r = t_j^r / t_j^S, \quad \text{and} \quad t_j^S = \sum_{i=1}^S t_{i,j}$$

where  $t_{i,j}$  is the  $i$ th rate of the  $j$ th season,  $N$  is the number of seasons, and  $S$  is the number of surveillance weeks per season. Now, using a smoothing regression of  $p_j^r$  over  $r$ , choosing an auto-parameter for the smoothing window, will give  $\check{p}_j^r$  as the smoothed MAP [21, 74].

Calculating the increase in the percentage  $p_j^r$  from one period of  $r$  weeks to the next one will give the optimum number of weeks for each season:

$$\Delta_j^r = \check{p}_j^{r+1} - \check{p}_j^r.$$

The optimum is found when this increase is lower than the predefined criterion:

$$r_{j*} = \min_{r=1, \dots, S-1} \{r : \Delta_j^r < \delta\}$$

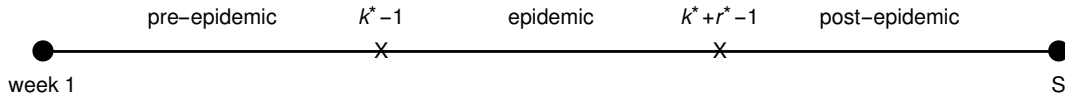
where  $r_{j*}$  denotes the optimum duration of the  $j$ th season and  $\delta$  is the minimum percentage we gain from adding one week to the epidemic period. Then that gain is smaller than  $\delta$ , we stop adding weeks and that is the optimum because there is no point in adding new weeks that add only a small percentage of the cases to what we already have. The higher the  $\delta$  parameter, the fewer weeks we are going to include in the epidemic (i.e. a shorter epidemic). The shorter the epidemic, in practice probably the higher the post-epidemic threshold. But in theory, they are not related. The parameter  $\delta$  is used in the determination of the epidemic, not the calculation of thresholds. Also, the post-epidemic

threshold does not have a significance itself but is a comparison of mean pre-epidemic and post-epidemic values, but never a threshold for ending an epidemic. The way it is calculated invalidates it as a threshold for the end of the epidemic. In fact, we never use it.

Once the optimum epidemic duration is found, the value  $k^*$  that maximises the expression used to calculate the  $p_j^r$  will determine the timing:

$$k^* : \sum_{i=k^*}^{k^*+r_j^*-1} t_{i,j} = \max_{k=1, \dots, S-r_j^*+1} \left\{ \sum_{i=k}^{k+r_j^*-1} t_{i,j} \right\}$$

where  $k^*$  gives the week when the epidemic begins, and  $k^* + r_j^* - 1$  gives the end week of the epidemic,  $(1$  to  $k^*-1)$  denote the weeks before the epidemic period (known as the pre-epidemic), and  $(k^* + r_j^*$  to  $S)$  give the weeks after the epidemic period (known as the post-epidemic):



Calculating the baseline and threshold using pre-epidemic values of historical seasons is the second step. The mean of all pre-epidemic rates of all historical seasons gives the baseline.

To calculate the threshold, we should consider pre-epidemic values only. Choosing the highest  $n$  values from the pre-epidemic period in each season where  $n = 35/(\text{number of seasons})$ ,

$$\forall j \rightarrow \left( t_{(1),j}, t_{(2),j}, \dots, t_{(n),j} \right)$$

where  $t_{(i),j}$  is the  $i$ th pre-epidemic highest rate.

Vega et al. (2013) [161,162] used 30 weeks not 35 as they assumed a winter period of 30 weeks. Here we are using a longer winter period.

A one-sided 95% confidence interval for the pre-epidemic threshold can be calculated from all these  $n \times N$  values by:

$$(0, \bar{t} + z_{0.05} \cdot S_t)$$

where  $z_{0.05}$  denotes the upper 5% percentile of a standard normal distribution,  $S_t$  the standard deviation of the sample, and  $\bar{t}$  is the arithmetic mean of the selected pre-epidemic values:

$$\bar{t} = \frac{\sum_{i=1}^n \sum_{j=1}^N t_{(i),j}}{n \cdot N}$$

The epidemic threshold is defined as the upper limit of the confidence interval as:  $t^* = \bar{t} + z_{0.05} \cdot S_t$  where  $S_t^2 = \sum_{i=1}^n \sum_{j=1}^N (t_{(i),j} - \bar{t})^2 / (n \cdot N - 1)$

This is used to calculate the pre-epidemic threshold, and in the same way we can calculate the post-epidemic threshold. The post-epidemic baseline can be found by calculating the mean of all post-epidemic rates. Finally, coming to the third step in the MEM method, rates in the post-epidemic period are expected to be lower than the post-epidemic threshold rate.

By using the  $n$  highest values for each epidemic period to build a sample of  $n \cdot N$  rates, thresholds can be evaluated for different levels of intensity. Using the arithmetic mean and 50%, 90% and 95% confidence levels with each sample, one-sided confidence intervals are evaluated to give upper limit thresholds for the different levels of intensity of the epidemic.

Some recent publications have been using the MEM method in different health institutions to detect influenza outbreaks in different places such as Spain [12], Europe [161], the United Kingdom [59,65], the United States [39], Cambodia [88], Portugal [115], France [120] and Scotland [100].

### 3.4 Cumulative Summation (CUSUM) Method

The CUSUM (or cumulative sum control chart) is a sequential analysis technique in statistical quality control. It was developed by E. S. Page of the University of Cambridge. The CUSUM makes use of the cumulative sum of distance from a standard level. The CUSUM chart draws the cumulative sum of distances from the standard level for the separate observations. It is usually used for monitoring change detection [61]. The Moving Average Cumulative Sums (Mov. Avg Cusum) method [146] is a method based on regression or time series analysis of ILI data. This method relies on use of a series of stationary data. New detection methods based on advanced statistical approaches, especially those aimed at real-time monitoring and projecting of influenza cases, are hindered when there is a change in level of rate of occurrence without this signifying an epidemic. The main challenge is how to use ILI surveillance data with a simple and efficient method to raise an alarm in the early stages of the pandemic influenza.

We describe several applications of the CUSUM method, to have a clear view about it, as many recent papers use this approach [37,40,122,146,192]. Assume that we observe a sequence of random variables  $X_1, X_2, \dots$ . Suppose the first observations  $x_1, x_2, \dots, x_{n-1}$  correspond to a pre-epidemic state, with a given distribution function for the rates, while the observations after  $x_{n-1}, x_n, x_{n+1}, \dots$ , correspond to an epidemic state, with a different distribution function for the rates. Say  $F_B$  is the first distribution function, for the observations  $x_1, x_2, \dots, x_{n-1}$ , and  $F_E$  is the distribution for  $x_n, x_{n+1}, \dots$ , with  $F_B \neq F_E$ . Assume



that we know all parameters except  $n$ ; that is, the distribution function changed but we do not know the time of this shift, so the aim is to determine when the shift has occurred the first time. The CUSUM approach finds this time  $n$  using the last  $k$  where:

$$S_n = \sum_{i=1}^n \log \frac{f_E(x_i)}{f_B(x_i)} - \min_{k \leq n} \sum_{i=1}^k \log \frac{f_E(x_i)}{f_B(x_i)} > L,$$

where  $f_B$  and  $f_E$  are densities corresponding to  $F_B$  and  $F_E$  respectively, and  $L$  is a constant that sets the operating characteristics of the system.  $L$  does not depend on  $n$ . The  $S_n$  can be calculated as:

$$S_n = \max \left( S_{n-1} + \log \frac{f_E(x_n)}{f_B(x_n)}, 0 \right).$$

The incipient point for the CUSUM,  $S_0$ , will be some number between 0 and  $L$  [67].

This method is used to detect sudden changes to the arithmetic average of a set of specific values and is commonly used in industrial operations control to supervise the level of production quality [135].

Let  $y_t, t = 1, 2, \dots$  be a series of observations. Then the upper CUSUM limit over a  $d$ -week period can be defined at time  $t, C_t^+$ , as:

$$\text{CUSUM}_t^+ = \max \left\{ 0, \frac{y_t - \tilde{y}_{(7)}}{\tilde{S}_{(7)}} - k + \text{CUSUM}_{t-1}^+ \right\}$$

with CUSUM  $C_{t-d}^+ = 0$  [37, 94]. The parameter  $k$  gives the lowest standardized difference from the operating mean to be ignored by the system. The operation mean  $\tilde{y}_{(7)}$  and also the operation variance  $\tilde{S}_{(7)}^2$  are calculated from the 7 weeks series  $y_{t-d-7}, \dots, y_{t-d-1}$  prior to the most recent  $d$  weeks.  $\Phi_{(1-\alpha/2)}$  where  $\Phi$  is a standard Normal deviate ( $z$  value), is the pre-specified threshold used to control the alarm. When CUSUM  $C_t^+$  surpasses the identified threshold; the alarm will be raised [37]. This CUSUM procedure assumes that the counts are normally distributed. There are other variations. Singh et al. in 2010 [142] used this method by analysing data from the 2001-02 season through to the 2008-09 season of seasonal ILI cases. These are normally reported weekly over a period of 33 weeks (from the first week of October to the third week of May) in different age- and sex-classes, by sentinel general practices (GPs) across Scotland. Scottish Enhanced Respiratory Virus Infection Surveillance (SERVIS) sentinel GPs cover all health boards in Scotland. They compared the performance of this method with the pandemic detection algorithm and the basic method called the threshold method [34, 137]. This showed the CUSUM method to

perform well, but it did not outperform other methods used in the same study.

Griffin and others [60] used the CUSUM method to generate early detection of influenza outbreaks using the District of Columbia Department of Health's syndromic surveillance system, starting from monitoring the daily statistics  $S_i$  defined as:

$$S_i = \max\left(0, S_{i-1} + (X_i - \mu) - k\right)$$

where  $X_i$  indicates the observation taken in day  $i$ ,  $\mu$  presents the mean daily count estimated from the data, and  $k$  is an off-set parameter selected by the user. An alarm is set off at time  $\hat{\tau} = \inf\{i : S_i > h\}$ , where  $h$  is calculated experimentally to guarantee a fixed false positive rate (specified by the user) in times out of the influenza seasons.

This method has become widely used in different health institutions in different countries such as China [189], Taiwan [49], Australia [169] and in different studies [53, 70, 114, 138].

Also as this method is robust, it has been used in recent published papers for comparison with other methods, e.g. [2, 5, 17, 27, 28, 89, 167, 168].

### 3.5 Hidden Markov Model

The Hidden Markov Model (HMM) can be defined as a twofold stochastic operation with an implicit stochastic operation which is not observable (it is hidden) but with other stochastic operations that output the series of observations [130]. Hidden Markov Models (HMMs) are a suitable statistical instrument to explain events showing serial dependency in data.

Let  $Y$  be  $y_t$  observations,  $t = 1, 2, \dots, n$ , which are an output of the stochastic operation. The general idea of the hidden Markov model is to join with each  $Y_t$  a hidden random state variable  $S_t$  to determine the conditional distribution of  $Y_t$ ; if  $S_t = j$ , the conditional distribution of  $Y_t$  will be  $f_{j_t}(y_t; \theta_j)$ , where  $f_{j_t}$  belongs to a family of certain distributions and  $\theta_j$  are the parameters which we aim to estimate. Assuming that the unobserved series  $S_t$  follows an  $m$ -state identical Markov chain of order 1 with stationary transition probabilities, then

$$\alpha_{ij} = P(S_t = j | S_{t-1} = i), \quad i, j = 1, \dots, m.$$

Given the heterogeneity, we note that the Markov chain is stable and homogeneous, but the reverse is not always true. The variables  $Y_t$  are independent conditionally on the random variables  $S_t$ . In this Markov-dependent mixture model the sequence of states is

unobserved, so the mixture model is called a hidden Markov model. If there are two underlying distributions correlating with (ILI or ARI) incidence rates, then  $y_t$  explains the observed (ILI or ARI) incidence rate in week  $t$  and is imagined to be created from one of the two distributions, and  $\alpha_{12}$  explains the switching probability from the first distribution 1 to the second distribution 2 in two sequential weeks [81].

Hidden Markov Models have been used with Bayesian methods to estimate the model parameters by Watkins and others [170] for disease surveillance of case counts reported daily (hepatitis A) in postcode areas in Western Australia. They found that the HMM is more highly responsive to the accumulation of events close in time than the cusum-based algorithms. Also the false alarm rate was less than 0.1.

In a previous study by Rath and others [132] the HMM method was used to detect influenza in France using ILI (1985-1996). They decided to use two states in the HMM, where the first state with non-epidemic rates was modelled with an Exponential distribution and the second state with epidemic rates was modelled with a Gaussian distribution. They found their model described the data well without needing to model seasonal and trend effects. They concluded that a better adjustment to the variations in the data is provided by the Exponential distribution, which also avoided the need for complicated parameter assessment that would be needed with Gaussian models.

### 3.6 Weekly Cases Ratio (WCR) method

One way to detect pandemic influenza is the method of Singh et al. published in 2010 [142]. This *WCR* algorithm is based on two main metrics obtained from weekly reported cases (e.g. ILI or ARI), namely the *WCR* and  $N_{HB}$  metrics:  $N_{HB}$  is the number of health boards (HBs) which had a reported increase in cases in the current week  $w$  compared to the previous week  $w - 1$ , i.e. which had a value of *WCR* above 1, where weekly case ratio (*WCR*) can be defined as:

$$WCR = \frac{\text{Rate in week } w}{\text{Rate in week } w-1} \quad (\text{defined for week 2 onwards}) \quad (3.6.1)$$

where Rate is the total rate of ILI or ARI cases reported to all SERVIS sentinel GPs. Because *WCR* is based on the previous week's reported ILI or ARI cases in the denominator, zero can simply be replaced in the denominator of equation 3.6.1 by 1.

Singh et al. [142] obtained the joint pmf of a discretised version of *WCR* (using bins of size

0.1) and  $N_{HB}$ . This pmf was based on large scale simulated data generated using Poisson distributions with a mean given by the weekly case ratio for a given HB from 6 seasons of historical data. A threshold probability is chosen and an alarm is raised in the first week when the joint pmf value of an observation of  $(WCR, N_{HB})$  lies below this threshold.

Testing the algorithm for specificity means checking if it does not detect a pandemic when no pandemic is occurring. To check this, firstly calculate  $WCR$  and  $N_{HB}$  for each week in the ILI (or ARI) time series. Secondly, with the chosen threshold probability value  $\delta$ , we count how many non-epidemic weeks there are in which the joint probability of the observed  $(WCR, N_{HB})$  calculated from the pmf of the simulated data falls below the threshold. The total counts give the number of false alarms ( $FA_s(\delta)$ ) in season  $s$ . Theoretically,  $FA_s^{max}$ , the maximum number of false alarms in the season  $s$ , is the total number of weeks minus 1 (because we cannot calculate  $WCR$  in the first week). Finally, the specificity of the detection algorithm for a given threshold probability can be calculated as:

$$1 - \left( \frac{\sum_{s=seasons} FA_s(\delta)}{\sum_{s=seasons} FA_s^{max}} \right)$$

The probability of threshold  $\delta$  gives rise to a pre-set specificity of this detection method. This  $\delta$  defines the detection threshold. To compare this algorithm with other algorithms with the same constant specificity,  $\delta$  can be adjusted in order to give realised specificities of 95% and 99%, for example.

### 3.7 Farrington method

The seasonal fluctuations of organisms causing diseases should be detected for public health reasons. According to Farrington et al. [47, 51], any increase on top of expected patterns should be noted and any count that is abnormally high should be focused on. Moreover, a routine scanning system that entails sensitivity, timeliness, specificity as well as easily interpretable outputs should be used. This will ensure that outbreaks are detected in time and necessary interventions are taken. This should be done in a way that the confidence of the system is not compromised [112, 131]. The three components (sensitivity, timeliness and specificity) determine the statistical features of the system. Data should be collected over a week and should undergo epidemiological investigations at the end of the week. A single robust algorithm should be used for all organisms. The system has to be sufficiently flexible to accommodate the wide range of organism seasonal patterns and underlying trends. With this view, a log-linear regression model was chosen and it was adjusted based on overdispersion, seasonality, secular trends and past outbreak of a given disease. The expected value for the week is calculated based on past data as well as a

threshold above which the outcome will be rated unusual [47].

The basis of the system was a network of laboratories, including the 49 laboratories of the Public Health Laboratory Service and over 200 National Health Service laboratories, which reported to the Communicable Disease Surveillance Centre (CDSC) in London. In this method, a log-linear regression model was used, taking into account overdispersion, seasonality, secular trends and past outbreaks of weekly data from the CDSC for more than one organism (such as *rotavirus*, *Clostridium difficile*, *Salmonella derby*, *Shigella sonnei*, *influenza B* and *Salmonella typ*).

One of the methods used in outbreak detection is the regression method. It relies on laboratory outcomes and reported infections for detecting outbreaks in a surveillance system. The approach basically uses standardised residuals. The threshold value is obtained from the distribution of the residuals when an outbreak does not happen. This method can be taken as an extension of the Shewhart chart. The Parametric model in which the projected infection count in a specific month  $x$  is computed as the average of the recorded count at months  $(x - 1)$ ,  $x$  and  $(x + 1)$  for a specified period of years, is another method used. The approach adopted ensures that effects that happen seasonally are catered for by the method instead of the explicit modelling hence an aspect of quality is provided [156].

When recording statistical data, the reference date is taken to be the reporting date on which the specimen is taken from the patient. This eliminates biases as a result of delay in reporting [47, 132, 173]. Despite this, the timeliness and sensitivity of the detection mechanism is influenced due to delay between infection and reporting as an additional source of variability. The effect of delay can be investigated formally as follows: Take  $F(x)$  to be the accumulated distribution of delay in reporting with  $x$  denoting the span between infection of an individual and the reporting date of the organism and let  $X(t)$  represent specimens collected at a period  $t$ . The expected number of reports at time  $t$  is given by:

$$Y(t) = \int_0^t X(u) f(t - u) du$$

where  $f(x)$  represents the density function of the reporting delay distribution, taking  $X(t) = 1$  for  $t \in [0, \partial]$  and  $X(t) = 0$  otherwise. It follows that if the threshold value is  $\alpha < 1$  which  $Y(t)$  must attain for the outbreak to be detected, then the detection time  $t_d$  satisfies the equation:

$$\alpha = F(t_d) - F(t_d - \partial)$$

where  $F(t - \partial) = 0$  for  $t < \partial$  [47].

Outbreaks of short duration are detected close to the mode  $t_m$  of  $f$  provided  $f(t_m) > \alpha/\partial$ . Long duration outbreaks are detected at  $t_d = F^{-1}(\alpha)$  [47].

The Farrington methods of detecting influenza are elaborate and reliable. The statistical approach makes the methods quantifiable as well. The Farrington methods have been widely used in other parts of the world such as Hong Kong and China [6]. Other researchers have equally considered similar approaches in influenza detection. Amorós et al. [6], Conesa et al. [33], and Cox et al. [38] consider statistical approaches in detecting influenza outbreak.

Deeper review of the literature reveals that the Farrington method is being used to detect influenza in some countries [48, 71, 90, 122, 132, 176].

System	Advantage	Disadvantage	Example of when and where used
Threshold Method	Simple, widely used.	Easily affected by historical data.	Scotland, England and Wales in 2009, Sweden 2014.
Time Series Technique	Simple.	Easily affected by historical data, not widely used currently.	Egypt 2014, The United States 2013.
Moving Epidemic Method (MEM)	Robust, accurate.	Relatively affected by historical data.	Europe 2015, Scotland 2016.
Cumulative Summation (CUSUM) Method	Robust, simple, widely used.		China 2017, Taiwan 2017, Australia 2008.
Hidden Markov Model		Not widely used recently.	France 1985.
Weekly Cases Ratio ( $WCR$ ) Method	Perform well with rapid change.	Need huge simulations with each use.	New method.
Farrington Method	Robust, accurate.	Performs best with repeated same scenarios.	Hong Kong 2015.

Table 3.1: Short list of advantages, disadvantages and examples of when and where the detection systems are used.

In the Farrington method performing best with repeated same scenarios is considered as a

disadvantage because influenza does not occur in the same scenario every year (especially when we think about outbreaks).

### 3.8 Conclusion

In this chapter, we have discussed and reviewed a number of disease surveillance methods that have been proposed for identifying outbreaks of infectious diseases as they emerge in the very early stages of the epidemic. Methods which are suitable for use to detect an influenza epidemic include the threshold method, ARIMA method, Moving Epidemic Method (MEM), Cumulative Summation (CUSUM) Method, Weekly Case Ratio (*WCR*), Hidden Markov Model and Farrington method. This review shows that some range of statistical methods have been proposed for prospective outbreak detection.

The choice of which statistical method to use depends on the intended application. Methods that are designed for the surveillance of a single infection or syndrome are tuned to the specific features of that infection and may need frequent user intervention. By contrast, systems designed for routine application for hundreds or thousands of potential infections, with different frequencies and time patterns, will require robustness and automation – such systems like the Farrington method and general time series methods are unlikely to be optimal in the case of flu surveillance.

Systems will also vary according to the features they are designed to detect; for example, trends and seasonal change may be relevant depending on the context. No statistical technique is likely to be ideal in every setting.

The threshold method is a simple method for deciding a threshold for any infectious diseases, especially seasonal and pandemic influenza, and aspects of this method are present in the MEM method. The ARIMA method and Hidden Markov Model have not been widely used recently to detect influenza cases each year. The Farrington method works very efficiently when we have a recurring scenario each year but influenza, as we have seen in historical data, has no fixed and repeated scenario every year. For example, in the Scotland Data we see the increase starting in October sometimes, sometimes in November, and sometimes there is a peak in January.

CUSUM is widely used in comparing detection systems in many publications, e.g. [2, 5, 17, 27, 28, 89, 167, 168] and is also used in comparison with the original *WCR* method [142]. MEM is now widely used in many European countries [12, 59, 161, 162] and by Health Protection Scotland (which supplied our data).

To compare the performance of any methods, we need to choose robust methods to get fair comparison. For the reasons provided above, in the next chapter, we will apply some

of these methods, namely MEM (in section 3.3) and *WCR* (in section 3.6) with real data. CUSUM is also used later (in chapters 7, 8 and 9).



## Chapter 4

# Applying and modifying detection systems

In this chapter, we apply some of the well-known established detection systems. Here, we will show how the Moving Epidemic Method (MEM) (section 3.3) and Weekly Cases Ratio (WCR) method (section 3.6) work with some real data (examined in chapter 2). We selected here the ARI and the ILI data from all 14 health boards in Scotland to apply these methods.

Since the *WCR* method is the subject of study for this thesis, we also analyzed the values of *WCR* in summer to compare them with the corresponding values in winter, and check whether the data could be normally distributed. We also linked the *WCR* values with  $N_{HB}$ , in summer and winter, and check whether these follow the same joint distribution. These results are used in the development of our extended *WCR* system in chapters 5 to 9.

### 4.1 Applying the MEM method to the data for Scotland

In this section, we show how MEM works and how to use the data from the 2009-2010 to 2013-2014 winter seasons to detect an increase in the 2014-2015 winter. We selected data for each winter, from the start of October to the end of May in the following year (we have used this period for winter at this stage, but we decided later on to use the end of March as the end of winter because there were lots of zeroes in the consultations reported in many HBs during April and May). Weekly aggregate numbers of acute respiratory infection (ARI) and influenza-like illness (ILI) were used for the 14 health boards in Scotland. The general idea can be written in three parts: firstly, dividing each winter into three

periods as follows: the period before the epidemic (pre-epidemic), the epidemic period and the period after the epidemic (post-epidemic); secondly, calculating a baseline and the epidemic threshold using the pre-epidemic data and post-epidemic data from historical data; and finally, computing thresholds for different levels of intensity.

In applying MEM to the data, we used the default setting of `mem` package [160] in R [129] which are the threshold parameter calculated from the arithmetic mean and point 95% confidence interval (standard deviations).

For each health board, we used ARI and ILI historical data from the winter of 2009-2010 to the winter of 2013-2014 to describe the pre-epidemic and post-epidemic thresholds. After that we compared those thresholds with the real period of the 2014-2015 season (pre-epidemic, epidemic and post-epidemic). The pre-epidemic threshold is expected to be higher than the pre-epidemic period rates, and the post-epidemic rates are expected to be lower than the post-epidemic threshold. Then we can assess rate of false alarms, the detection lag and timeliness for the method using the 2014-2015 season. We can define some terms which we will use with the 2014-2015 season [162]:

- Alert week: The first week in the 2014-2015 season which shows a rate above the pre-epidemic threshold. We consider this as the start week of epidemic.
- False alarm: A weekly rate from the modelled pre-epidemic period of the 2014–2015 season which is above the pre-epidemic threshold.
- Detection lag: The delay between the alert week and the start of the epidemic period modelled by MEM for the 2014-2015 season. If there is no delay, the lag is zero.
- Timeliness: The number of weeks between the first week of the epidemic period modelled by MEM for the 2014–2015 season and the alert week; either the alert week is after the start of the epidemic period or vice versa.

For example, if the MEM system raised the alarm in week 7, and the first week of the epidemic period modelled by MEM is week 7, then the detection lag will be zero and timeliness will be zero. If the MEM system raised the alarm in week 10, and the first week of the epidemic period modelled by MEM is week 7, then the detection lag will be  $10-7=3$  weeks and timeliness will be  $10-7=3$  weeks. If the MEM system raised the alarm in week 2, and the first week of the epidemic period modelled by MEM is week 7, then the detection lag will be zero (because the alarm was already raised before the epidemic) and timeliness will be  $7-2=5$  weeks.

These last two are often similar in our results.

Cross-validation was used to evaluate the accuracy of the method [162]. Each single season was taken away from the historical data, to be used as the objective season of interest for

each health board. The beginning and the end of the epidemic period of the 2014-2015 season and the pre- and post-epidemic thresholds on the basis of the remaining seasons, excluding the target season, were calculated by MEM. Then we repeated this procedure for each season available per health board.

The rates of the target season before and after the epidemic period set by MEM were compared with the thresholds calculated using all historical information to assess the quality of the method. The following measures were used to assess performance of the threshold:

1. Sensitivity: calculated as number of epidemic weeks above the pre-epidemic threshold (before the peak) and number of epidemic weeks above the post-epidemic threshold (after the peak) divided by the epidemic length in weeks.
2. Specificity: calculated as number of non-epidemic weeks below the pre-epidemic threshold (before the peak) and number of non-epidemic weeks below the post-epidemic threshold (after the peak) divided by the number of non-epidemic weeks.
3. Positive predictive value (PPV): number of epidemic weeks above the threshold divided by the total number of weeks above the threshold.
4. Negative predictive value (NPV): number of non-epidemic weeks below the threshold divided by the total number of weeks below the threshold.
5. Median timeliness: Median of the time to detection.

The ARI and ILI data from the 14 health boards with a period from October 2014 to the end of May 2015 were selected as the target season. The results for the modelled 2014-2015 season are summarised in table 4.1 for the ARI data and in table 4.2 for the ILI data, and are also plotted by health boards in figure 4.1 for ARI data and in figure 4.2 for the ILI data. Considering the historical data from 2009-2010 to 2013-2014, for ARI data, the pre-epidemic threshold per 100,000 population ranged from 86.45 in GGC to 260.49 in SH and from 3.08 in LO to 24.61 in OR for ILI data. Post-epidemic thresholds were a little different to pre-epidemic thresholds. For most HBs these varied in a range of  $\pm 10\%$  above or below the pre-epidemic ones except for GGC, SH and WI for ARI data and LN, LO and OR for ILI data. In tables 4.1 and 4.2, we can see some missing values in the detection lag and timeliness, such as for DG, LN, LO and SH for the ARI data and OR for ILI data. The reason for this is that there were high rates in the previous seasons, which affects the pre-epidemic threshold and leads to failing to raise an alarm.

No.	HB	Pre-threshold ( $n/100,000$ )	Post-threshold ( $n/100,000$ )	length of epidemic 2014-2015	False alarms 2014-2015 ( $n$ )	Detection lag 2014-2015 ( $w$ )	Timeliness 2014-2015
1	AA	123	118	19	0	3	0
2	BR	150	144	20	0	3	0
3	DG	204	204	16	0	0	0
4	FF	130	136	17	0	0	0
5	FV	106	129	19	0	0	0
6	GGC	86	139	20	0	5	0
7	GR	107	111	18	0	5	0
8	HG	185	205	18	0	0	0
9	LN	133	135	17	0	6	0
10	LO	91	95	17	0	6	6
11	OR	89	98	16	0	0	0
12	SH	260	447	10	0	1	1
13	TY	160	172	18	0	3	0
14	WI	223	358	13	0	0	0
Tot. Scot		118	122	18	0	3	0.5

Table 4.1: Pre-epidemic and post-epidemic thresholds based on historical data (ARI) until 2013-2014, and comparison with the observed 2014–2015 season;  $n$  denotes number of cases and  $w$  are weeks.

No.	HB	Pre-threshold ( $n/100,000$ )	Post-threshold ( $n/100,000$ )	length of epidemic 2014-2015	False alarms 2014-2015 ( $n$ )	Detection lag 2014-2015 ( $w$ )	Timeliness 2014-2015
1	AA	4	4	14	0	0	0
2	BR	8	4	14	0	5	0
3	DG	7	7	12	0	0	0
4	FF	4	5	13	0	0	0
5	FV	4	6	14	0	0	0
6	GGC	3	6	11	0	4	0
7	GR	4	5	13	0	0	0
8	HG	10	9	12	0	0	0
9	LN	4	7	11	0	3	0
10	LO	3	6	15	0	9	0
11	OR	25	15	13	0	1	1
12	SH	6	13	9	0	2	2
13	TY	5	7	13	0	0	0
14	WI	20	15	11	0	8	0
Tot. Scot		3.6	6.3	11	0	2	0.21

Table 4.2: Pre-epidemic and post-epidemic thresholds based on historical data (ILI) until 2013-2014, and comparison with the observed 2014–2015 season;  $n$  denotes number of cases and  $w$  are weeks.

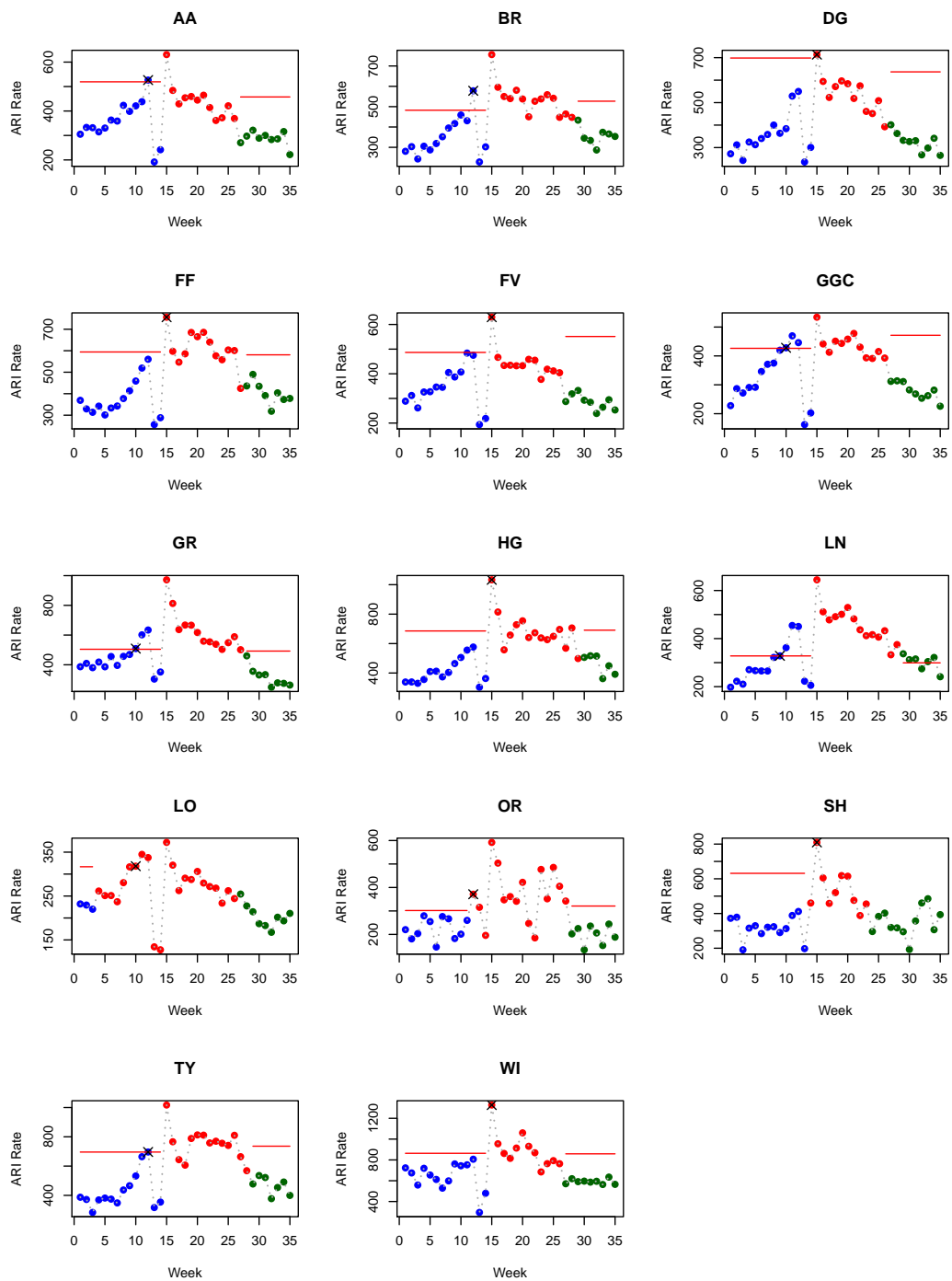


Figure 4.1: Epidemic threshold, levels of intensity and modelled ARI season 2014-2015, by HBs. MEM plots include the pre-epidemic period (blue), epidemic period (red), and post-epidemic (green), while the triangle gives the alert week and the red line presents the pre-epidemic threshold, while the post-epidemic threshold is presented by the green line. A 5 year series (2009/10 to 2013/14) is used in MEM.

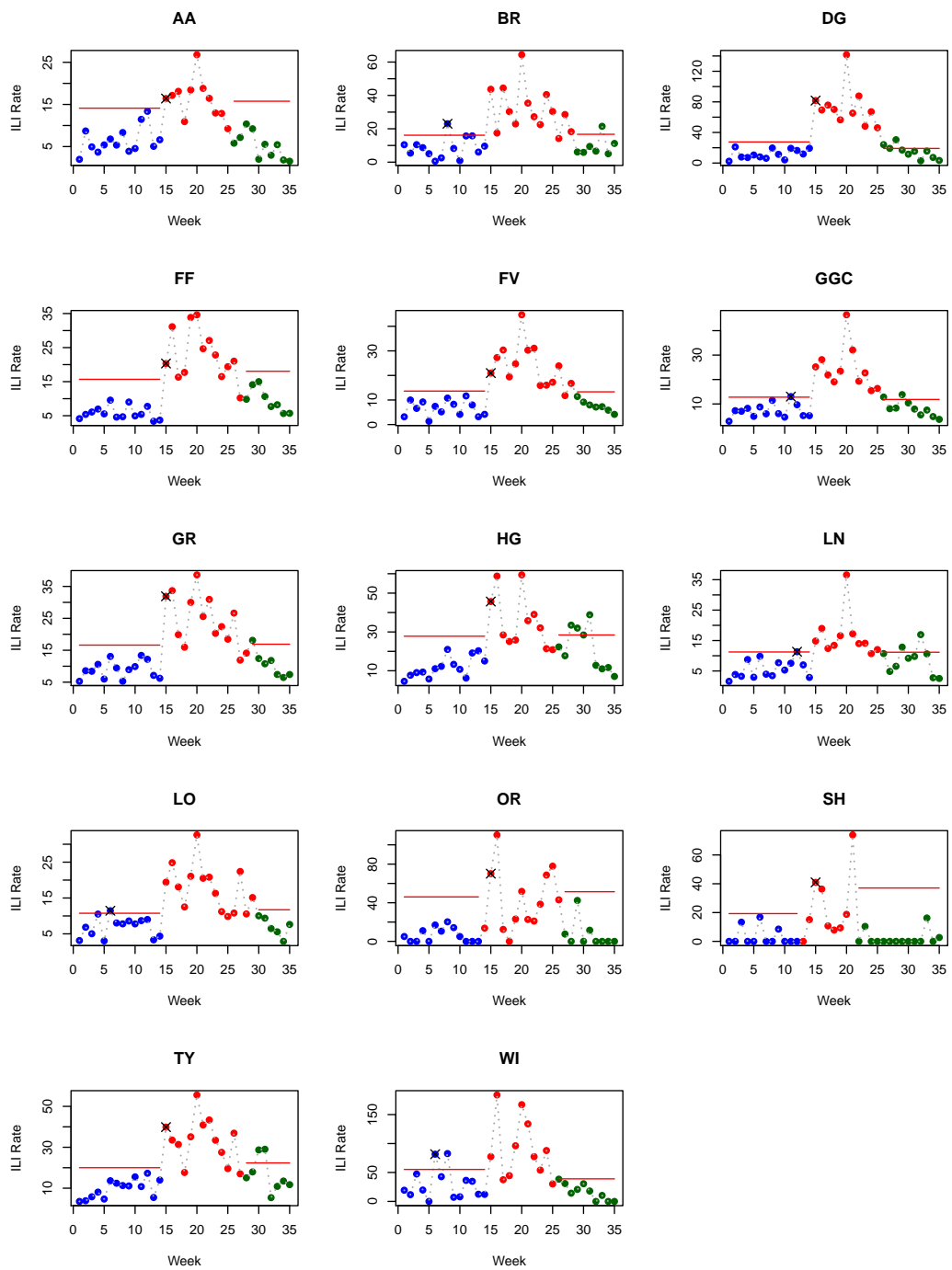


Figure 4.2: Epidemic threshold, levels of intensity and modelled ILI season 2014-2015, by HBs. MEM plots include the pre-epidemic period (blue), epidemic period (red), and post-epidemic (green), while the triangle gives the alert week and the red line presents the pre-epidemic threshold, while the post-epidemic threshold is presented by the green line. A 5 year series (2009/10 to 2013/14) is used in MEM.

Now, we want to see how MEM works with the total ARI and ILI data. Firstly, we apply

MEM for ARI and ILI for 5 years of historical data (from October to the end of May) to detect any increase in winter 2014-2015, then we reduce the number of years of historical data. Figure 4.3 shows how MEM works to detect the pandemic in 2014/2015 using 5, 4, 3 and 2 years historical data. Sensitivity and specificity for the same periods are shown in figure 4.4. We see that the time of detection for ILI data is shorter than for ARI data, which gives an advantage to using ILI data, according to median time of detection. Sensitivity and specificity are better using ARI data in these periods. The second step is trying to consider each two years to detect the following year, then calculating sensitivity and specificity to see the overall sensitivity and specificity curve. From the different curves in figure 4.4, we have not found a curve that we can judge as the best in performance because, as we observe, if we find high sensitivity we find low specificity at the same time.

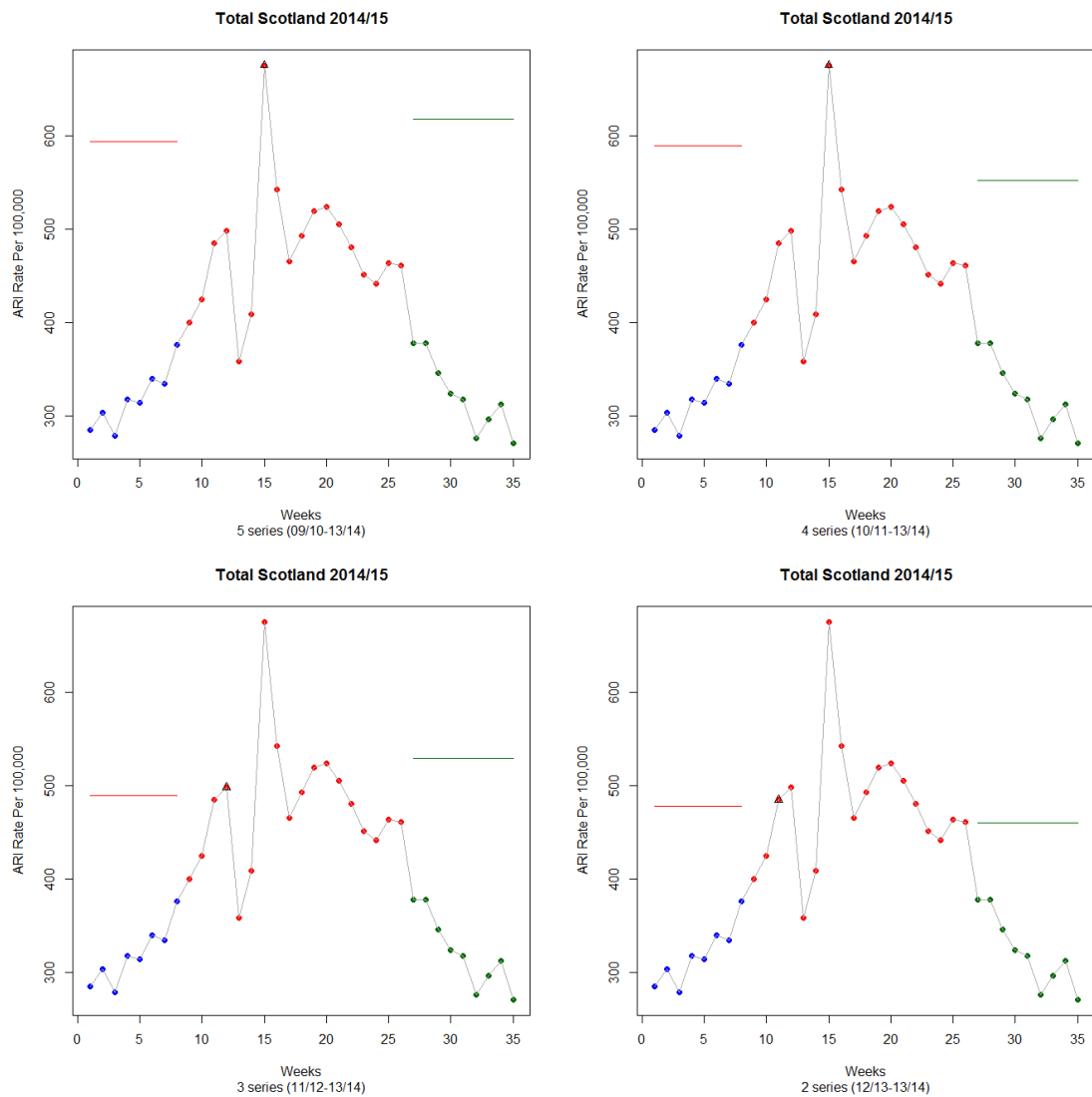


Figure 4.3: MEM figures for ARI data using 5, 4, 3 and 2 years historical data to detect a pandemic in winter 2014/15. MEM plots include the pre-epidemic period (blue), epidemic period (red), and post-epidemic (green), while the triangle gives the alert week and the red line presents the pre-epidemic threshold, while the post-epidemic threshold is presented by the green line. We used the default setting of `mem` package [160] in R [129] which is the threshold parameter ( $\delta$ ) calculated from the arithmetic mean and point 95% confidence interval (standard deviations).



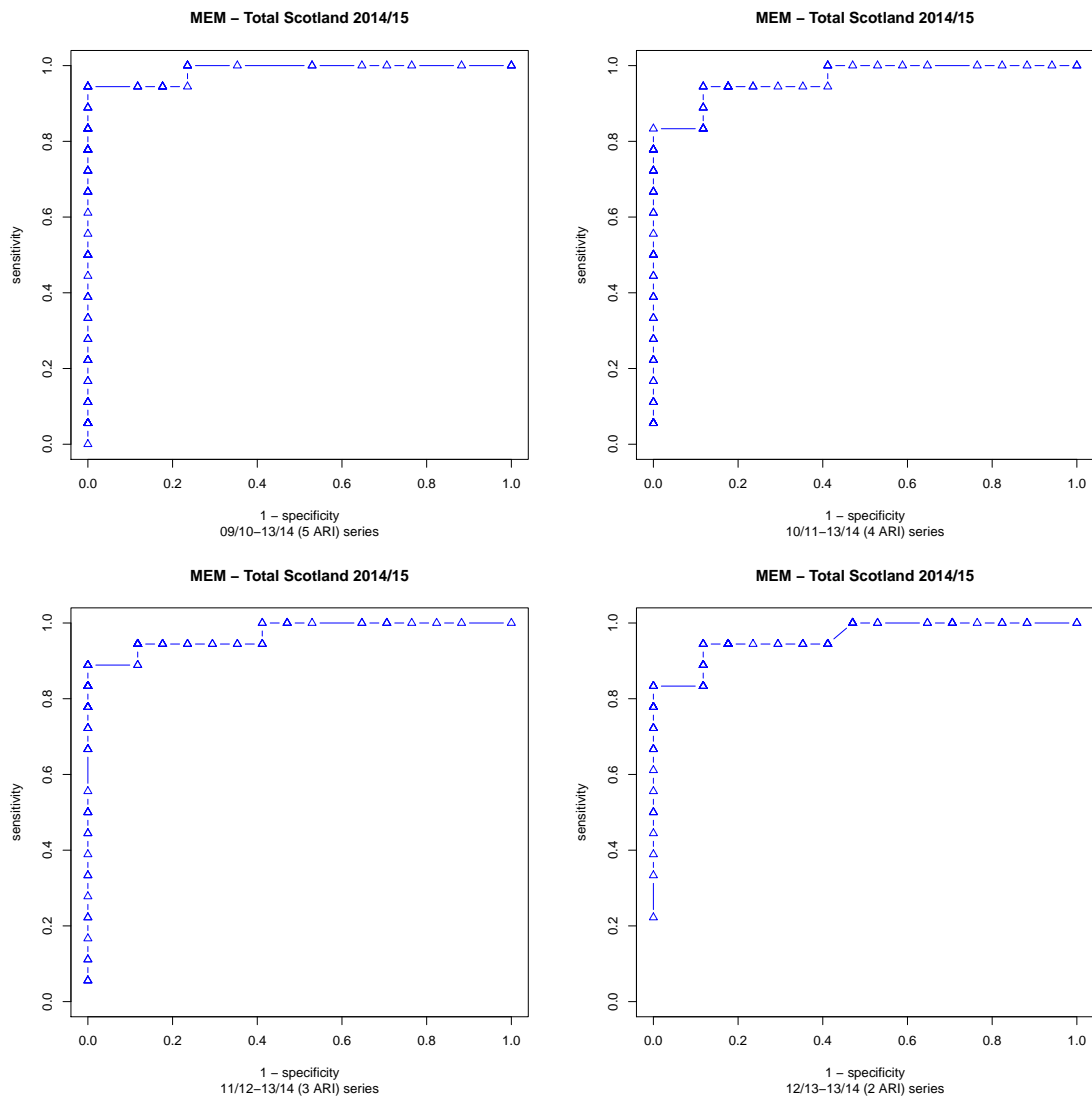


Figure 4.4: Sensitivity and specificity using MEM for ARI data with 5, 4, 3 and 2 years historical data to detect a pandemic in winter 2014/15.

As for the ARI data, we do the same with the ILI data to assess MEM's performance. MEM detects the outbreak with less timeliness, as presented in figure 4.5. However, the best performance is not achieved so far, because there is a clear imbalance when we review the sensitivity and specificity plots in figure 4.6. This may be due to the presence of weeks during the period identified by the system as a epidemic, but which have lower rates, less than the pre-epidemic threshold. This is clearly influenced by the calculation of sensitivity according to this system.

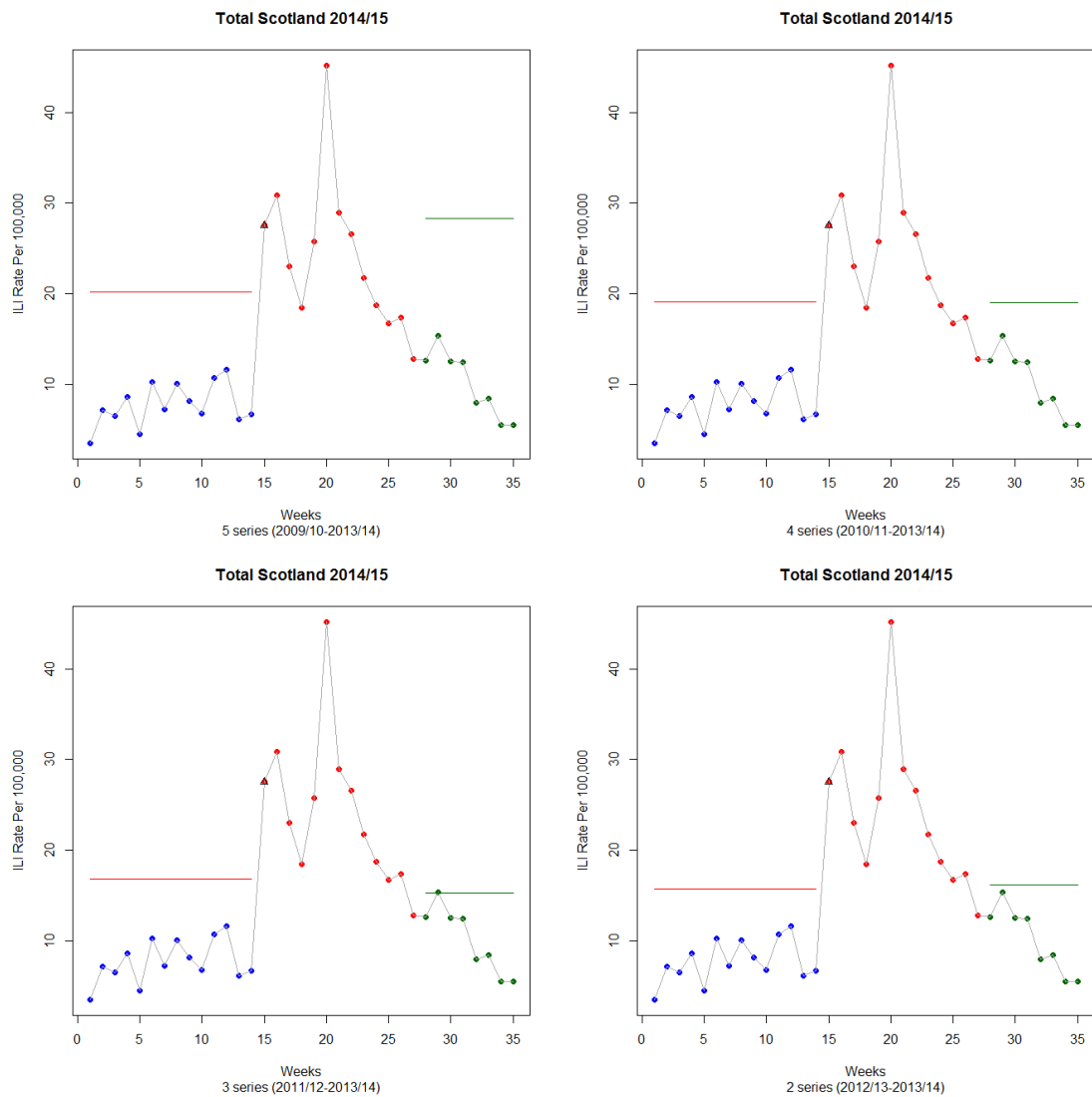


Figure 4.5: MEM figures for ILI data using 5, 4, 3 and 2 years historical data to detect a pandemic in winter 2014/15. MEM plots include the pre-epidemic period (blue), epidemic period (red), and post-epidemic (green), while the triangle gives the alert week and the red line presents the pre-epidemic threshold, while the post-epidemic threshold is presented by the green line. We used the default setting of `mem` package [160] in R [129] which is the threshold parameter ( $\delta$ ) calculated from the arithmetic mean and point 95% confidence interval (standard deviations).

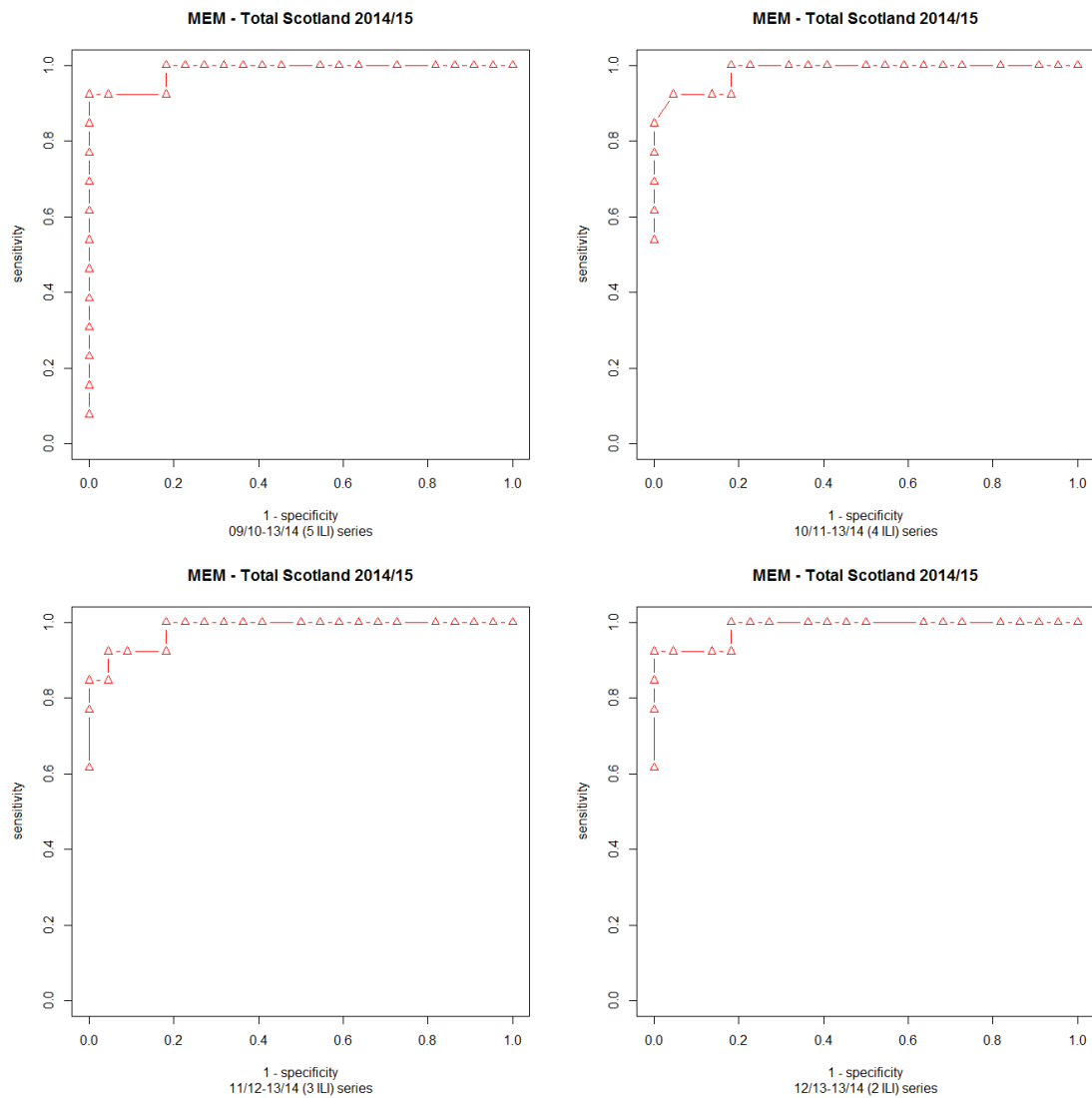


Figure 4.6: Sensitivity and specificity using MEM for ILI data with 5, 4, 3 and 2 years historical data to detect a pandemic in winter 2014/15.

To try getting better performance for detecting the outbreak (high sensitivity and specificity) with MEM, we use the 2009/10 and 2010/11 seasons to detect the next outbreak for 2011/12, and so on until we use 2012/13 and 2013/14 to detect the 2014/15 outbreak, so we have 4 MEMs. For the ARI data, figure 4.7 shows sensitivity and specificity plots of each MEM. For the ILI data, we see sensitivity and specificity plots of each MEM presented in figure 4.8.

Here there are similar results, as there is no curve which we can confirm as having the best performance. We have found a curve that we will say is the worst, which is the top left curve in figures 4.7 and 4.8. The simple reason for this is that including the 2009 data in the historical data will raise the pre-epidemic threshold and thus affect the sensitivity

directly. The remaining curves indicate a high sensitivity with low specificity.

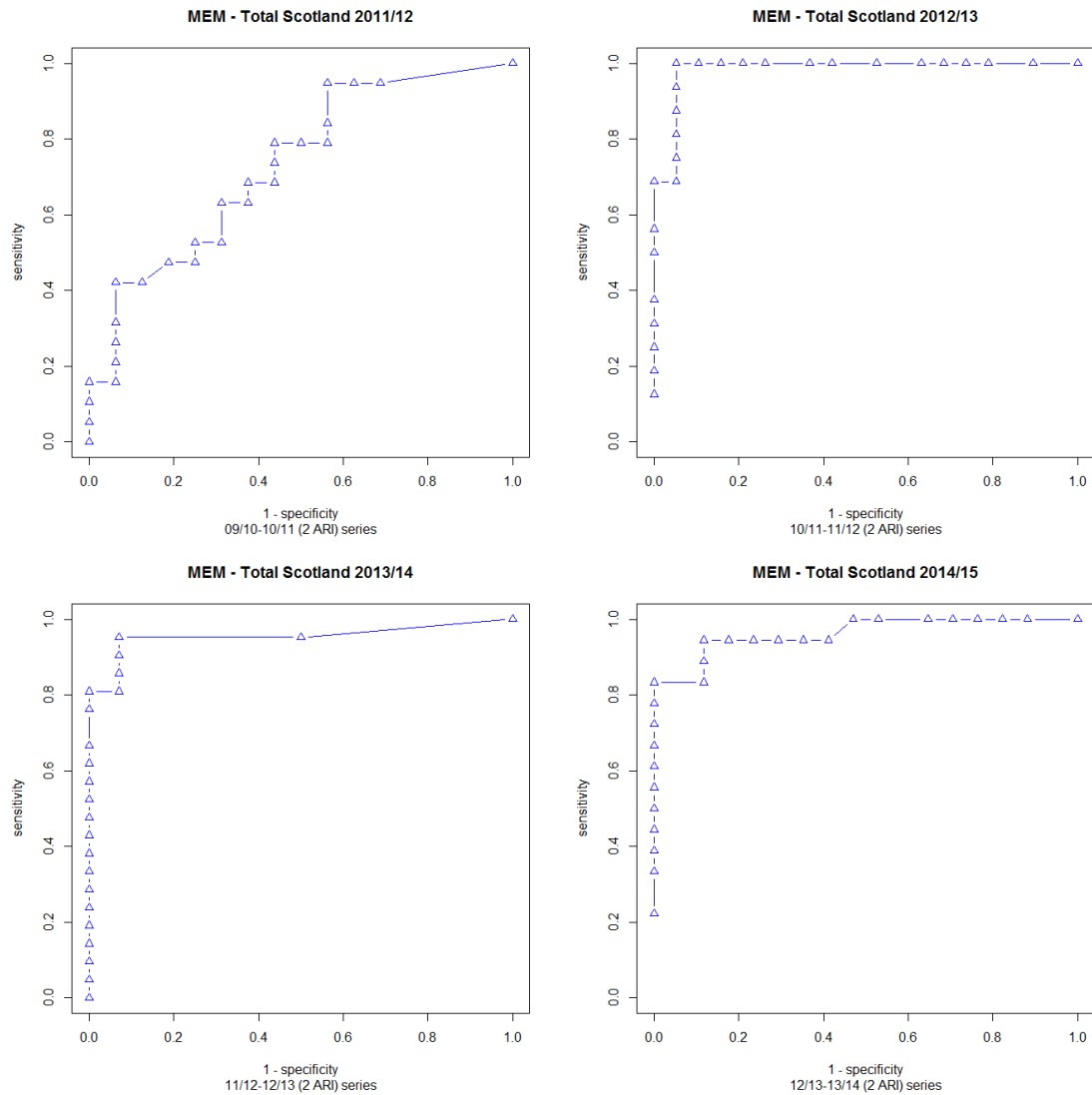


Figure 4.7: Sensitivity and specificity using MEM for ARI data with 2 years historical data to detect a pandemic in the following winter.

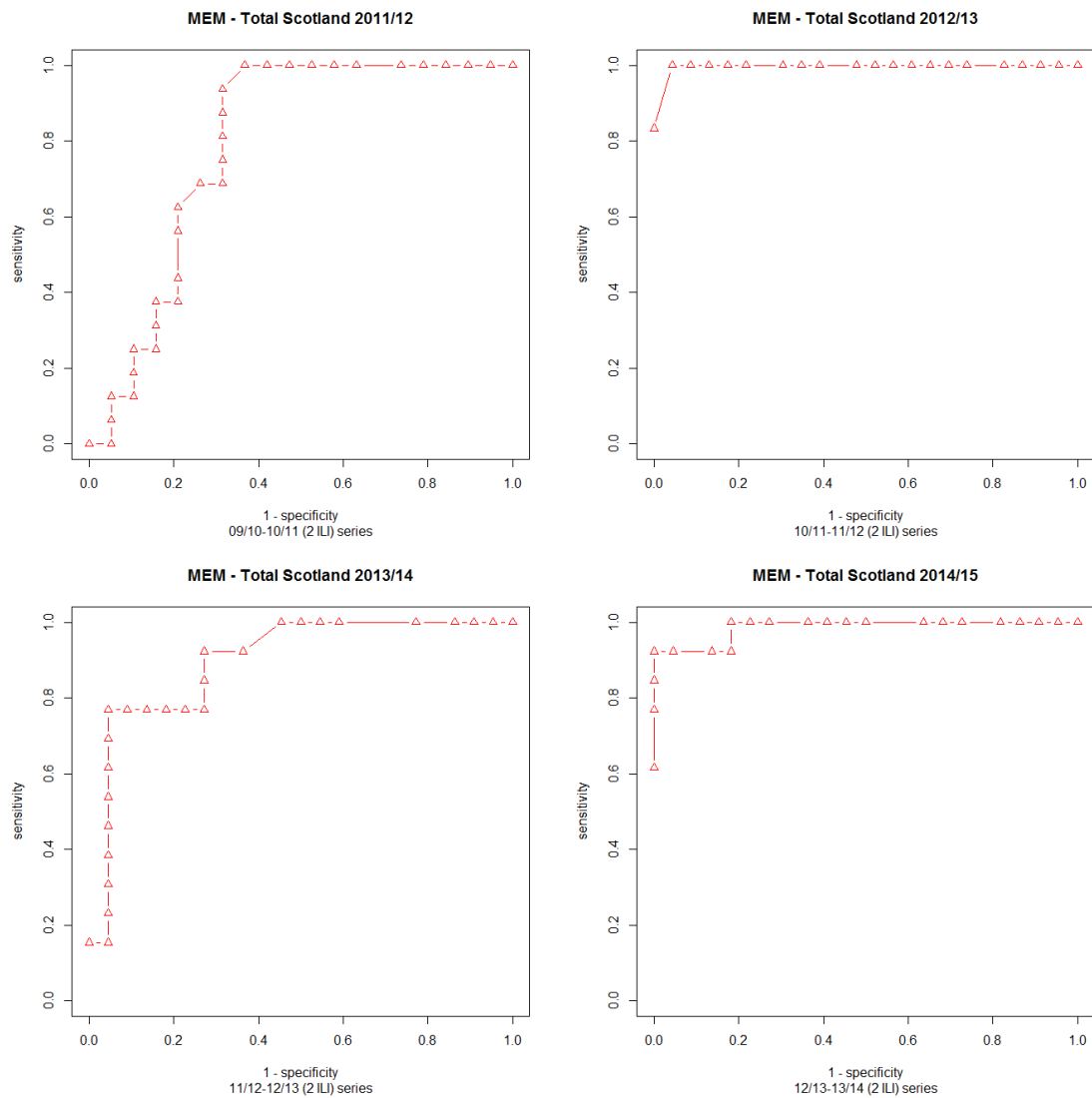


Figure 4.8: Sensitivity and specificity using MEM with ILI data with 2 years historical data to detect a pandemic in the following winter.

As we did not yet find a suitable plot to choose a cut-off point, we will consider another approach and see the performance. Now, we consider using 3 years of data to detect the following year, which mean we will have 3 MEMs. Figure 4.9 shows how this works with the ARI data, while figure 4.10 shows plots for the ILI data. Again, they are similar to when we used 2 years of historical data, as there is no curve we can confirm as having the best performance. We have found a curve that we will say is the best, which is the top left curve in figures 4.9 and 4.10. The remaining curves indicate a high sensitivity with low specificity. Combined performance (average of sensitivity and specificity) with ARI data in figure 4.9 (the bottom right curve) may indicate better performance than for the ILI data in figure 4.10 (the bottom right curve), and this may be because ARI data is more general.

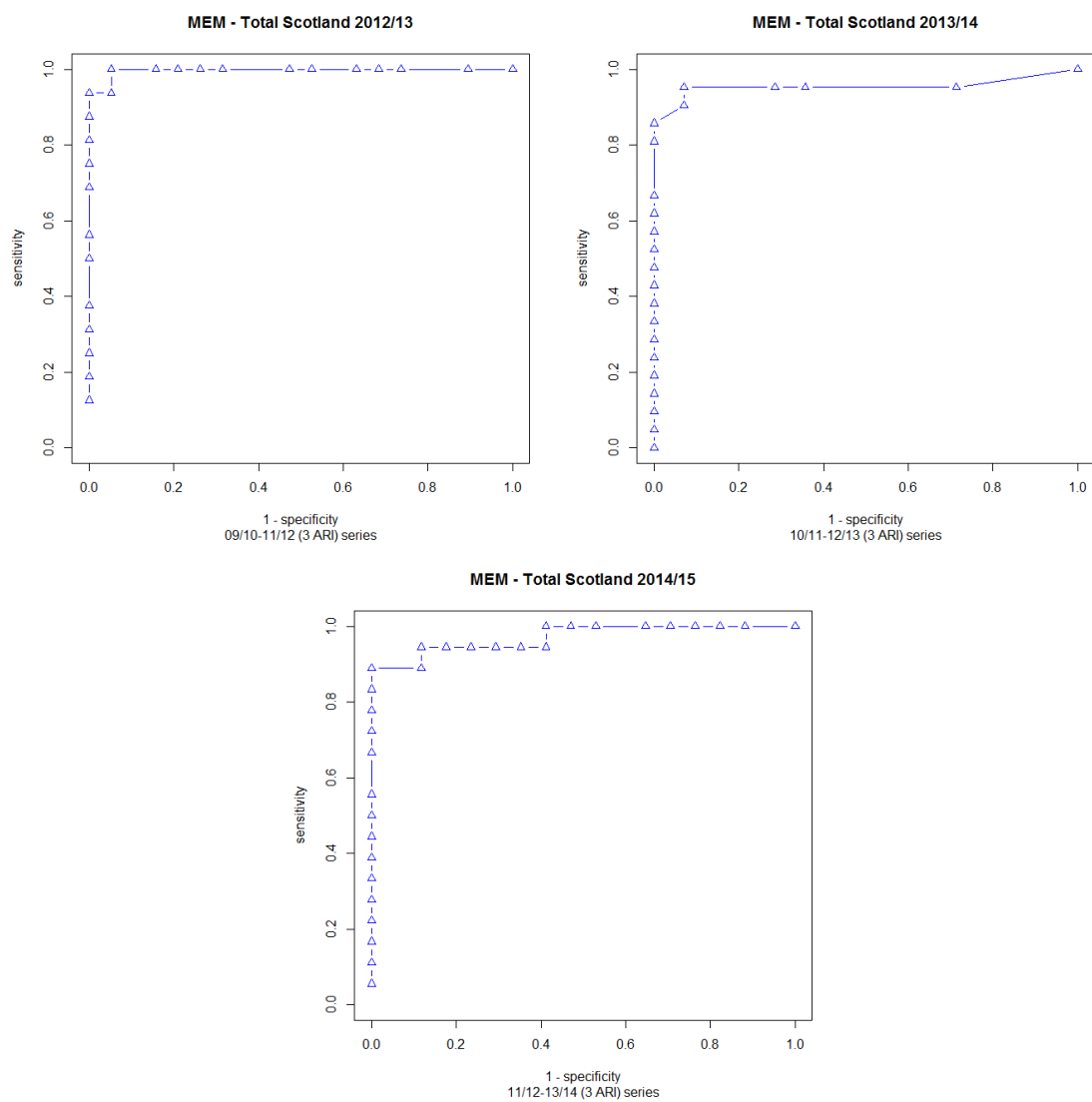


Figure 4.9: Sensitivity and specificity plots using MEM for ARI data with 3 years historical data to detect a pandemic in the following winter.

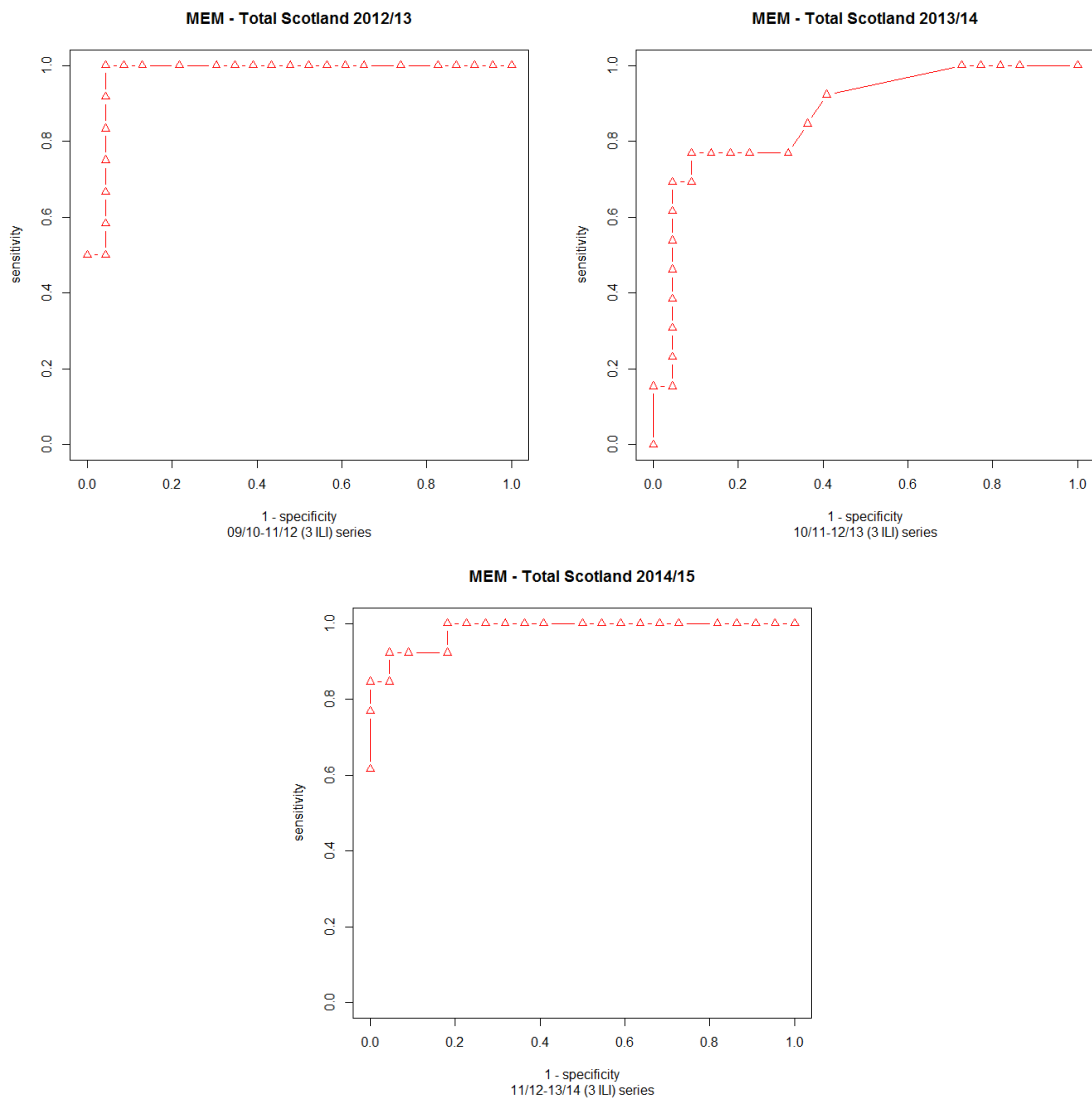


Figure 4.10: Sensitivity and specificity plots using MEM for ILI data with 3 years historical data to detect a pandemic in the following winter.

As we are looking get the best performance, we will now use the last chance for MEM, by using 4 years data to detect the fifth year, which means we have 2 MEMs for our data. Figure 4.11 shows the sensitivity and specificity plots for the ARI data, while figure 4.12 shows the output with the ILI data. Here we come to the last scenario where 4 years of historical data were used (the default situation requires 5 years and this is what we used in the first attempt and then tried to modify it). So far, a there is a similar result as higher specificity came with lower sensitivity, and vice versa, as the higher the sensitivity the lower the specificity, and away from the point of preference in this curve.

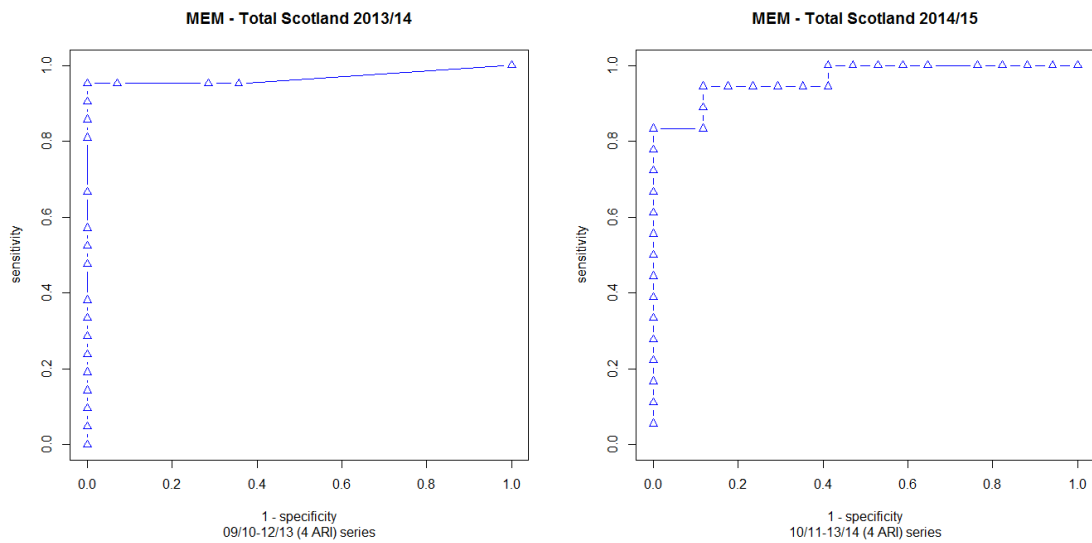


Figure 4.11: Sensitivity and specificity plots using MEM for ARI data with 4 years historical data to detect a pandemic in the following winter.

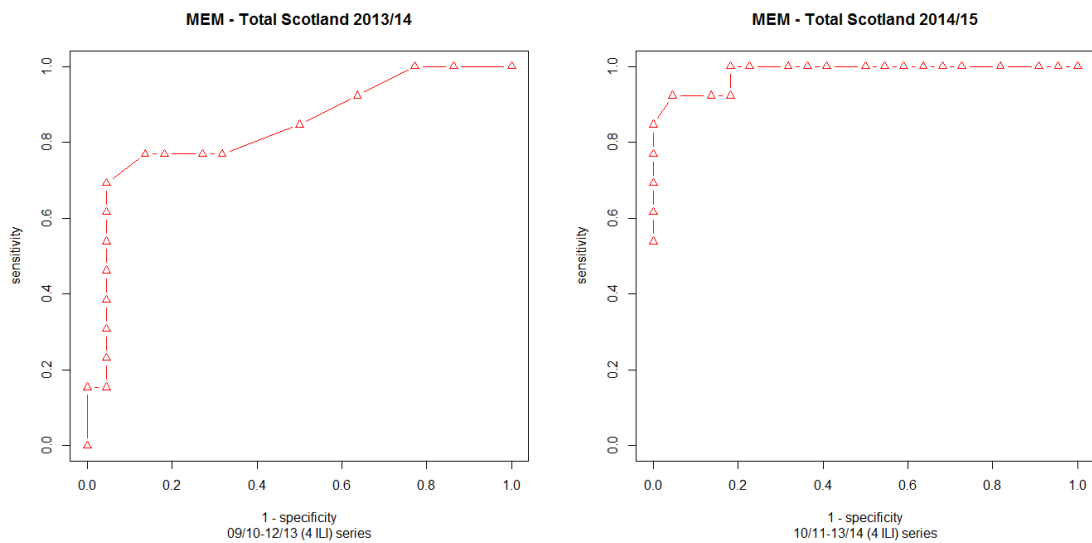


Figure 4.12: Sensitivity and specificity plots using MEM for ILI data with 4 years historical data to detect a pandemic in the following winter.

One example for monitoring influenza used by Vega et al. [162] included weekly aggregate numbers of influenza-like illness (ILI) and acute respiratory infections (ARI) for 19 European countries or regions since 1996. For each region or country, thresholds (pre-epidemic and post-epidemic) were defined, based on historical data from the start of surveillance in the country until the end of the 2008-2009 season. Then, those thresholds were compared in the modelled period in the next year (the 2009-2010 season) with all levels (pre-epidemic, epidemic and post-epidemic). Rates in the epidemic period were expected to be higher



than pre-epidemic period rates, while lower levels were expected again in the post-epidemic period. Cross-validation was used to evaluate the accuracy of this method, and each single season was taken from the historical data to be used as the objective season for each country. The MEM method was used. The performance measures can be standardized using these terms: sensitivity (across all countries providing data this was 71.8%), specificity (above 90% in all countries), positive predictive value (PPV) (defined by number of epidemic weeks above the threshold divided by the number of weeks above the threshold) and negative predictive value (NPV) (defined by number of non-epidemic weeks below the threshold divided by the number of weeks below the threshold). The PPV and NPV of the MEM exceeded 80% in all countries. The last measure is the median timeliness of epidemic detection, which was 1 week. Some of these measures are also used in the work in this thesis.

## 4.2 Applying the *WCR* method to the Scotland data

We first analyse the data in Scotland by breaking down the whole year to two periods, which are summer and winter:

1. Summer period from April to the end of September.
2. Winter period from October to the end of March.

Singh et al. in 2010 [142] defined the value of *WCR* as:

$$WCR = \frac{\text{ILI or ARI rates reported to all sentinel GPs in week } w}{\text{total ILI or ARI rates reported to all sentinel GPs in week } w-1}$$

Here, in our work, we defined  $WCR = \frac{Rate(w)}{Rate(w-1)}$  where  $Rate(w)$  is the number of cases in week  $w$  divided by the population size in week  $w$ . Because we cannot expect a big change in the population from week  $w$  to week  $w - 1$ , we will expect to get a very similar value of *WCR* as using the definition of *WCR* from Singh et al. [142].

First we look at the summer period.

### 4.2.1 Using the *WCR* method for summer

We present the *WCR* in the summer period (April - September) in 2010. Figure 4.13 shows *WCR* against weeks in this period. Overall the level of *WCR* show neither increasing not a decreasing trend over the period. Instead they fluctuate about a roughly constant level. Here the mean value of *WCR* should be around one. For health board BR there are very high weeks followed by very low weeks so the *WCR* oscillates, the average *WCR* is still around one. Some health boards show similar violation cases to the general pattern, like

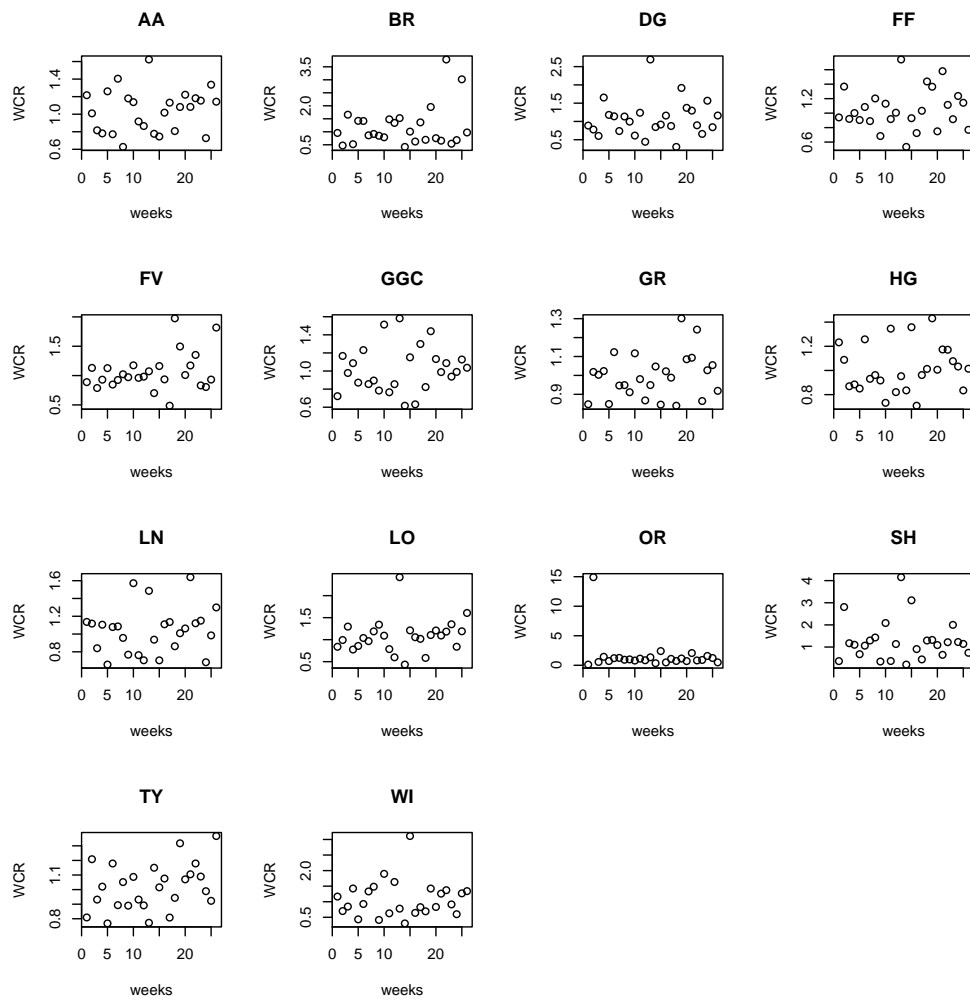


Figure 4.13: *WCR* against weeks in 2010 (April to September).

those cases in health boards DG, FF, FV, GR, LN, LO, SH and WI. Another case to mention is health board OR. It shows stability in *WCR* during most weeks, except for an unusual case in week 2 with a very high level and  $WCR = 15$ . Table 4.3 gives the mean of *WCR* for each HB in summer 2010. The mean for OR is 1.55, for BR it is 1.18, and for SH it is 1.28, but all others are around 1.

The second year 2011 in the same period (April-September), presented in figure 4.14, shows *WCR* less than 1 in most cases. This means that mean cases reported in week  $w$  are less than cases reported in week  $w - 1$ . As for the previous year, some health boards showed different patterns, like week 5 for all health boards (apart from SH), and week 23 for FF. Health board SH shows very unusual cases (25 and 15) in weeks 14 and 24 respectively. Table 4.4 gives the mean of *WCR* for each HB in summer 2011. For all except SH those are close to 1.

Now, we want to see if there are any patterns in the third year, 2012, for the same period

No.	HB	Mean of <i>WCR</i>	No.	HB	Mean of <i>WCR</i>
1	AA	1.0393	8	HG	1.0178
2	BR	<b>1.1803</b>	9	LN	1.0372
3	DG	1.0737	10	LO	1.0797
4	FF	1.0509	11	OR	<b>1.5512</b>
5	FV	1.0577	12	SH	<b>1.2824</b>
6	GGC	1.0219	13	TY	1.0179
7	GR	0.9967	14	WI	1.0864

Table 4.3: Means of *WCR* in 2010 (April to September) for each health board; higher values are shown in bold.

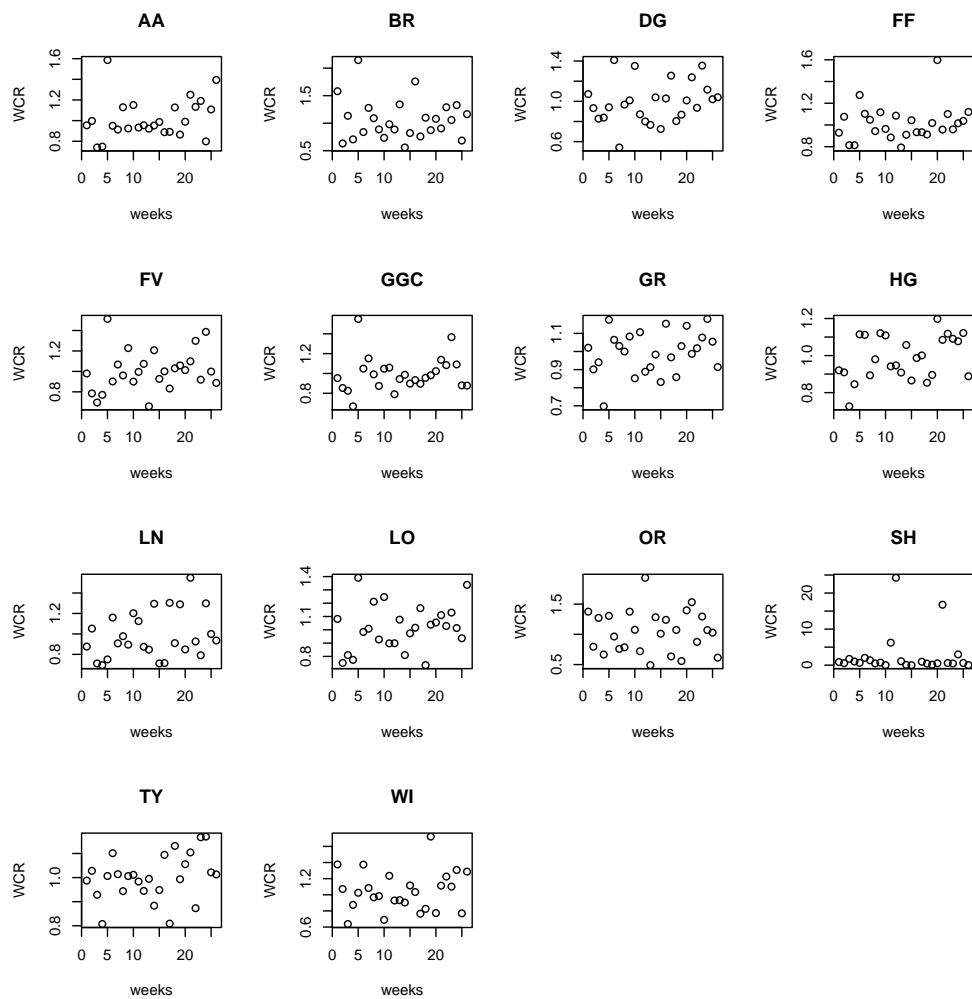


Figure 4.14: *WCR* against weeks in 2011 (April to September).

No.	HB	Mean of <i>WCR</i>	No.	HB	Mean of <i>WCR</i>
1	AA	1.0184	8	HG	0.9910
2	BR	1.0624	9	LN	0.9864
3	DG	0.9911	10	LO	1.0155
4	FF	1.0148	11	OR	1.0435
5	FV	1.0072	12	SH	<b>3</b>
6	GGC	0.9949	13	TY	1.0009
7	GR	0.9950	14	WI	1.0434

Table 4.4: Means of *WCR* in 2011 (April to September) for each health board; higher values are shown in bold.

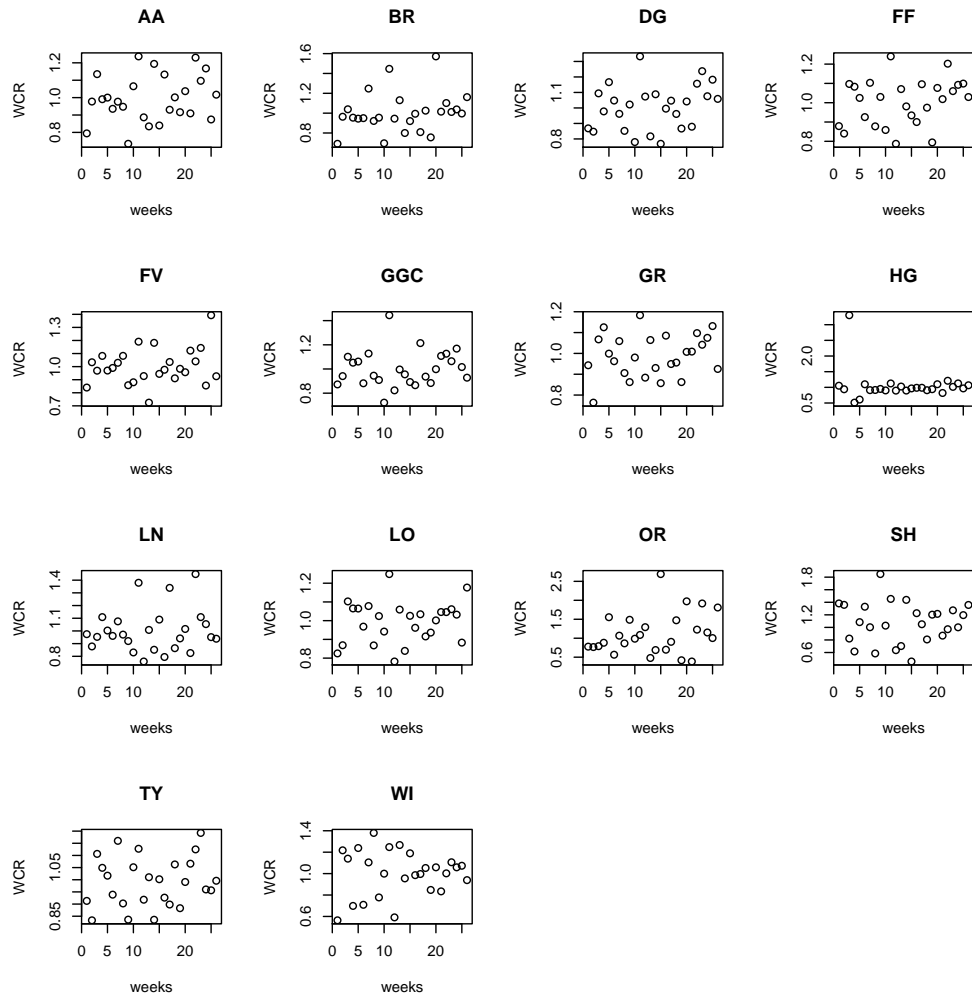
No.	HB	Mean of <i>WCR</i>	No.	HB	Mean of <i>WCR</i>
1	AA	0.9947	8	HG	1.0473
2	BR	1.0033	9	LN	1.0018
3	DG	1.0076	10	LO	0.9945
4	FF	1.0029	11	OR	<b>1.1132</b>
5	FV	1.0018	12	SH	1.0737
6	GGC	1.0014	13	TY	0.9907
7	GR	0.9895	14	WI	1.0013

Table 4.5: Means of *WCR* in 2012 (April to September) for each health board; higher values are shown in bold.

(April to September). Health boards (apart from HG) show increases and decreases in *WCR* within the range 0.8 to 1.4. The health board of HG (as mentioned in the previous section) shows a very unusual increase in week 5, with a *WCR* of nearly 3. These can be seen in figure 4.15. Table 4.5 gives the mean of *WCR* for each HB in summer 2012. The highest value 1.11 is for OR, while all others are close to 1.

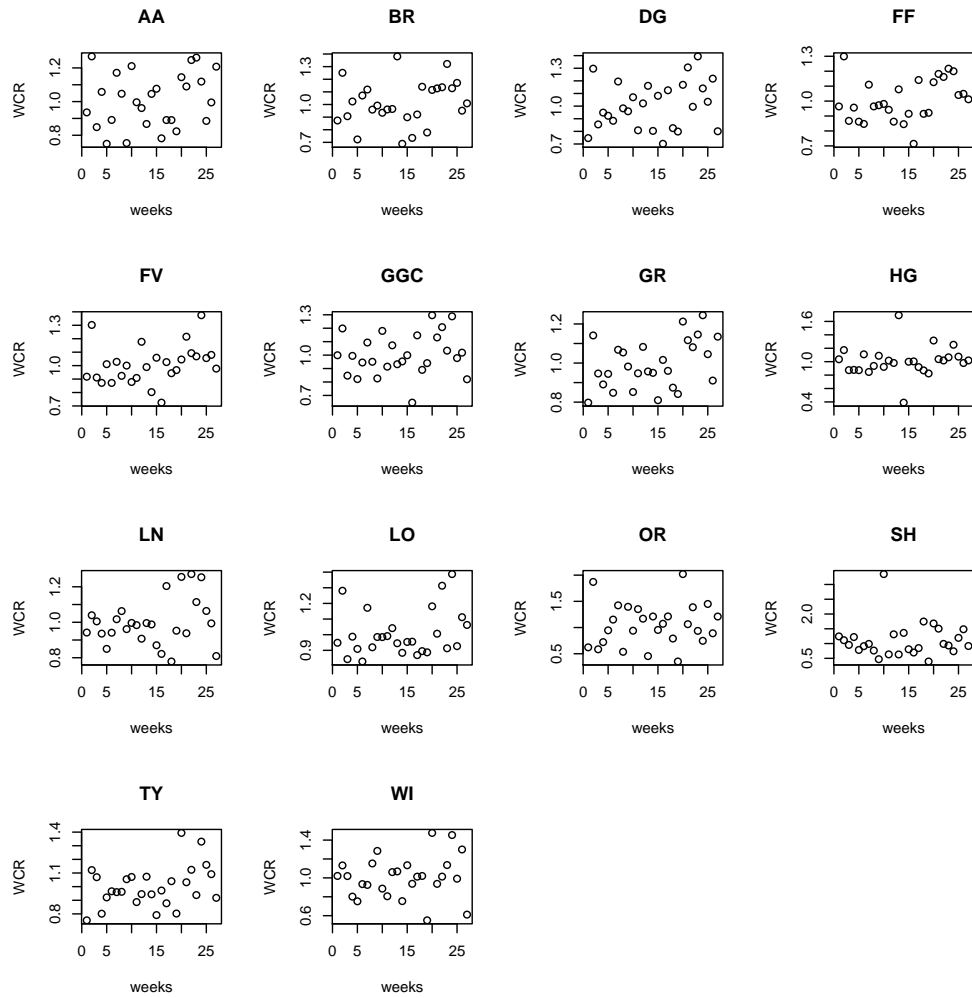
Now we look at *WCR* in the period of April to September 2013. Figure 4.16 shows stability in a fluctuation in the ups and downs for the different health boards, in the range from 0.6 to 1.4. As usual, some health boards present differently. Some reported very high *WCR* values, which happened with OR, SH and HG in some weeks. In general, all health boards showed fluctuations which did not have a clear pattern. Table 4.6 gives the mean of *WCR* for each HB in summer 2013. The highest value is about 1.1 for SH, while the others are close to 1.

When we come to the last year (2014), it is different from previous years, as we can see a clearly defined pattern in figure 4.17 for each health board, and in most cases 0.4 to 1.4 is the general range of *WCR*. Health boards AA, BR, FF, GGC, FV and TY show a less specific pattern. We must point out that there are again some anomalous points, as is evident in OR, SH and WI. Table 4.7 gives the mean of *WCR* for each HB in 2014. The

Figure 4.15: *WCR* against weeks in 2012 (April to September).

No.	HB	Mean of <i>WCR</i>	No.	HB	Mean of <i>WCR</i>
1	AA	1.0080	8	HG	1.0073
2	BR	1.0105	9	LN	0.9983
3	DG	1.0091	10	LO	1.0096
4	FF	1.0054	11	OR	1.0543
5	FV	1.0083	12	SH	<b>1.0989</b>
6	GGC	1.0046	13	TY	0.999
7	GR	0.9946	14	WI	1.0046

Table 4.6: Means of *WCR* in 2013 (April to September) for each health board; higher values are shown in bold.

Figure 4.16: *WCR* against weeks in 2013 (April to September).

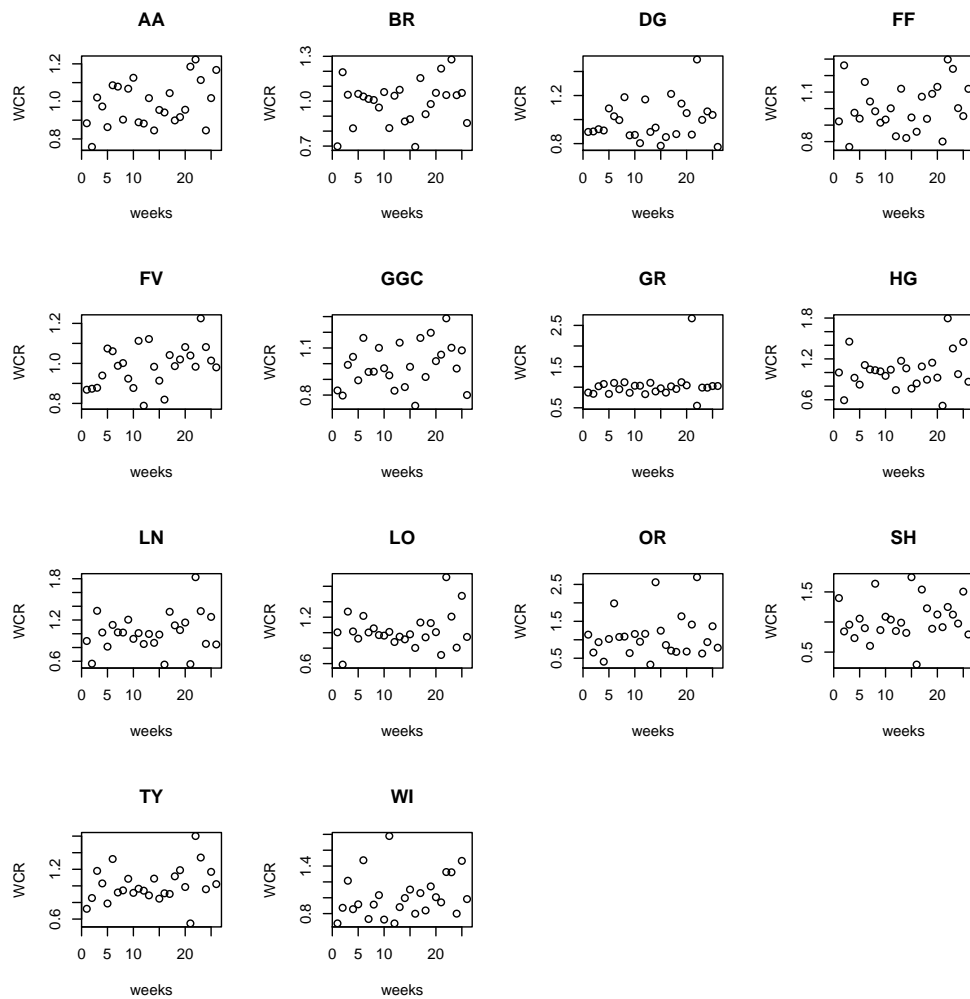


Figure 4.17: *WCR* against weeks in 2014 (April to September).

highest value of 1.10 is for OR, while all others are close to 1.

We can conclude with the important and clear note that the mean *WCR* in all summers is generally close to 1 (sometimes higher and sometimes lower than 1), which means the prevailing situation in summer is stable numbers of GP consultation. Exceptions are generally for the smaller health boards.

#### 4.2.2 Using the *WCR* method for winter

Now we investigate what happens in the cold seasons during the year when many people complain of influenza (as seen in the analysis of the rise in GP consultations in November, December and January) in section 2.4. We present *WCR* in the period October - March in 2010/2011. *WCR* is shown against weeks in this period in figure 4.18. Most health boards showed wide spread in *WCR* values with a range of 0.6 to 1.7. There is limited

No.	HB	Mean of <i>WCR</i>	No.	HB	Mean of <i>WCR</i>
1	AA	0.9868	8	HG	1.0220
2	BR	0.9939	9	LN	1.0193
3	DG	0.9862	10	LO	1.0235
4	FF	1.0051	11	OR	<b>1.1049</b>
5	FV	0.9876	12	SH	<b>1.0444</b>
6	GGC	0.9898	13	TY	1.0093
7	GR	1.0322	14	WI	<b>1.1333</b>

Table 4.7: Means of *WCR* in 2014 (April to September) for each health board; higher values are shown in bold.

No.	HB	Mean of <i>WCR</i>	No.	HB	Mean of <i>WCR</i>
1	AA	1.0107	8	HG	1.0306
2	BR	<b>1.0748</b>	9	LN	1.0166
3	DG	<b>1.0529</b>	10	LO	1.0189
4	FF	1.0236	11	OR	<b>1.1902</b>
5	FV	1.0195	12	SH	<b>1.3341</b>
6	GGC	1.0122	13	TY	1.0108
7	GR	1.0377	14	WI	1.0277

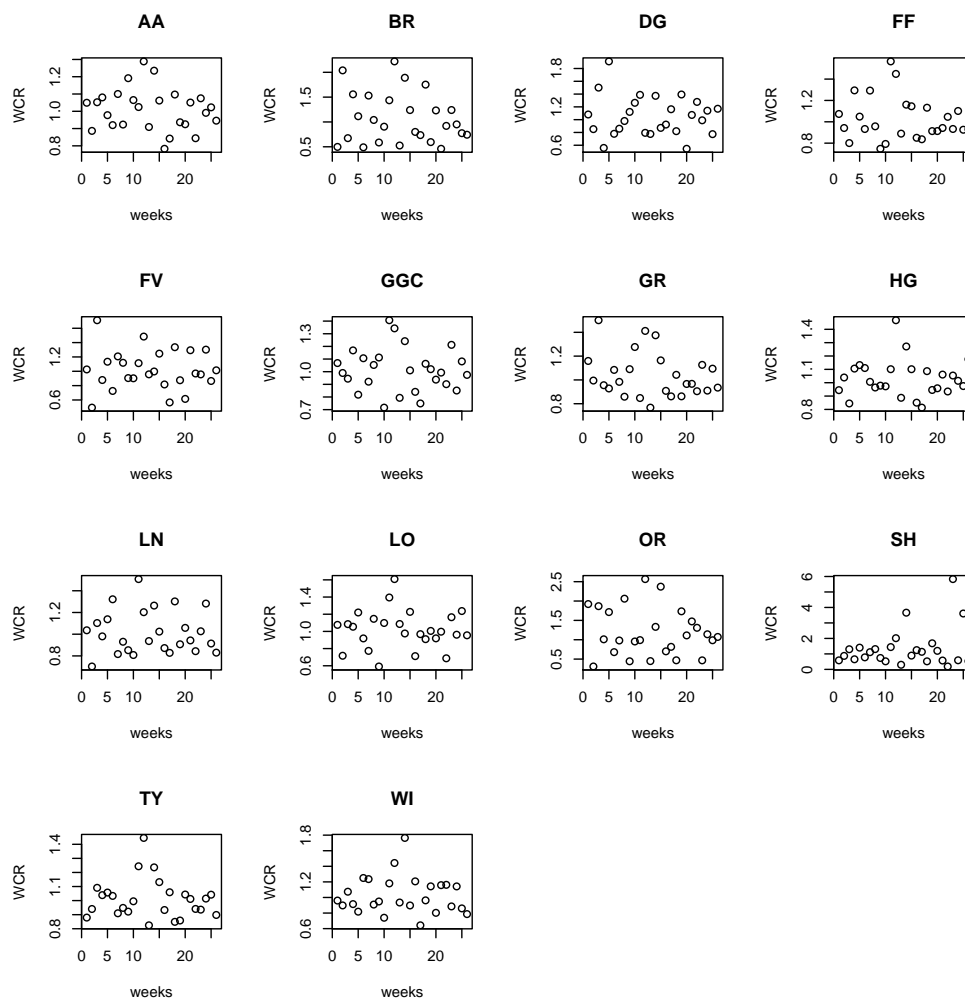
Table 4.8: Means of *WCR* in 2010/11 (October to March) for each health board; higher values are shown in bold.

pattern, with clear oscillation in *WCR*. In some health boards there are some anomalous values, as in DG, FV, TY and WI. Health boards OR and SH again has a completely different value of *WCR*, an unusual ratio over 2. Table 4.8 gives the mean of *WCR* for each HB in winter 2010/11. The mean for BR is 1.07, for DG it is 1.05, OR it is 1.19 and for SH it is 1.33, but all others are around 1 and we can see all of them are at least a little higher than 1.

Cold seasons in 2011/12 showed a similar pattern to the previous year, with a range of *WCR* (0.7 to 1.4) in the majority of health boards and they showed clear pattern in the high values from 0.9, as presented in figure 4.19, which shows some oddities about the general pattern in the health boards of BR, DG, FV, LO, OR and WI. SH is a health board which does not conform to the general pattern in other health boards. It showed a lot of zeroes and some unusual high points in *WCR*, up to a value of 30. The means of each health board are presented in table 4.9 and the highest value of 3.28 is for SH, while all others are fairly close to 1.

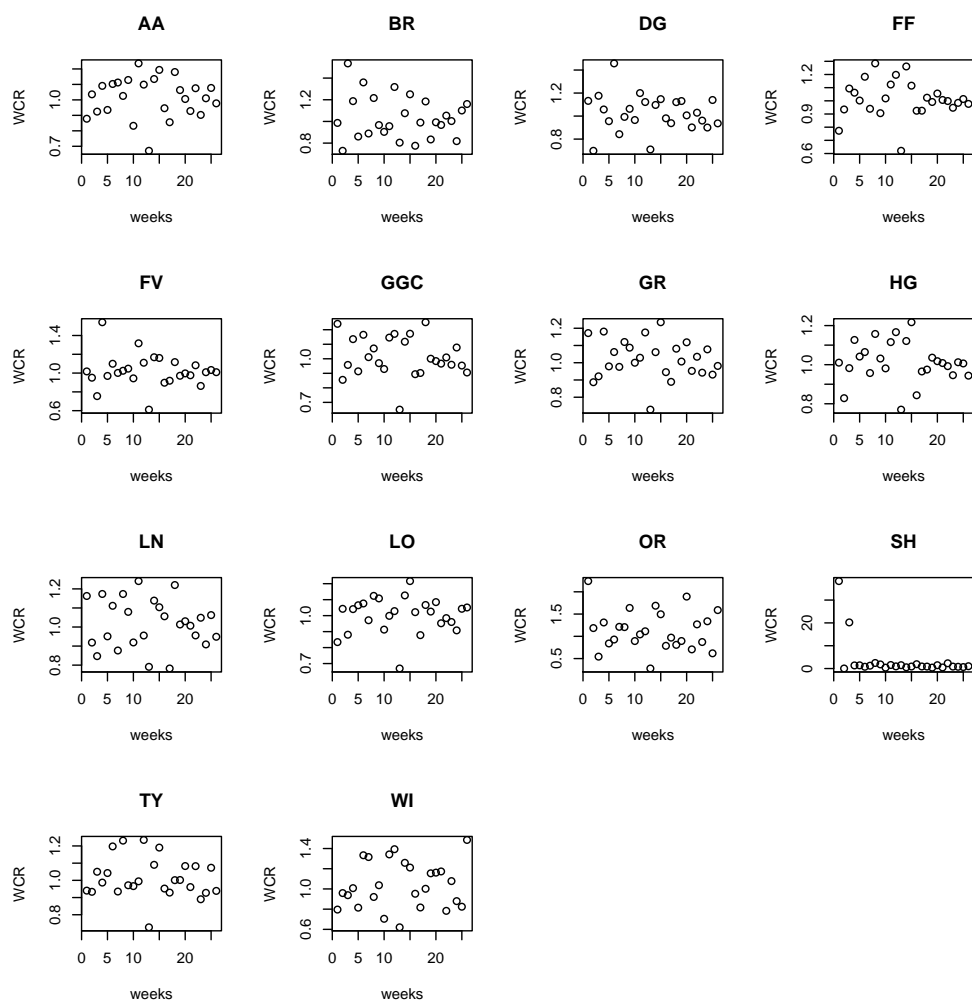
It appears that in the year 2012/13 there is consensus over most health boards, as in general they have *WCR* in the range 1.004 to 1.048 (apart from OR). Figure 4.20 presents *WCR* for 2012/13. Table 4.10 shows the means of *WCR* for the different health boards



Figure 4.18: *WCR* against weeks in 2010/11 (October to March).

No.	HB	Mean of <i>WCR</i>	No.	HB	Mean of <i>WCR</i>
1	AA	1.0162	8	HG	1.0123
2	BR	1.0353	9	LN	1.0183
3	DG	1.0256	10	LO	1.0027
4	FF	1.0140	11	OR	<b>1.1295</b>
5	FV	1.0228	12	SH	<b>3.27631</b>
6	GGC	1.0158	13	TY	1.0128
7	GR	1.0222	14	WI	1.0371

Table 4.9: Means of *WCR* in 2011/12 (October to March) for each health board; higher values are shown in bold.

Figure 4.19: *WCR* against weeks in 2011/12 (October to March).

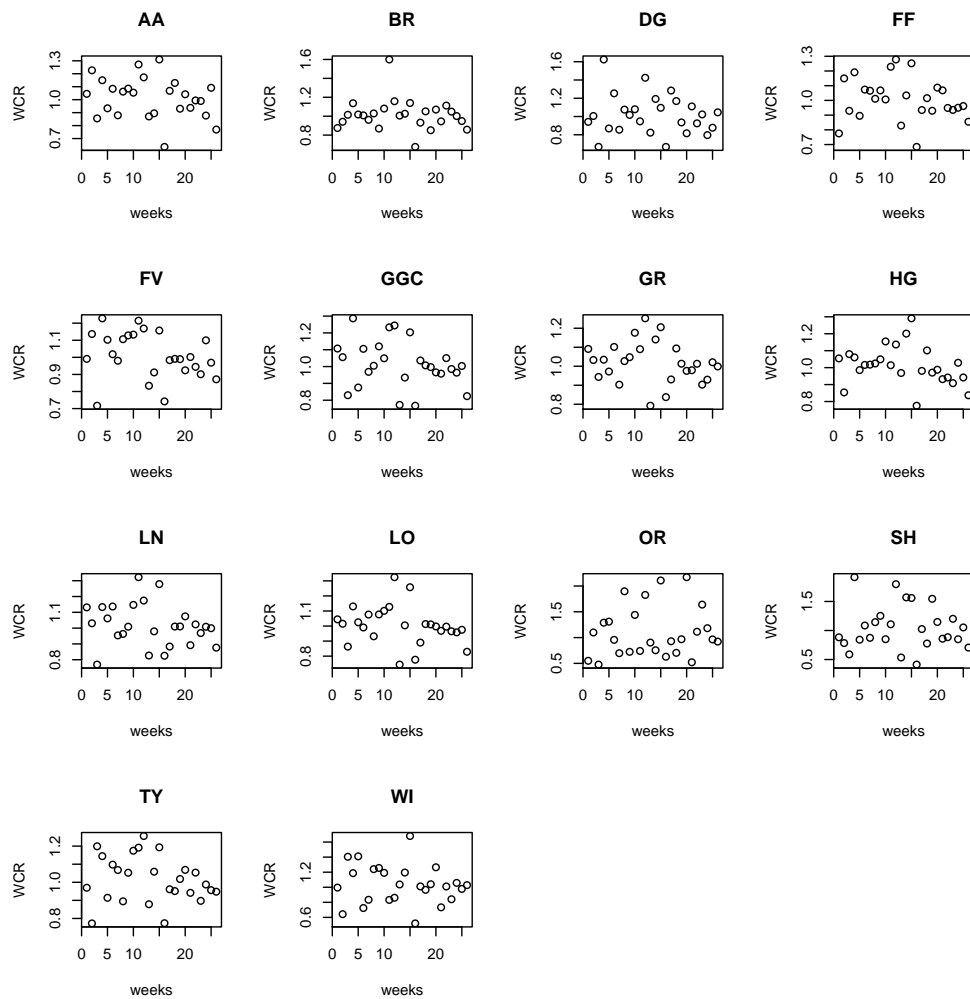


Figure 4.20:  $WCR$  against weeks in 2012/13 (October to March).

and the highest value of 1.1 is for OR, while all others are quite close to 1.

For this period from October 2013 to March 2015, Figure 4.21 shows  $WCR$  values which are spread in the range of 0.7 to 1.2. If  $WCR > 1$ , this means there are more cases in week  $w$  compared with week  $w - 1$ . There are some health boards which have some points which can be classed as outside the general pattern in their  $WCR$ , like OR, SH and WI. Health board AA does not show any pattern in this year. The means of  $WCR$  in this period are presented in table 4.11. The mean for OR is 1.05, for SH it is 1.05 and for WI it is 1.16, but all others are around 1 and we can see that all of them are higher than 1.

We can conclude with an important and clear note that the mean  $WCR$  in all winters is higher than 1 in all health boards, which means that the prevailing situation in winter is increasing numbers of consultations. Higher mean values tend to be for smaller health boards.

No.	HB	Mean of <i>WCR</i>	No.	HB	Mean of <i>WCR</i>
1	AA	1.0183	8	HG	1.0120
2	BR	1.0138	9	LN	1.0189
3	DG	1.0203	10	LO	1.0036
4	FF	1.0060	11	OR	<b>1.0969</b>
5	FV	1.0094	12	SH	<b>1.0481</b>
6	GGC	1.0134	13	TY	1.0162
7	GR	1.0194	14	WI	<b>1.0363</b>

Table 4.10: Means of *WCR* in 2012/13 (October to March) for each health board; higher values are shown in bold.

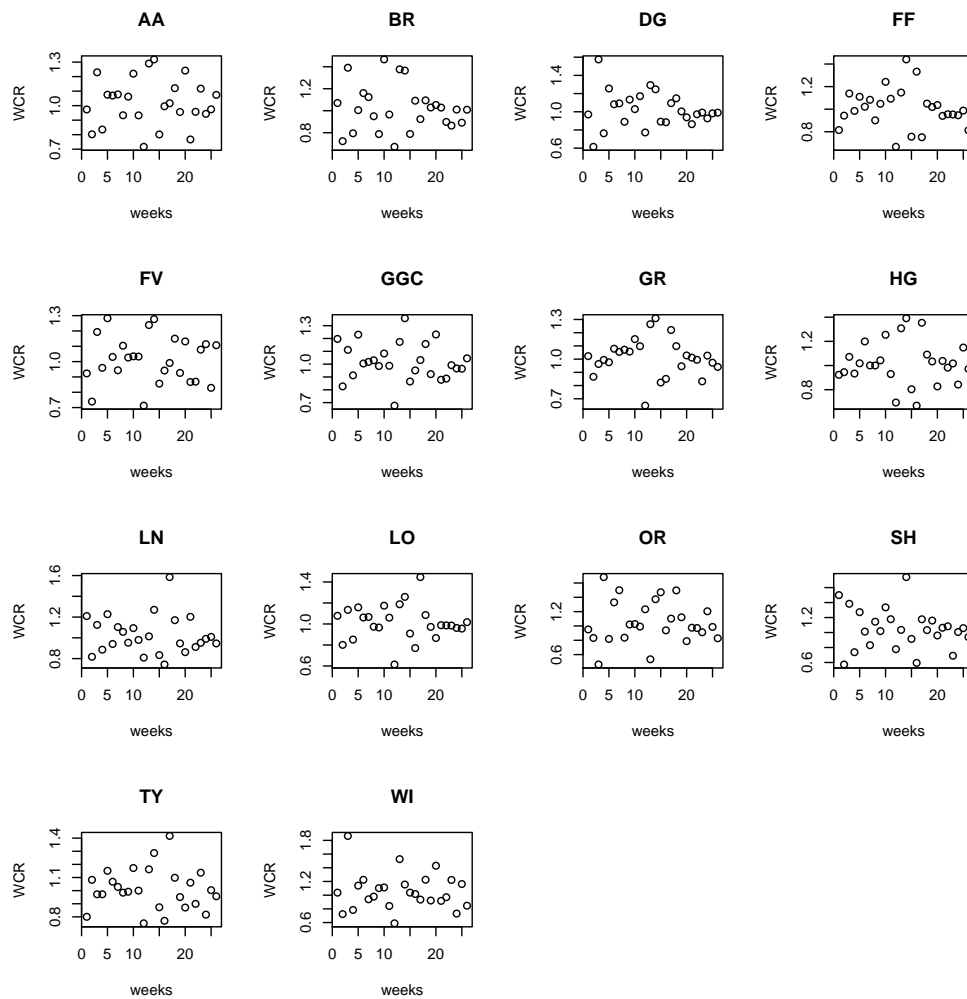


Figure 4.21: *WCR* against weeks in 2013/14 (October to March).

No.	HB	Mean of $WCR$	No.	HB	Mean of $WCR$
1	AA	1.0194	8	HG	1.0188
2	BR	1.0208	9	LN	1.0247
3	DG	1.0203	10	LO	1.0127
4	FF	1.0225	11	OR	<b>1.0535</b>
5	FV	1.0067	12	SH	<b>1.0469</b>
6	GGC	1.0138	13	TY	1.0107
7	GR	1.0116	14	WI	<b>1.1608</b>

Table 4.11: Means of  $WCR$  in 2013/14 (October to March) for each health board; higher values are shown in bold.

### 4.2.3 The joint probability distribution of $(WCR, N_{HB})$

Above we examined the  $WCR$  rates for each health board in each year, in summer and winter. Now we consider probability distributions for these data.

Figure 4.22 represents the joint frequency distribution of  $(WCR, N_{HB})$  over all health boards and in all years over summer (defined from April to September) and winter (defined from October to March).  $N_{HB.Increase}$  in the plot now represents  $N_{HB}$ . There is a pattern, that as  $N_{HB}$  increases,  $WCR$  also increases. We also present summer and winter in separate figures, figure 4.23 and 4.24 to compare the different patterns.

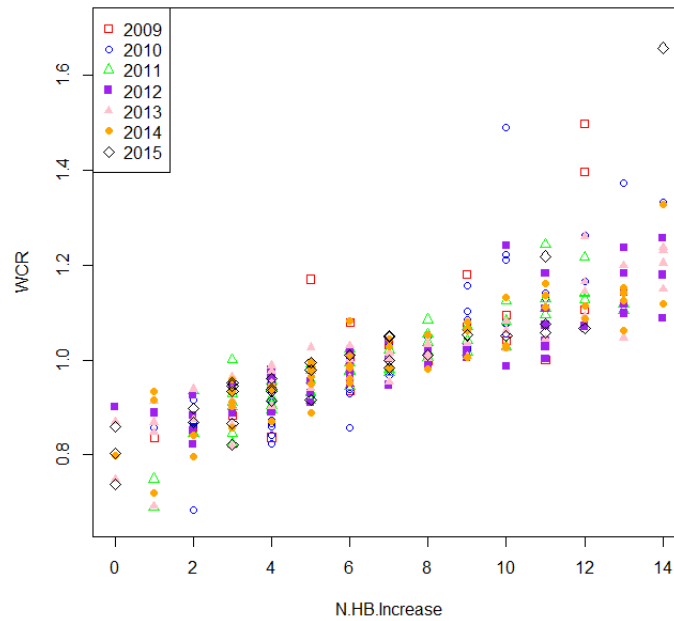


Figure 4.22: The joint frequency distribution of  $(WCR, N_{HB})$  over all health boards in all years, for summer and winter combined. Individual points represent different weeks within any one year.

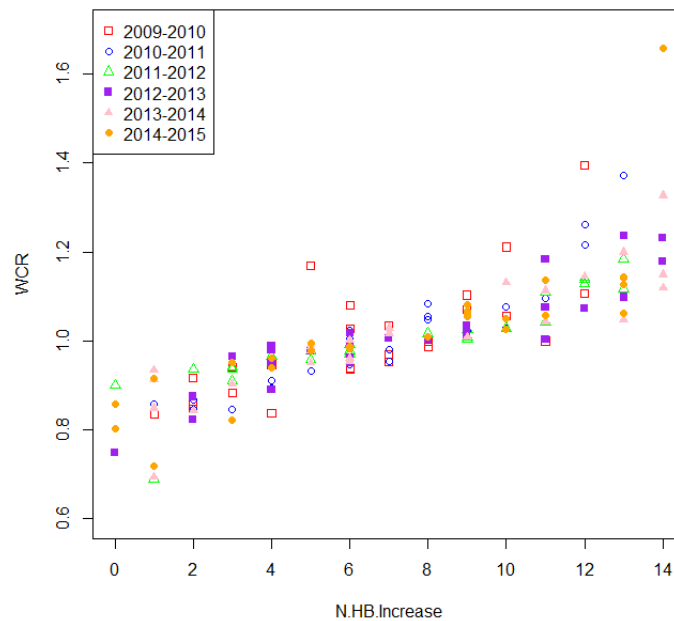


Figure 4.23: The joint frequency distribution of  $(WCR, N_{HB})$  over all health boards in all years, for winter. Individual points represent different weeks within any one year.

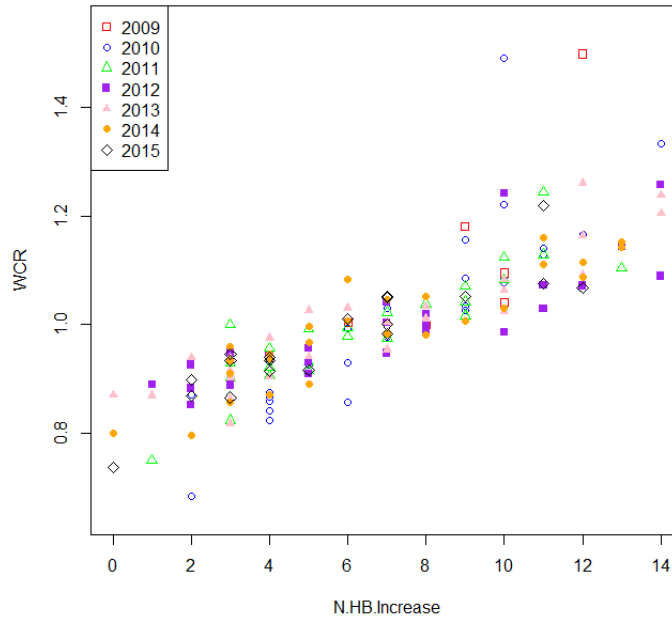


Figure 4.24: The joint frequency distribution of  $(WCR, N_{HB})$  over all health boards in all years, for summer. Individual points represent different weeks within any one year.

When we compare figures 4.23 and 4.24, we see the same increasing trend (as  $N_{HB}$  increases, so does  $WCR$ ) but the joint distribution of  $(WCR, N_{HB})$  in summer is lower down than the joint distribution in winter, meaning that the rates reported increase in winter, as expected. We will test this below.

### Univariate and bivariate testing

In this section, we want to determine whether the different samples (from summer and winter) came from the same distribution or not. There are different tests which can be used to see whether the data in the summer period (defined from April to September) follows the same distribution as the data in the winter period (defined from October to March) or not.

In the graph of the joint distribution,  $N_{HB}$  is on the horizontal axis and  $WCR$  is on the vertical axis, and we want to compare the distributions in three situations:

1.  $WCR$  in summer against  $WCR$  in winter.
2.  $N_{HB}$  in summer against  $N_{HB}$  in winter.
3. Two-dimensional distributions for  $WCR$  and  $N_{HB}$  in summer against winter.

The first and second of these can be tested by using the Kolmogorov-Smirnov test [177]

to see if they follow the same distribution, and the Anderson-Darling test can be used to check if the *WCR* are normally distributed. As  $N_{HB}$  takes numbers 0 to 14, it is clear that it cannot be normally distributed. Firstly, we check the hypotheses for the test of the *WCR* data over the total period in summer compared to winter:

$H_0$ : *WCR* in summer and winter follow a single distribution and  $H_1$ : *WCR* in summer and winter do not follow the same distribution. The  $p$ -value for the Kolmogorov-Smirnov test [177] is 0.09, so the null hypothesis cannot be rejected since the  $p$ -value is higher than 0.05, the default value of the level of significance. According to this test, the difference between the two samples is not significant enough to say that they have different distributions. We can break down the period for each year individually, and we get the same conclusion, as in table 4.12 below, which presents  $p$ -values for the Kolmogorov-Smirnov tests of each summer period against the corresponding winter period.

Test no.	Year	$p$ -value
1	Summer 2010 and winter 2009-2010	0.7327
2	Summer 2011 and winter 2010-2011	0.9966
3	Summer 2012 and winter 2011-2012	0.3071
4	Summer 2013 and winter 2012-2013	0.1531
5	Summer 2014 and winter 2013-2014	0.9966

Table 4.12: The Kolmogorov-Smirnov test  $p$ -values for tests of *WCR* in each summer period against the corresponding winter period.

We now consider whether the distribution of *WCR* could be normal. We check if *WCR* in summer and winter is normal or not by using the Anderson-Darling test: as before, checking the total period in summer and winter. The hypotheses tested are  $H_0$ : *WCR* follows a normal distribution and  $H_1$ : *WCR* does not follow a normal distribution. In winter, the  $p$ -value for the Anderson-Darling test is 0.0001, while the  $p$ -value in summer is 0.009, so we strongly reject the null hypothesis since each  $p$ -value is much less than  $\alpha=0.05$ , the significance level. According to this powerful test, the *WCR* for all years combined in summer and winter are both not normally distributed. As before, we can break down the period, for each year individually, as in table 4.13 below, which presents  $p$ -values for the Anderson-Darling tests of each summer and winter period. Now every test is non-significant, so we cannot reject normality (at significance level 0.05) of *WCR* in any season or year.



No.	Year	$p$ -value	Year	$p$ -value
1	Summer 2010	0.36	winter 2009-2010	0.15
2	Summer 2011	0.66	winter 2010-2011	0.06
3	Summer 2012	0.06	winter 2011-2012	0.11
4	Summer 2013	0.12	winter 2012-2013	0.09
5	Summer 2014	0.93	winter 2013-2014	0.46

Table 4.13: The Anderson-Darling  $p$ -values for the normality test of *WCR* in each summer period and winter period.

We now come to step 2, to do the Kolmogorov-Smirnov test for the  $N_{HB}$  in the total period in summer and winter. The hypotheses are  $H_0$ :  $N_{HB}$  in summer and winter follow the same distribution, and  $H_1$ :  $N_{HB}$  in summer and winter do not follow the same distribution. The  $p$ -value for the Kolmogorov-Smirnov test is 0.59, so we cannot reject the null hypothesis since the  $p$ -value is higher than  $\alpha = 0.05$ . According to this test, the difference between two samples is not significant enough to say that they have different distributions.

As we did with *WCR*, we can break down the period for each year individually, and we get the same conclusion, as in table 4.14 below, which presents  $p$ -values for the Kolmogorov-Smirnov tests of each summer period against the corresponding winter period.

Test no.	Year	$p$ -value
1	Summer 2010 and winter 2009-2010	0.72
2	Summer 2011 and winter 2010-2011	1
3	Summer 2012 and winter 2011-2012	0.72
4	Summer 2013 and winter 2012-2013	0.79
5	Summer 2014 and winter 2013-2014	0.99

Table 4.14: The Kolmogorov-Smirnov test  $p$ -values for tests of  $N_{HB}$  in each summer period against the corresponding winter period.

Finally, we come to step 3, to compare the two-dimensional distributions for *WCR* and  $N_{HB}$  in summer and winter. Using the Crossmatch test [4, 69], we test  $H_0$ : *WCR* and  $N_{HB}$  in summer and winter follow the same joint distribution, against  $H_1$ : *WCR* and  $N_{HB}$  in summer and winter do not follow the same joint distribution. The  $p$ -value for this test is 0.75, higher than 0.05, so we cannot reject the null hypothesis and we conclude that the joint distributions for *WCR* and  $N_{HB}$  in summer and winter are also not different. This means that we can combine the data over summer and winter.

We check here the whole period (not each year individually) as we expect to get the same result and because the joint distribution in any one year contains less data.

The above work applies to the ARI data.

When the ILI data was investigated, interesting results were obtained. The results did not follow a similar distribution pattern in summer and winter as was observed with the ARI data. Testing *WCR* in summer against *WCR* in winter gave a  $p$ -value of  $0.02 < 0.05$ , so the test of the difference between the two samples is significant enough to conclude that they have different distributions. Testing  $N_{HB}$  in summer against  $N_{HB}$  in winter gave a  $p$ -value of  $0.3 > 0.05$  from the Kolmogorov-Smirnov test, so the difference between the two samples is not significant enough to say that they have different distributions, while testing the two-dimensional distributions for *WCR* and  $N_{HB}$  in summer against winter using the Crossmatch test gave a  $p$ -value  $< 0.0001 < 0.05$ . We conclude that the joint distributions for *WCR* and  $N_{HB}$  in summer and winter are different for the ILI data, as are the distributions for *WCR* alone, although for  $N_{HB}$  they could be assumed to be the same. So for the ILI data on the basis of these tests it would not make sense to combine the data for summer and winter but we did in fact do this later because we wished to use the ILI data which is more specific to influenza, and as presented in section 2.4, there are lots of zeroes in the numbers of ILI consultations in the summer period compared to the winter period.

We did not check the normality situation here as we know from chapter 2, section 2.4, that ILI consultations in Scotland included lots of zeroes (especially in the summer and also in small HBs). Later in chapters 5 and 6, we do assume normality of the *WCR* data, but this is based on millions of simulations.

Singh et al. in 2010 [142] analysed Scottish ILI data from the 2001-02 season through to the 2008-09 season. The ILI cases are normally reported weekly over a period of 33 weeks (from the first week of October to the third week of May) in different age and sex classes, by sentinel general practices (GPs) across Scotland. Scottish Enhanced Respiratory Virus Infection Surveillance (SERVIS) sentinel GPs cover all health boards in Scotland. The Singh et al. paper does not consider the health board of the Western Isles as it did not then take part in the SERVIS network. The population size differed widely between health boards, from 20,000 to 1,360,000. The percentage of the population registered with sentinel GPs range from 2.7% to 4.9% of the Scottish population of around 5.15 million. This system was designed for general use as a national surveillance scheme, not for each health board separately. Weekly ILI cases were aggregated at the HB level in the data analysis. Historical data in the seasonal influenza for 2001-2002 until 2006-2007 were used in estimating the background pattern of the seasonal ILI cases. In season 2007-08, the ILI data from 23 sentinel GPs recorded only 93 cases while the other seasons had over 300 cases recorded, so this season was excluded from the weekly case ratio (*WCR*) method. The 2008-09 data included real pandemic influenza A (H1N1), so

it was used for performance testing with the detection algorithm and not as part of the historical data. Singh et al. compared the *WCR* method with the well-known CUSUM method and the threshold method. Different values were used for the case reporting rate ( $\alpha$ ) in simulated data and pre-set specificity (Sp) values. The specificity was used to set the threshold probability for the *WCR* method. Results for this comparison were based on sensitivity (Sen), and median detection time (MDT) with three case reporting rates ( $\alpha = 0.5\%$ ,  $1\%$  and  $5\%$ ) and two levels of specificity (Sp) of  $95\%$  and  $99\%$ . In general, the *WCR* method showed satisfactory performance in sensitivity, as it was almost  $100\%$  sensitive ( $98\%$ , with a very low case reporting rate of  $0.5\%$ , was the lowest sensitivity value). The worst performance in terms of sensitivity was the CUSUM method with  $77\%$  to  $97\%$ , while the threshold method showed satisfactory performance of  $100\%$ . Regarding MDT, *WCR* was the best as it achieved only 3-5 weeks, compared with 4-6 weeks in the CUSUM and threshold methods. The ILI rate threshold method used a threshold of 24 and 34 cases per 100,000 population for  $95\%$  and  $99\%$  specificity respectively. The *WCR* method and also the CUSUM method outperformed the ILI threshold method in influenza seasons in terms of time to detection, though CUSUM was less sensitive than the other methods. Rapid detection is very important because the aim from any method is buying time before a pandemic occurs. The *WCR* method outperformed the other two methods by quick exposure in most model runs at Sp of  $99\%$  and  $\alpha$  of  $0.5\%$ . In the first six weeks, it detected a pandemic in more than  $50\%$  of all runs. The CUSUM and the threshold methods detected pandemics, respectively, in  $<25\%$  and  $<35\%$  of total runs in this case. In general, under the same conditions *WCR* was superior to other methods in terms of rapid detection.

In this chapter, we applied the MEM method (section 3.3) and the *WCR* method (section 3.6) working with the real data, the ARI and the ILI data from all health boards in Scotland.

In section 4.1, we can see that MEM with ARI data outperforms the ILI data in terms of sensitivity and specificity. We cannot determine a certain number of years to use from the historical data to say that it will give the best performance. This is because sensitivity and specificity of MEM depend on a threshold which is directly related to data from past years.

Using the *WCR* method in section 4.2, we do not see much difference between the data of ARI and ILI when we applied the first part of the *WCR* method and did not apply the second part using  $N_{HB}$ . We did not neglect this inadvertently but tried to investigate whether we could use winter and summer data together in section 4.2.3. We analyzed the

values of *WCR* in summer to compare them with corresponding values in winter and check the distributions, and similarly for  $N_{HB}$ . We also linked the *WCR* with  $N_{HB}$  in summer and winter and checked whether they follow the same joint distribution in summer and winter. We concluded that the joint distributions for *WCR* and  $N_{HB}$  in summer and winter are not different for the ARI data though they are different for the ILI data (this may be because we have lots of zeroes in numbers of ILI consultations; especially in the summer season).

The problem of the Singh et al. *WCR* system for routine use is that it requires a simulation each time it is used to decide whether or not an observation of  $(WCR, N_{HB})$  is abnormal.

The challenge now is to find a simpler system defining values of *WCR* and  $N_{HB}$  that we would consider unusual (indicating an increasing rate), and this is what we will discuss in the next chapters 5 to 9. This is an extension to the work of Singh et al. (2010) for detection of epidemics using *WCR* and  $N_{HB}$ .

## Chapter 5

# Overview of the system - Simulation

In chapter 4, we applied MEM and *WCR* systems to the data from Scotland and we identified that the *WCR* method is not easy to use because we must have a reference point for *WCR* each time. The major difficulty with the application of the *WCR* system is that there is no simple way of deciding whether or not an increase in the rate has occurred. In order to do this, we try to understand the joint distribution between the *WCR* and  $N_{HB}$  when the rate is constant. By considering the joint distribution when the rate is constant (meaning there is no change in the rate) that will enable us to set up a system to detect when there has been a change. The region in the top right corner in figures 4.22, 4.23 and 4.24 will be the region where the signal will be and this will be covered in detail in chapter 6.

In this fifth chapter, we will see whether or not we can find a simple way of describing the joint bivariate distribution. We will describe the main points of our detection system using the *WCR* method. As we already have a real case, from the data for Scotland we will generate another environment similar to this real environment. This lets us consider the possibility of generalizing the use of these data to detect an outbreak when it occurs in other environments which differ in the number of geographical regions and population size. We make use of the ILI data to set up simulations.

### 5.1 Simulation with different spatial locations and structures

In the Scottish data, we have 14 locations (HBs) and we have the total population of those locations. Each location has a different size of population. We start from the total

Scotland population in 2014 of 5.4 million people [103], and also the individual population size for each health board, so that gives us a general view about total population size in different areas and the percentage of the population in those big areas which is in each sub-area or health board.

Different scenarios can be generated from the Scottish data, so we start by explaining those scenarios. The first case we consider is the total population in big area A which equals 5 million, similar to the Scotland population, and 25 million in big area B which is similar to Saudi Arabia [52], North Korea [123], Netherlands [185] and Romania [186], and in area C we consider a total population of 50 million people like the population in Colombia [124], South Korea [125], Italy [187] and Spain [188]. The second factor is the number of sub-area locations (HBs) in each area. We will consider 14 sub-area locations in area A (like Scotland), 30 sub-area locations in area B and 50 sub-area locations in area C. The third factor is the population in each sub-area, with two structures considered here:

- using the same proportions of the total population which we already have in Scottish HBs, which we call structure  $P_1$ .
- using an equal split which divides the total population by the number of regions, which we call structure  $P_2$ .

It is also possible to use other structures which may be of interest concerning the number of geographic regions or population. The aim from this kind of work is to attempt to generalise the actual situation (Scotland) in terms of the division of places and the population, so that we can apply the early-warning system to other countries and we can answer questions such as:

1. Does the *WCR* method work better when there are more sub-areas (regions)? or does it not make any difference?
2. Does the *WCR* method give better performance with a higher population size? or does it not make any difference?
3. Can we use the simulated cases to find a general model, so we can apply the *WCR* method by giving the number of sub-regions?

Table 5.1 presents three different scenarios for simulating with different spatial locations.

Case	A	B	C
Population (in million)	5	25	50
no. of regions	14	30	50
Pop. structure	<b>a.</b> Scotland ( $P_1$ )		<b>b.</b> Equal ( $P_2$ )

Table 5.1: Population structures considered.

In this chapter, we will use structure  $P_2$ , while structure  $P_1$  will be used later in chapter 9.

## 5.2 Different spatial locations - Steady State

In this section, we create a virtual environment for a state of stability in rates, without outbreaks. The aim of this is to see whether we can consider the opposite of this environment to detect of increasing rates.

We start in the first step to assume that the influenza rates are flat during the year. We need a baseline in order to simulate cases around it, and we choose to consider the base point of this steady state as  $\frac{\max(ILI) \text{ count in } L + \min(ILI) \text{ count in } L}{2}$  through the years in each location  $L$ . This is the mid-range, a simple average, which equals 478, 430 and 525 in the case of 14, 30 and 50 HBs respectively. They are not exactly the same as we take samples from the Scottish HBs to generate data for more than 14 HBs. The results from this scenario can be used to see if the mean of  $WCR$  increases as  $N_{HB}$  increases in other situations (HBs= 30, 50) like it does in the actual Scottish data. We simulated 10,000 weekly values for  $WCR$  and  $N_{HB}$  over 6 years using a Poisson distribution, to give 3,110,000 cases in the end. The simulation uses HBs= 14, 30 and 50, and generates a joint distribution of  $WCR$  and  $N_{HB}$  from these 3,110,000 cases. For HBs= 14, we have for each HB the baseline count and population size. To generate results for HBs= 30 and 50, we randomly sample HBs from the 14 HBs which we start with. We see from this flat rate that the mean of  $WCR$  increases as  $N_{HB}$  increases, for all population sizes and number of sub-area locations, either 14, 30 and 50 (case A, B and C), as in the real Scottish data considered in chapter 4. At  $N_{HB}= 7$  the mean  $WCR$  is 1.001146, while at  $N_{HB}= 4$  the mean is 0.8917849 and at  $N_{HB}= 10$  the mean is 1.020732. Furthermore the range of values is 0.9817849, 1.0207320. We present the joint relative frequency distribution between  $WCR$  and  $N_{HB}$  in all situations in table 5.1, using the population structure  $P_2$ , in figures 5.1, 5.2 and 5.3 for 14, 30 and 50 HBs respectively.

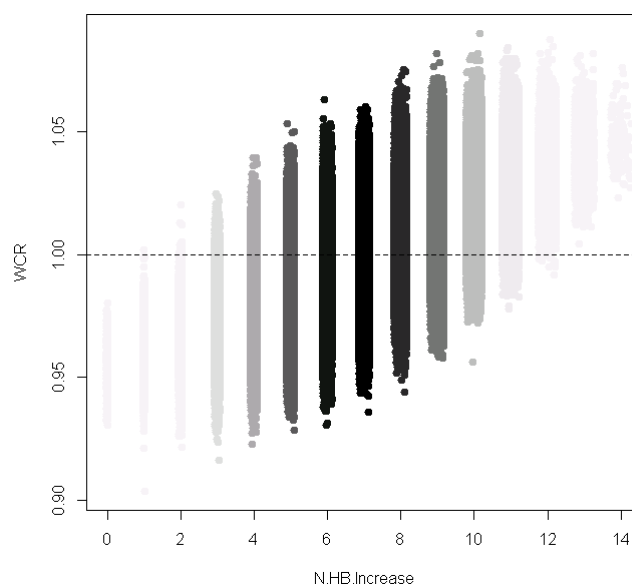


Figure 5.1:  $WCR$  against  $N_{HB}$  for flat rates with 14 HBs and population of 5 million. The black gradient reflects the intensity of the number of simulations. This uses 10,000 cases each week for 6 years.

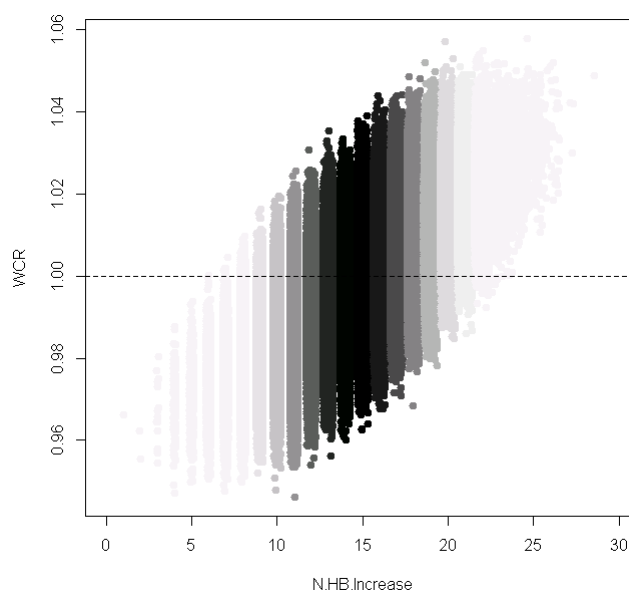


Figure 5.2:  $WCR$  against  $N_{HB}$  for flat rates with 30 HBs and population of 25 million. The black gradient reflects the intensity of the number of simulations. This uses 10,000 cases each week for 6 years.



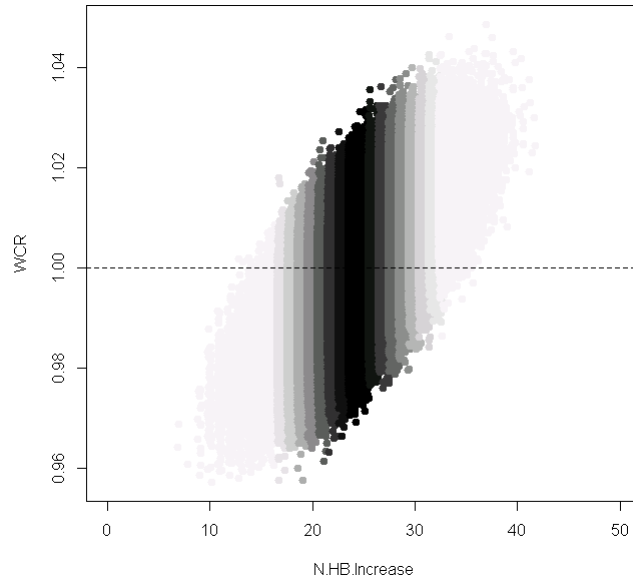


Figure 5.3:  $WCR$  against  $N_{HB}$  for flat rates with 50 HBs and population of 50 million. The black gradient reflects the intensity of the number of simulations. This uses 10,000 cases each week for 6 years.

There is an interesting outcome, as the plots for  $WCR$  and  $N_{HB}$  all look similar, with a possible vertical axis of symmetry at  $\frac{N_{HB}}{2}$  and mean  $WCR$  equal to 1. As expected, in each case as  $N_{HB}$  increases, the distribution of  $WCR$  shifts upwards. We will check the normality of  $WCR$  at each  $N_{HB}$  by plotting the histogram of each set of  $WCR$  values. We found that they appear to be normally distributed, as shown in figures A.1, A.2, A.3, A.4, A.6, A.7 and A.8 in the Appendix. (It is worth mentioning that in fact not all the graphs are normal. They are when there are a lot of observations but not in the tails of the distribution of  $N_{HB}$  when there are relatively few values of  $WCR$ ). We treat them as symmetric and we will use the cases from  $N_{HB}$  from 0 to  $\frac{N_{HB}}{2} - 1$  and add them to the cases from  $\frac{N_{HB}}{2} + 1$  to the end of  $N_{HB}$ . So now we get nearly double the number of cases in our simulation in the rest of this chapter. For example, with HBs= 14, if we have values of  $WCR$  of (0.85, 0.88, 0.95) for  $N_{HB}$  of 5, we can add them up to the cases for  $N_{HB}$  of 9 but increase them to be symmetric about 0, so they will be (1.15, 1.12, 1.05).

### 5.2.1 Approaches to choosing different baselines for the simulation

When we want to do simulations, we have different baselines which we can consider as the mean of a Poisson distribution to simulate from. Table 5.2 shows how to deal with each kind of baseline as each one has two choices of population structure. We can list those approaches as below:

1. Considering the Scotland ILI count in each HB, we choose to divide (*max (ILI) count in L + min (ILI) count in L*) by different suitable numbers (say 2, 10 and 20). We can choose either population structure, equal sub-area population sizes or like the Scottish structure. We considered here the case that the populations are equal in each HB. This is assumed in section 5.2.
2. Considering the baseline of simulation as a fixed rate in each HB. We can also choose either the equal population structure or the Scottish structure.
3. Considering the baseline of simulation as varying rates in each HB. We can also choose either the equal population structure or the Scottish structure.

### 5.2.2 Choosing from Scottish ILI counts

In table 5.1, justified by logical considerations, there we choose to sample ILI counts for 30 and 50 HBs from the Scottish 14 HBs. Here, we choose to divide (*max (ILI) count in L + min (ILI) count in L*) by different suitable numbers (say 2, 10 and 20). The Scottish population in our data is around 5 million. When we take randomly 30 and 50 ILI counts to achieve new datasets for 30 and 50 HBs, we take a sample from the ILI counts in different HBs but we considered different populations in these 30 and 50 HBs (we considered the weekly population around one million in each HB). This approach 1 is used in this chapter and tested in chapter 8.

### 5.2.3 Fixed rate in all locations

In this approach to calculate the baseline counts for simulation, we will consider the rate as fixed in each location and we can consider it as 1, 3 or 5 per 100,000 (we can choose any number but we select those to compare with the other structure below). The same logical consideration to sample ILI counts for 30 and 50 HBs from the Scottish 14 HBs in section 5.2.2 was considered.

When we come to calculate the rate, our samples from Scottish data correspond to 25 and 50 millions of population in area B and C respectively (in section 5.1). This affected the rate calculation as the new population was consistent each week and higher than what was reported in 14 HBs. So, here we will take every sample from the Scottish ILI counts data with its corresponding population. Here also we will consider two population structures which are:

- a Considering equal populations in each different location.
- b Considering the Scottish population structure.

### 5.2.4 Varying rate in different locations

Here, we can consider the base point of simulation as using varying rates in each location, but keeping the average as in the Scottish data. To calculate the rate in the end, we can use the same Scottish structure as in section 5.2.3.

For example, if we consider the rate as 1 per 100,000, with a fixed rate, all different locations will take the same rate 1 per 100,000, but with varying rates, the rate will move up and down (say 0.8 per 100,000 in  $L1$  and 1.2 in  $L2$ ) between different locations, but keeping the average for all locations as 1 per 100,000.

Approaches 2 and 3 (section 5.2.3 and 5.2.4) are used and tested in chapter 9. We decide to use different approaches to simulation because the  $WCR$  is a ratio of rates which are affected in the end by the population size used.

No.	Baseline point	Population structure
1	$\frac{\max(ILI) \text{ count in } L + \min(ILI) \text{ count in } L}{2}$	a. equal population in each HB b. Scottish structure
2	fixed rate in each HB	a. equal population in each HB b. Scottish structure
3	varying rate in each HB	a. equal population in each HB b. Scottish structure

Table 5.2: Approaches to choosing different baselines for the simulation.

## 5.3 Finding functions to relate $WCR$ and $N_{HB}$ (1)

The second aim from using the simulations using a constant rate of consultations scenario is trying to check the possibility of finding functions to describe the distribution of  $WCR$  with a given  $N_{HB}$ . To investigate this we will check the normality of  $WCR$  for each  $N_{HB}$  with the different population sizes. We already checked their histograms in section 5.2 but we will use other approaches for confirmation. We present normal quantile-quantile plots (Q-Q plots) [190] in the Appendix A.1.2, which are in figures A.9, A.10, A.11, A.12, A.14, A.15 and A.16, to check the normality situation before giving the equation relating  $WCR$  and  $N_{HB}$ , and we found that all of them appear to be normally distributed. (Again, it is worth mentioning that in fact not all the graphs are normal. They are when there are a lot of observations but not in the tails of the distribution of  $N_{HB}$  when there are relatively few values of  $WCR$ ). We now try to model the mean and standard deviation of  $WCR$  values for each  $N_{HB}$ .

First we consider the mean of  $WCR$ . Below is shown the formulae of models m14, m30,

m50 (in formula 5.3.1) used to model the means for each number of HBs, and table 5.3 shows a short summary of each model, while details of those models are shown in Appendix A.2:

$$\mu_{ij} \sim \alpha_j + \beta_j \times n_{ij} \quad (5.3.1)$$

where  $\mu_{ij}$  represents the mean for  $WCR$  in each  $N_{HB}$ ,  $n_{ij}$  represents  $N_{HB}$ ,  $i$  starts from the mid-point to the end of HBs (e.g. for 14 HBs,  $i$  takes values 7 to 14), and  $j$  represents number of HBs (14, 30, 50). This model is fitted by weighted GLM regression. Weights are used because we cannot assume that all  $N_{HB}$  have the same frequency in the simulations, as presented in figures 5.1, 5.2 and 5.3 so, weights reflect the number of simulations in each  $N_{HB}$ . Table 5.3 shows that all the models have significant coefficients. The intercepts are similar and the slopes are all small.

HBs= 14 (m14)			
coefficient	estimate	$p$ -value	AIC
Intercept	0.9570	<2e-16	-96.37
slope	0.00623	<2e-16	
HBs= 30 (m30)			
Intercept	0.9545	<2e-16	-166.53
slope	0.00306	<2e-16	
HBs= 50 (m50)			
Intercept	0.9573	<2e-16	-225.86
slope	0.00172	<2e-16	

Table 5.3: Short summary of models m14, m30 and m50 for the mean of  $WCR$ .

In order to find a relationship between  $WCR$  and  $N_{HB}$ , we plot the relation between the mean ( $\mu$ ) and standard deviation ( $\sigma$ ) of  $WCR$  for each  $N_{HB}$  separately. Figure 5.4 presents the relation between  $N_{HB}$  and  $\mu$  in the cases of 14, 30 and 50 HBs. We see the fitted linear models in the red lines, and we see that the points seem to follow a line in each case. The fitted lines have similar intercepts though the slopes differ a bit (and are all very small). The size of the circles represents the number of simulations in each  $N_{HB}$ . We can see that the red line crosses the majority of simulated means, except at the end of  $N_{HB}$  because of the low number of simulations in that region.

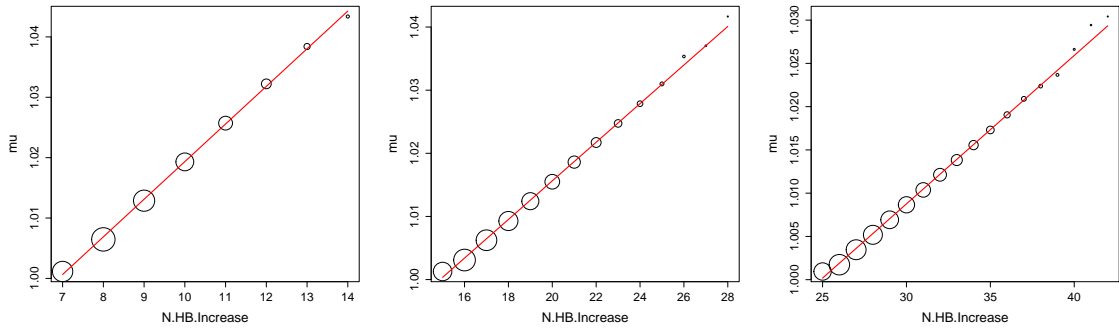


Figure 5.4: The relation between  $N_{HB}$  and  $\mu$  ( $mu$ ) in the cases of 14, 30 and 50 HBs (left to right). The red lines are straight lines fitted to the mean of  $WCR$  in each HB using models `m14`, `m30`, `m50` in Appendix A.2. The size of circles represents the number of simulations at each  $N_{HB}$ .

The next step is predicting  $\mu$  for the case of 30 and 50 HBs and showing them together with the case of 14 HBs. The approach taken uses values of  $\mu$  between the maximum  $N_{HB}$  and the minimum  $\frac{N_{HB}}{2} - 1$  in the case of 14, 30 and 50 HBs, then taking a sequence of values between these two extremes. Figure 5.5 presents the predicted means for the case of 14 HBs and predicted means from the cases of 30 and 50 HBs using the three fitted models shown in figure 5.4. We can see from figure 5.5 that the predicted means for the case of 14 HBs and predicted means from the cases of 30 and 50 HBs using the fitted models follow similar linear patterns and seem to be close to each other, which suggests that the models fitted give useable results and we could use the model for 14 HBs to predict the mean  $WCR$  for any value of  $N_{HB}$  in the range considered above, with a new number of HBs.

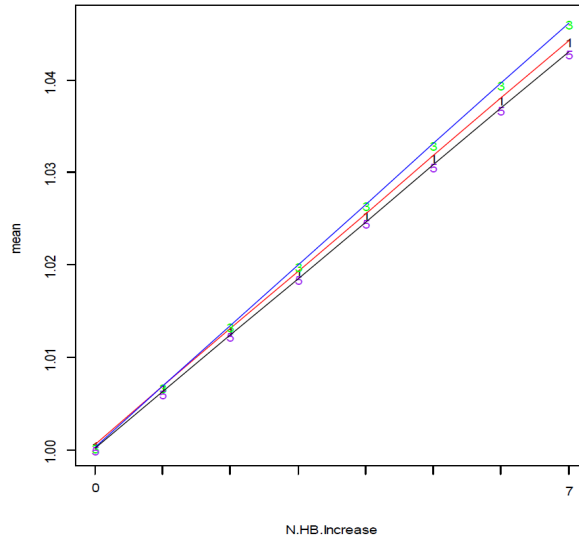


Figure 5.5: Predicted mean of  $WCR$  for the case of 14 HBs (red) and predicted mean  $WCR$  from models based on the cases of 30 (blue) and 50 (black) HBs from figure 5.4.

Now we consider modeling the standard deviation of  $WCR$ . Below is shown the formulae of models  $sd14$ ,  $sd30$ ,  $sd50$  (in formula 5.3.2) used to model the standard deviation for each number of HBs, and table 5.4 shows a short summary of each model, while details of those models are shown in Appendix A.2:

$$\sigma_{ij} \sim \alpha_j + \beta_j \times n_{ij} \quad (5.3.2)$$

where  $\sigma_{ij}$  represents the standard deviation for  $WCR$  in each  $N_{HB}$ ,  $n_{ij}$  represents  $N_{HB}$ ,  $i$  starts from the mid-point to the end of HBs (e.g. for 14 HBs,  $i$  takes values 7 to 14), and  $j$  represents number of HBs (14, 30, 50). This model is fitted by weighted GLM regression.

HBs= 14 (sd14)			
coefficient	estimate	$p$ -value	AIC
Intercept	0.0144	<0.00001	-102.77
slope	-0.00025	0.00315	
HBs= 30 (sd30)			
Intercept	0.00981	<0.00001	-191.48
slope	-0.000047	0.0148	
HBs= 50 (sd50)			
Intercept	0.000693	<0.00001	-236.57
slope	-0.0000285	0.0793	

Table 5.4: Short summary of model sd14, sd30 and sd50 for the standard deviation of  $WCR$ .

The relationship between  $N_{HB}$  and  $\sigma$  is presented in figure 5.6. The size of circles represents the number of simulations at each  $N_{HB}$ . As well as figure 5.6 clearly showing the poor fit of the model in formula 5.3.2, results from table 5.4 confirm this poor fit as the slope is non-significant for sd50. The models differ in terms of slope and intercepts.

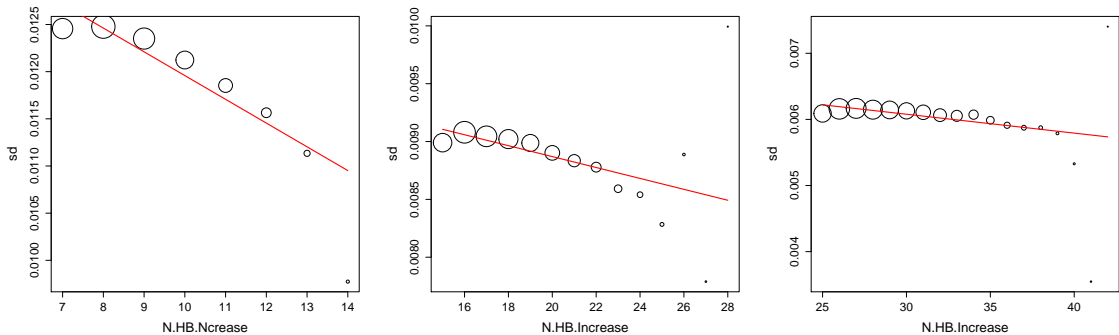


Figure 5.6: The relation between  $N_{HB}$  and  $\sigma$  ( $sd$ ) in the cases of 14, 30 and 50 HBs (left to right). The red lines are straight lines fitted to the standard deviations of  $WCR$  in each HB using models sd14, sd30, sd50 in Appendix A.2. The size of circles represents the number of simulations at each  $N_{HB}$ .

It is clear that the relationships between the standard deviation of the  $WCR$  and  $N_{HB}$  are not linear, as shown by the red lines in figure 5.6. Consequently we will consider suitable transformations to investigate if it is possible to derive a simple model. The Box-Cox approach [136] is one method which can provide the appropriate power of model transformation. Figure 5.7 shows the appropriate power for each  $\sigma$  in different cases of  $N_{HB}$ . Suitable powers to get a linear model from the Box-Cox method in the case of 14,

30 and 50 locations are 14, -1 and 3 respectively. Figure 5.8 presents the relations in the different cases after transformation of the values of  $\sigma$  with these powers. The red lines present the generalized linear models [91] with the power of 14, -1 and 3 in the cases of HBs equal to 14, 30 and 50 respectively. The first plot in figure 5.8 with HBs= 14 fits well but the other two with HBs= 30 and 50 still do not fit well. Also, the levels are different for different HBs (14, 30 or 50).

Below is shown the formulae of models `sd14BC`, `sd30BC`, `sd50BC` (in formula 5.3.3) used to model the transformed standard deviation for each number of HBs, and table 5.5 shows a short summary of each model, while details of those models are shown in Appendix A.2:

$$\sigma_{ij}^{\lambda_j} \sim \alpha_j + \beta_j \times n_{ij} \quad (5.3.3)$$

where  $\sigma_{ij}$  represents the standard deviation for  $WCR$  in each  $N_{HB}$ ,  $n_{ij}$  represents  $N_{HB}$ ,  $i$  starts from the mid-point to the end of HBs (e.g. for 14 HBs,  $i$  takes values 7 to 14), and  $j$  represents number of HBs (14, 30, 50). This model is fitted by weighted GLM regression, and  $\lambda$  represents suitable powers to get a linear model from the Box-Cox method for each number of HBs.

HBs= 14 (sdBC14)			
coefficient	estimate	$p$ -value	AIC
Intercept	4.599e-27	3.46e-06	-996.39
slope	-3.145e-28	3.92e-05	
HBs= 30 (sdBC30)			
Intercept	100.1242	<0.00001	73.489
slope	0.6368	0.0114	
HBs= 50 (sdBC50)			
Intercept	2.978e-07	2.59e-06	-571.7
slope	-2.406e-09	0.1	

Table 5.5: Short summary of models `sdBC14`, `sdBC30` and `sdBC50` for the transformed standard deviation of  $WCR$ .

As well as figure 5.8 clearly showing the poor fit of the model in formula 5.3.3, the results from table 5.5 support this poor fit as the slope is non-significant for `sdbc50`.



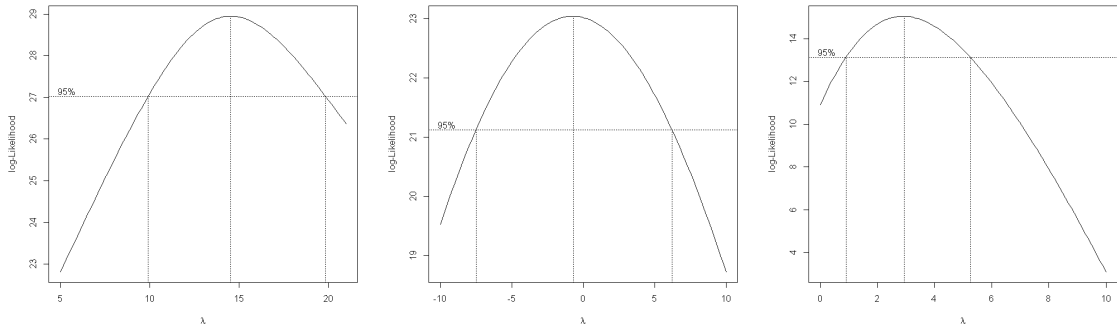


Figure 5.7: The suitable power ( $\lambda$ ) of  $\sigma$  in the cases of 14, 30 and 50 HBs using the Box-Cox technique, shown as the maximum point of the log-likelihood curve.

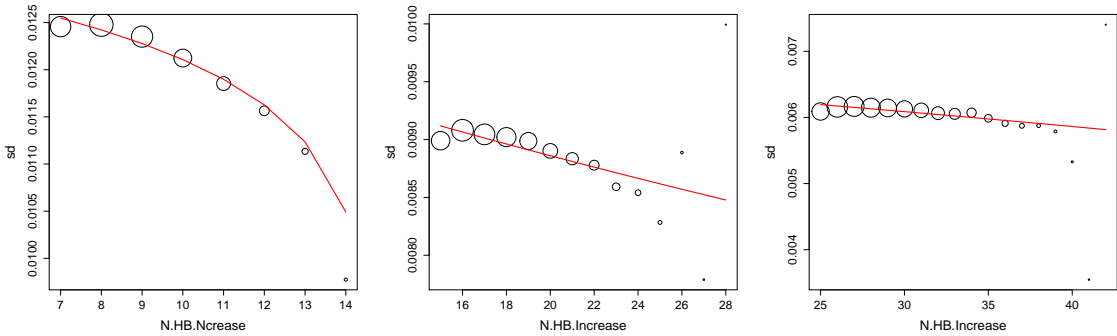


Figure 5.8: The relation between  $N_{HB}$  and  $\sigma$  in the cases of 14, 30 and 50 HBs. The red lines present the models with the power of 14, -1 and 3 in the cases of 14, 30 and 50 HBs respectively using models  $sd14BC$ ,  $sd30BC$ ,  $sd50BC$  in Appendix A.2. The size of circles represents the number of simulations in each  $N_{HB}$ .

We tried predicting  $\sigma$  for the case of 30 and 50 HBs and presenting them together with the case of 14 HBs. We used the same approach which we used with  $\mu$  to fit the minimum and maximum limits of  $N_{HB}$  to consider, but as we found the  $\sigma$  were not linear functions of  $N_{HB}$ , we will predict the  $\sigma$  with the suitable power models suggested from the Box-Cox technique. The predicted standard deviation for the case of 14 HBs and predicted standard deviations from the cases of 30 and 50 HBs are presented in figure 5.9. We can see from figure 5.9 that the predicted standard deviation for the case of 14 HBs and predicted standard deviations from the cases of 30 and 50 HBs using the fitted models do not have the same patterns and seem to be far away from each other, which suggests that the models fitted do not give useable results. They do all follow a straight line trend, but different lines. No single model describes the situation for all values of HBs, so we could not use the line for HBs= 14 to predict the standard deviation of  $WCR$  for any  $N_{HB}$  and a new value of HBs.

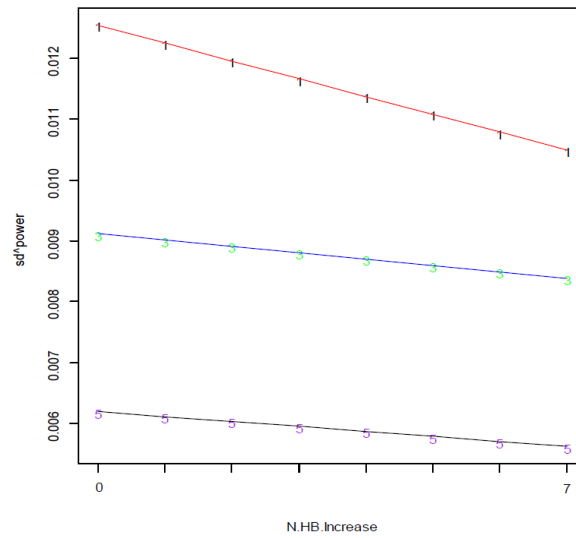


Figure 5.9: Predicted standard deviation of  $WCR$  for the case of 14 HBs with the power of 14 (red) and predicted standard deviations of  $WCR$  from models based on the cases of 30 HBs with the power of -1 (blue), and 50 HBs with the power of 3 (black) from figure 5.8.

We can see from this section that the model for the means showed a promising result which we can use to investigate the possibility of having an equation for the mean of  $WCR$  with any given  $N_{HB}$  from 14 to 50. Regarding the standard deviation, it did not show a promising result for an equation for the standard deviation of  $WCR$  with any given  $N_{HB}$  from 14 to 50 using the normal distribution. We will not give up here as we still have another approach to use, using a different distribution, for example.

### 5.3.1 New simulation for $WCR$ and $N_{HB}$

As the lines showing the relation between the standard deviation of  $WCR$  and  $N_{HB}$  in figure 5.9 are at very different levels depending on the number of locations used in the model, unlike the situation for mean  $WCR$  in figure 5.5, we cannot find one model for the relationships between the standard deviation of  $WCR$  and  $N_{HB}$  in a Normal distribution. We will use a new simulation with a different dataset and try to apply  $\mu$  and  $\sigma$  from the previous cases to compare 95% confidence intervals for HBs = 14, 30 and 50 and present those to compare them with the new simulation case, to see whether they have the same trend or not. We consider HBs= 40 in the new simulation. A plot of  $WCR$  against  $N_{HB}$  for HBs = 40 is presented in figure 5.10. As before for other values of HBs, this shows an increase in the mean of  $WCR$ , in that the distribution for  $WCR$  shifts upwards as  $N_{HB}$  increases.

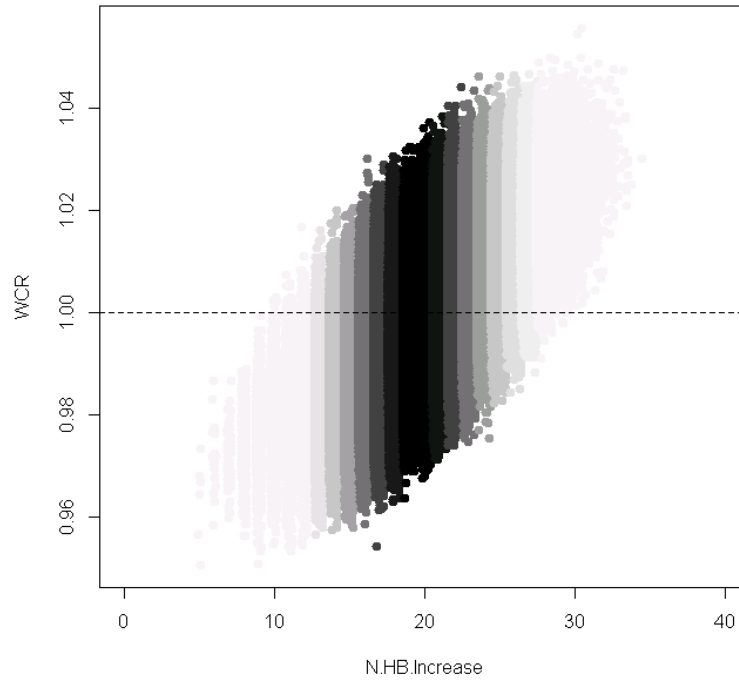


Figure 5.10:  $WCR$  against  $N_{HB}$  for flat rates with 40 HBs and population of 50 million. The black gradient reflects the intensity of the number of simulations.

### Setting coverage intervals (CI)

We decided to continue looking for possible useable equations of means and standard deviations of  $WCR$ , and to see how confidence intervals for  $WCR$  using  $\mu$  and  $\sigma$  match with new data for the HBs. As we have a new simulation for 40 HBs from section 5.3.1, we will use  $\mu$  and  $\sigma$  from the models in figures 5.5 and 5.9, which means we calculate  $\mu$  and  $\sigma$  individually for each group of HBs, to calculate 95% confidence intervals [72] for  $WCR$  with  $N_{HB}$  from 21 to 40, according to the following equation:

$$\mu \pm 1.96 \times \sigma \quad (5.3.4)$$

assuming a normal distribution for  $WCR$ .

Figure 5.11 shows the result of simulation of 40 HBs, giving the joint distribution of  $(WCR, N_{HB})$ , with 95% confidence intervals for  $WCR$  at each value of  $N_{HB}$  calculated from the case of HBs =14, 30 and 50. It shows that the new simulated dataset of 40 HBs, compared to the 95% confidence intervals from the case from HBs =14, 30 and 50, is consistent with these confidence intervals. This is discussed further below.

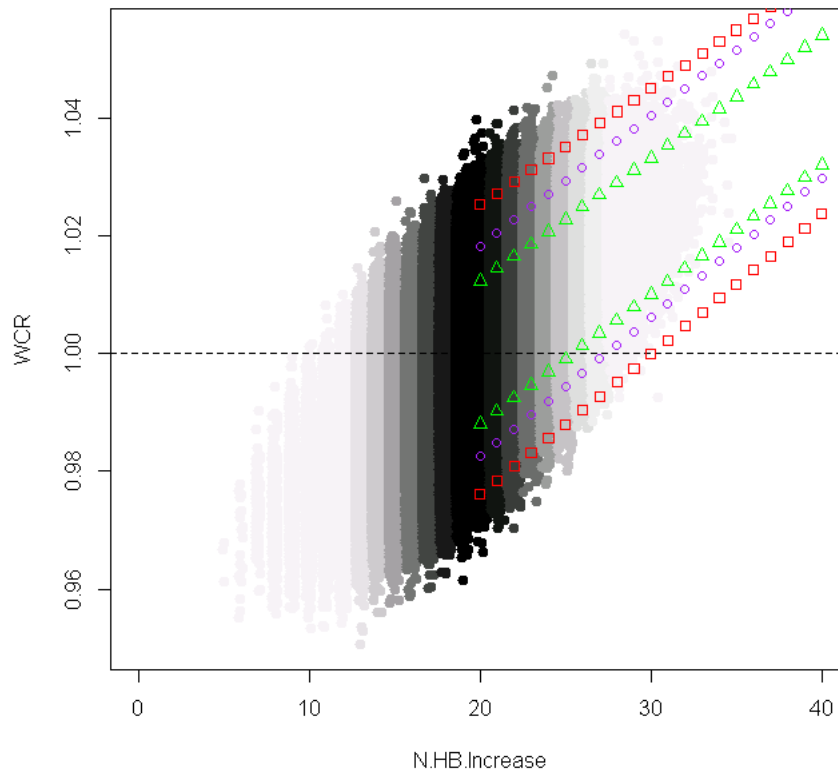


Figure 5.11:  $WCR$  against  $N_{HB}$  for flat rates with 40 HBs and population of 50 million. The black gradient reflects the intensity of the number of simulations. This uses 10,000 cases each week for 6 years. Red squares show the upper and lower 95% confidence intervals for  $WCR$  from the  $\mu$  and  $\sigma$  from the model from the case of 14 HBs, and purple circles represent the same for HBs =30, and green triangles represent the same for HBs =50.

## 5.4 Finding functions to relate $WCR$ and $N_{HB}$ (2)

As presented in section 5.3, we cannot model the standard deviation of  $WCR$  with the Normal distribution. But when we see the confidence interval limits in figure 5.11, we cannot ignore the strong relation in the increase in the pattern of all HBs =14, 30, 40 and 50, so we now continue trying to link  $WCR$  and  $N_{HB}$  using a Gamma model for the variance of  $WCR$ . First of all, we will show the general framework of the data for clearer understanding when we model the mean and the variance. Table 5.6 shows the key parameters needed as:

- $ng$ : this gives the number of geographical locations. We called those HBs before, with values (14, 30, 40, 50).
- $ni$ : this gives how many locations reported an increase ( $WCR > 1$ ) (we set the

starting points as  $\frac{ng}{2}$  to the end of  $ng$ ). This is the same as  $N_{HB}$  but starts from the middle of the range of values for  $N_{HB}$ .

- *mu*: this gives the mean of  $WCR$  in each  $ni$ .
- *var*: this gives the variance of  $WCR$  in each  $ni$ .
- *n.i.midway*: this gives number of increases above midway in  $N_{HB}$ , to get all increases on the same scale, which can be calculated as  $n.i.midway = ni - \frac{ng}{2}$ . In case of having odd number of  $ng$ , we will subtract the first integer number before  $\frac{ng}{2}$ , for example,  $ng=17$  then  $n.i.midway=8$ .
- *weights*: this is used to calculate a weight for each  $N_{HB}$  as follows:
  1. Choosing a Binomial probability as the  $N_{HB}$  are integers, then calculating each probability for each  $ni$  in each  $ng$ . For example: probability  $p$  when  $ng=14 = dbinom(ni(0,1,...,14),14,0.5)$  and the same with the remaining  $ng$ .
  2. Calculating the maximum probability from step (1) for each  $ng$ .
  3. Dividing each probability from step (1) by each maximum calculated in step (2) to give the weight for each  $N_{HB}$ .

Again weights are used because we cannot assume that all  $N_{HB}$  have the same number of simulations as they are really different. If we did not use weights, that means we assumed that all  $N_{HB}$  occur in the same number of simulations and their impact will be equal, which is not a correct approach.

#### 5.4.1 Modelling of means

When we have more than one possible model for predicting data, the simplest model with least variables is generally the best to predict the data [193]. We now consider some different models. As presented in section 5.3 about the possibility to model the mean, now we will again find a linear model (using `lm` in R) to relate the mean and  $N_{HB}$  for each  $ng$ . Tables 5.7, 5.8, 5.9 and 5.11 show the models used in this section ( $z1$ ,  $z2$ ,  $z3$ ,  $z4$ ), while details of those models are shown in Appendix A.3.

ng	ni	mu	var	n.i.midway	weight
14	7,8,...,14	$\mu_7, \mu_8, \dots, \mu_{14}$	$\sigma_7^2, \sigma_8^2, \dots, \sigma_{14}^2$	0,1,...,7	$w_7, w_8, \dots, w_{14}$
30	15,16,...,28	$\mu_{15}, \mu_{16}, \dots, \mu_{28}$	$\sigma_{15}^2, \sigma_{16}^2, \dots, \sigma_{28}^2$	0,1,...,13	$w_{15}, w_{16}, \dots, w_{28}$
40	20,21,...,35	$\mu_{20}, \mu_{21}, \dots, \mu_{35}$	$\sigma_{20}^2, \sigma_{21}^2, \dots, \sigma_{35}^2$	0,1,...,15	$w_{20}, w_{21}, \dots, w_{35}$
50	25,26,...,42	$\mu_{25}, \mu_{26}, \dots, \mu_{42}$	$\sigma_{25}^2, \sigma_{26}^2, \dots, \sigma_{42}^2$	0,1,...,17	$w_{25}, w_{26}, \dots, w_{42}$

Table 5.6: General framework used for modelling  $\mu$  and  $\sigma^2$  for each HBs = 14, 30, 40, 50. The reason that  $\mu$  and  $\sigma^2$  for HBs= 30, 40, 50 do not continue to the end of HBs is because the joint distributions of  $WCR$  and  $N_{HB}$  in those cases do not continue to the end of HBs, as shown in figures 5.2, 5.3, 5.10 and 5.11.

From the summary of the regression model (z1) shown in table 5.7, we can see that the intercept in the model for  $ng=14$  (used as the baseline) is significant, but the adjustments for  $ng= 30, 40, 50$  are not significant at the 0.05 significance level.

z1= lm(m ~ n.i.midway * factor(ng))		
coefficient	estimate	p-value
Intercept	0.9998	<2e-16
n.i.midway	0.0064	<2e-16
factor(ng)30	-0.00004	0.935
factor(ng)40	0.0004	0.410
factor(ng)50	-0.00011	0.809
n.i.midway:factor(ng)30	-0.00326	<2e-16
n.i.midway:factor(ng)40	-0.00425	<2e-16
n.i.midway:factor(ng)50	-0.00465	<2e-16

Table 5.7: Short summary of model (z1) for the mean of  $WCR$ .

The slopes in the linear models for different  $ng$ , namely 14, 30, 40, 50, are presented in table 5.8 (z2). From the model (z2) shown in table 5.8, (n.i.midway:factor(ng)) for  $ng= 30, 40$  and 50, we see from the  $p$ -values below 0.05 that the different slopes between  $ng= 14$  and  $ng= 30, 40, 50$  from  $n.i.midway$  to the end of the range of values for  $ng$  are significant.

z2= lm(m ~ n.i.midway:factor(ng))		
coefficient	estimate	p-value
Intercept	0.9998	<2e-16
n.i.midway:factor(ng)14	0.00646	<2e-16
n.i.midway:factor(ng)30	0.00319	<2e-16
n.i.midway:factor(ng)40	0.0025	<2e-16
n.i.midway:factor(ng)50	0.001805	<2e-16

Table 5.8: Short summary of model (z2) for the mean of  $WCR$ .

We can see that the different slopes for all  $ng= 14, 30, 40, 50$  from *n.i.midway* to the end of the range of values for  $ng$  are significant at significance level 0.05, but we should consider the weight for each  $N_{HB}$ , as the means are decreasing, so we will add the weights into the model.

Model (z3) in table 5.9 includes the weights and we see from the  $p$ -values less than 0.05 that the different slopes for all  $ng= 14, 30, 40, 50$  from *n.i.midway* to the end of the range of values for  $ng$  are significant.

z3= lm(m ~ n.i.midway:factor(ng), weights= weight)		
coefficient	estimate	p-value
Intercept	0.9994	<2e-16
n.i.midway:factor(ng)14	0.0067	<2e-16
n.i.midway:factor(ng)30	0.0033	<2e-16
n.i.midway:factor(ng)40	0.0024	<2e-16
n.i.midway:factor(ng)50	0.0019	<2e-16

Table 5.9: Short summary of model (z3) for the mean of  $WCR$ .

Figure 5.12 represents the means for all  $ng= 14, 30, 40, 50$  against  $N_{HB}$  and the predicted values from model (z3) in Appendix A.3. This figure shows the predicted values in dark points using model (z3) in Appendix A.3 and they show clearly that the model fitted the means reasonably well. The model for  $ng= 14$  looks slightly different from the others.

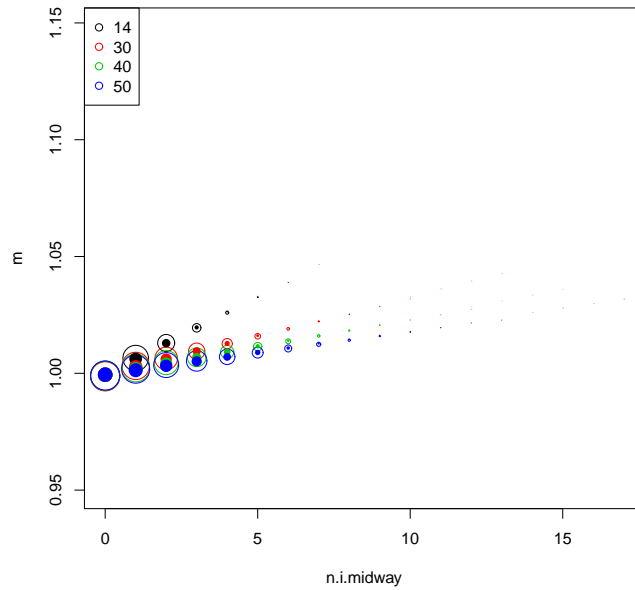


Figure 5.12: Means ( $m$ ) of  $WCR$  against  $n.i.midway$  for  $ng=14, 30, 40, 50$  represented as circled points. Dark points represent the predicted values from model (z3) table 5.9. The size of circles represents the number of simulations for each  $ni$ .

We can now use the fitted slopes from the above model (z3) in a new equation (z4) to give the mean  $WCR$  for  $ng$  equal to any value between 14 and 50. Table 5.10 gives slopes in the model (z3) of the mean for each  $ng$  and these are now used in a general model (z4) below.

### Results

$ng$	slope	$p$ -value
14	0.006745331	<2e-16
30	0.003341988	<2e-16
40	0.002432858	<2e-16
50	0.001905165	<2e-16

Table 5.10: Slopes in the model (z3) for the mean  $WCR$  for each  $ng$ .

From model (z3) in table 5.9, we can get the mean with a given slope from equation 5.4.6 for any  $ng$ , even above 50, as:

$$mean = slope \times (n.i.midway) + 0.9994 \quad (5.4.5)$$

From model (z4) in table 5.11, we can write a general equation for the slope with any  $ng$ ,



even more than 50, as:

$$\log(\text{slope}) = -0.98682 \times \log(\text{ng}) - 2.38015 \quad (5.4.6)$$

Figure 5.13 shows the suggested model for the slope for  $ng=14, 30, 40, 50$ , (and also  $ng=20$ ) and the blue line represents the predicted slope for each  $ng$ . We can see from the figure that the blue curve is close to all slopes, which confirms that the model fitted the slopes well.

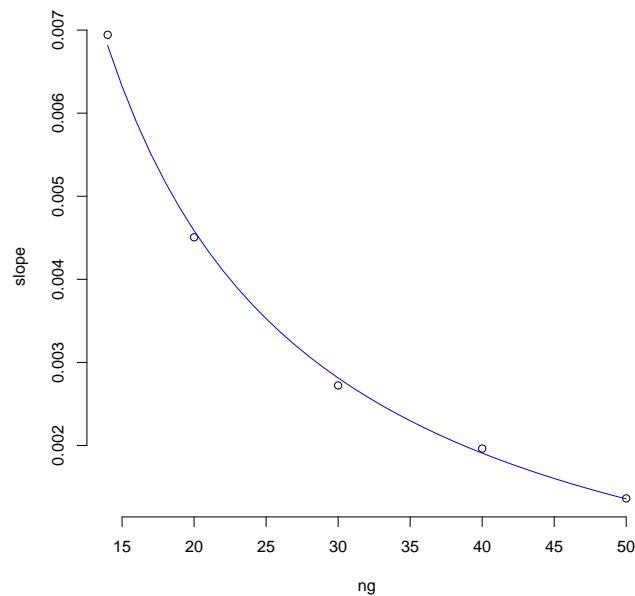


Figure 5.13: Slopes of the model for the mean  $WCR$  against  $ng=14, 20, 30, 40, 50$ . The blue line represents the predicted slope for each  $ng$  (from the model (z4) in table 5.11).

z4= $\text{lm}(\log(\text{slope}) \sim \log(\text{ng}))$		
coefficient	estimate	$p$ -value
Intercept	-2.38015	0.00222
slope	-0.98682	0.00109

Table 5.11: Short summary of model (z4).

Using equations 5.4.6 and 5.4.5, we can conclude that the mean  $WCR$  can be given as:

$$\text{mean} = 0.0925367 \times (\text{ng})^{-0.98682} \times (\text{n.i.midway}) + 0.9994 \quad (5.4.7)$$

So we can use equation 5.4.7 to model the mean  $WCR$  for any value of HBs or  $N_{HB}$ ,

assuming a normal distribution for  $WCR$ . We should note that an approximation has not been validated for  $ng > 50$ .

### 5.4.2 Modelling of variance

In section 5.4.1, we have been relatively successful in deriving an equation for the mean of the  $WCR$  at all values of the number of increases for a range of groups from 14 to 50. We now turn attention to the standard deviation or variance of the  $WCR$  at all values of the number of increases. From section 5.3, we found no possibility to model the standard deviation of  $WCR$  as a single function of  $N_{HB}$  in the different  $ng$  using the normal distribution. So we consider a new general model using a different distribution.

Statistical models attempt to separate a signal from noise in data. For example, a simple linear regression model involves fitting a straight line, explaining how a response variable varies with an explanatory variable. The responses seldom fall perfectly on the fitted line; their departures from the line are assumed to reflect noise in the data. As with ordinary linear models, with generalized linear models (GLMs), we seek one mathematical abstraction: a mathematical description of the uncertainty or noise as captured by a probability distribution. For linear models, the first is taken to be a straight line relationship between the response and a single explanatory variable, and the second is taken to be a normal distribution with unknown but constant variance [58].

GLMs extend ordinary least squares (OLS) linear regression models to allow distributions other than the normal distribution to explain random variation. These include the Gamma distribution. They also allow various response types, which include counts and positive continuous data. GLMs extend OLS models in that they allow for an additional function to link the mean of the response to the linear predictor. In this part of the work, we will consider how to fit such models to data, and how to test hypotheses about and find confidence intervals for the model parameters. We will also consider how to assess the adequacy of our models [104].

Here we will try to model the variance of  $WCR$  (not the standard deviation) using  $N_{HB}$  and a generalized linear model (GLM) with a Gamma distribution (which is often used to model a variance). The processes for a GLM can be listed in three steps, which are: choosing the model (model selection), then parameter estimation, and finally, prediction. Estimation proceeds by defining a measure of goodness of fit between the observed data and the fitted values, generated by the model [91].

Figure 5.14 shows all variances of  $WCR$  over all  $ng$  from the mid-point of  $N_{HB}$ , and the trend is not always linear. It is clear that there are differences either in the intercept or in the slopes for different  $ng$ . We examine the normality of the variance of  $WCR$  for all

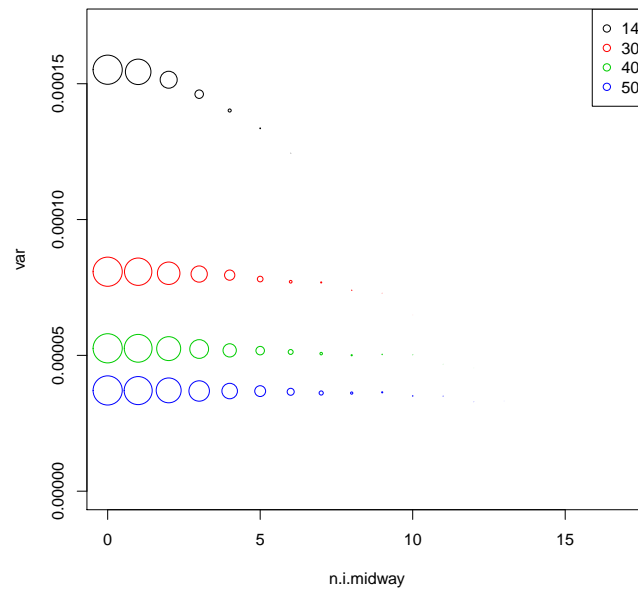


Figure 5.14: Variances ( $var$ ) of  $WCR$  against  $n.i.midway$  for  $ng=14, 30, 40, 50$  represented as circles. The size of circles represents the number of simulations for each  $ni$ .

$ng$  by plotting the histogram presented in figure 5.15, which shows the variance is not normal. Figure 5.16 shows box plots of the variance of  $WCR$  for each  $ng$  to show the distribution, the median level and any outliers. As shown in figure 5.15, the distributions of the variance for different  $ng$  are varied but not especially normal. The median levels of variance in different  $ng$  are clearly different, there are clear outliers for  $ng=30, 40$  and  $50$ , and there are no outliers for  $ng=14$ .

Firstly, we will present the models which compare variances for  $ng=14$  with other  $ng$ .

Tables 5.12, 5.13 and 5.15 show the models used in this section ( $z5, z6, z7$ ), while details of those models are shown in Appendix A.3.

From the  $p$ -values in the result of the regression model ( $z5$ ) in table 5.12,  $n.i.midway * factor(ng)$  for  $ng=30, 40$  and  $50$ , which are less than  $0.05$ , we can see that the different intercepts between  $ng=14$  and  $ng=30, 40, 50$  from  $n.i.midway$  to the end of the range of values for  $ng$  are significant. But when we see the scales of the models between  $ng=14$  and the other  $ng$ , we found the scales are small compared to the change in intercepts.

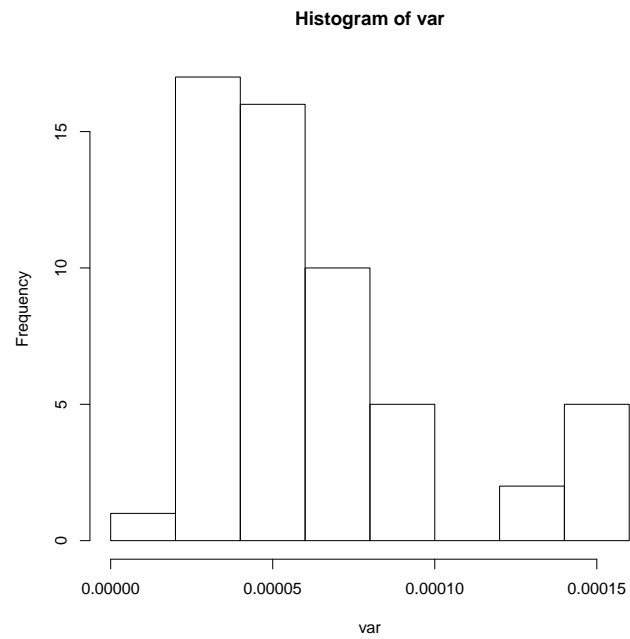


Figure 5.15: Histogram of variance of  $WCR$  over all  $N_{HB}$  in all  $ng$ , which clearly shows the variances are not normally distributed.

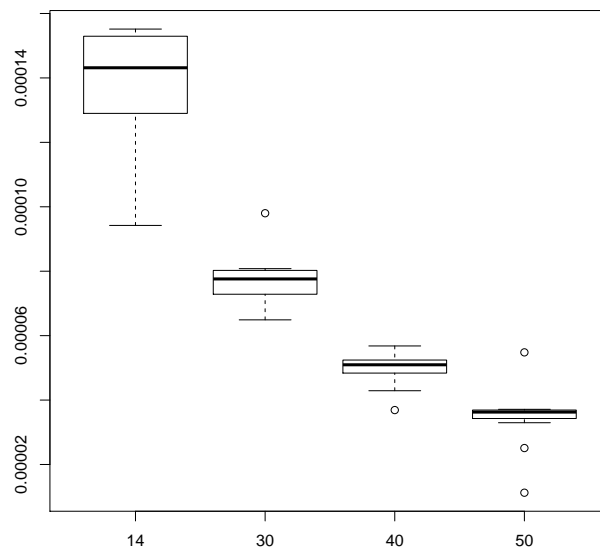


Figure 5.16: Box plots of variance of  $WCR$  for each  $ng$ , to show the distribution.

z5= glm(var ~ n.i.midway * factor(ng), family = Gamma, weights = weight)		
coefficient	estimate	$p$ -value
Intercept	6386.82	<2e-16
n.i.midway	145.10	<2e-16
factor(ng)30	5931.42	<2e-16
factor(ng)40	12544.25	<2e-16
factor(ng)50	20493.32	<2e-16
n.i.midway:factor(ng)30	-62.39	0.0024
n.i.midway:factor(ng)40	-57.32	0.0133
n.i.midway:factor(ng)50	-66.64	0.0117

Table 5.12: Short summary of model (z5) where  $var$  represents the variance of  $WCR$  in each  $N_{HB}$  and  $n.i.midway$  represents  $N_{HB}$  starting from the mid-point to the end of HBs. Weights reflects the number of simulated cases in each  $N_{HB}$  and  $ng$  reflects HBs (14, 20, 30, 40, 50).

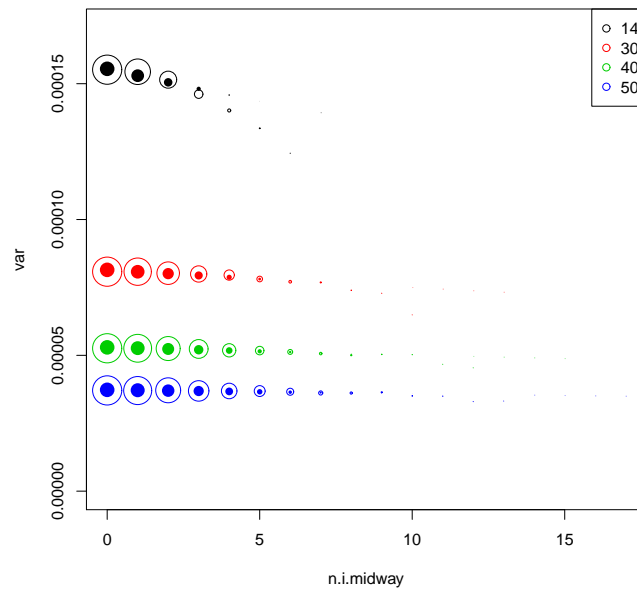
From the summary of model (z6) in table 5.13, we see from the  $p$ -values (less than 0.05) that the different intercepts between all  $ng=14, 30, 40, 50$  from  $n.i.midway$  to the end of the range of values for  $ng$  are also significant.

z6= glm(var ~ n.i.midway + factor(ng) - 1, family= Gamma, weights = weight)		
coefficient	estimate	$p$ -value
Intercept	106.483	<2e-16
factor(ng)14	6431.593	<2e-16
factor(ng)30	12274.279	<2e-16
factor(ng)40	18890.132	<2e-16
factor(ng)50	26810.344	<2e-16

Table 5.13: Short summary of model (z6) for the variance of  $WCR$ .

Figure 5.17 shows the variances for all  $ng=14, 30, 40, 50$  against  $N_{HB}$  and the predicted values from the (z6) model. We can see that the predicted variances from model (z6) (dark points) almost coincide with the actual variances (circled points) which indicates that model (z6) fitted the variances reasonably well. The trends are different for each  $ng$ .

Results		
$ng$	Coefficient ( $ng$ )	$p$ -value
14	6431.593	<2e-16
30	12274.279	<2e-16
40	18890.132	<2e-16
50	26810.344	<2e-16

Table 5.14: Intercepts of the model (z6) of the variance of  $WCR$  for each value of  $ng$ .Figure 5.17: Variances against  $n.i.midway$  for  $ng= 14, 30, 40, 50$  represented as circled points. Dark points represent the predicted values from model (z6) in table 5.13. The size of circles represents the number of simulations for each  $N_{HB}$ .

We can now use intercepts from the above model (z6) for the variance of  $WCR$  to generalize the model for the variance to a new equation, even for  $ng$  above 50. Table 5.14 gives the intercepts in the model (z6) of the variance of  $WCR$  for each  $ng$ .

From model (z7) in table 5.15, we can write a general equation for the intercept with any  $ng$  (even above 50):

$$\log(\text{Coefficient}(ng)) = 0.0400227 \times (ng) + 8.2160336. \quad (5.4.8)$$

Figure 5.18 shows  $ng=14, 30, 40, 50$  (and also  $ng= 20$ ) and the blue line represents from model (z7) the predicted intercept for each  $ng$ , which fits well. Now, from model (z6) we

z7= $\ln(\log(\text{Coefficient}(ng)) \sim ng)$		
coefficient	estimate	p-value
Intercept	8.2160336	0.000018
$ng$	0.0400227	0.000595

Table 5.15: Short summary of model (z7).

can get the variance with given intercepts as in equation 5.4.8 from any  $ng$  (even above 50) following this equation:

$$Variance = \frac{1}{106.483 \times (n.i.midway) + Coefficient(ng)}. \quad (5.4.9)$$

Using equations 5.4.8 and 5.4.9, we can conclude that the variance equals:

$$Variance = \frac{1}{106.483 \times (n.i.midway) + 3699.686 \times \exp(0.0400227 \times (ng))}. \quad (5.4.10)$$

Again as we mentioned with the mean in section 5.4.1, we should note that an approximation has not been validated for  $ng > 50$ .

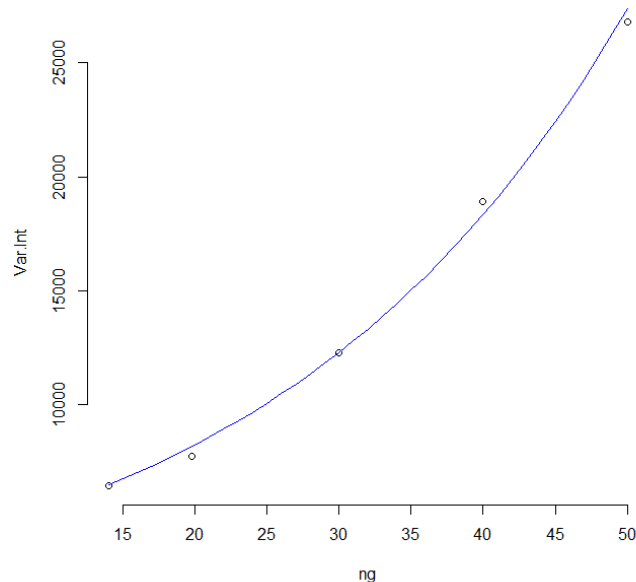


Figure 5.18: Intercepts of the model (z6) for variance of  $WCR$  (Coefficient ( $ng$ )) against  $ng=14, 20, 30, 40, 50$ . The blue line represents the predicted intercept for each  $ng$  from model (z7) in table 5.15.

At the end of sections 5.4.1 and 5.4.2, for a fixed rate based upon the level given by  $\frac{\max(ILI) \text{ count in } L + \min(ILI) \text{ count in } L}{2}$  and equal populations, we have got 2 equations which give the relationship between the mean and variance for  $WCR$  as it changes as the value of  $ng$  (HBs) increases.

## 5.5 Testing coverage intervals (CI) for the link between $WCR$ and $N_{HB}$

In section 5.4, we presented general equations for the mean of  $WCR$  in equation 5.4.7 and the variance in equation 5.4.10 with given  $ni$  and  $ng$ . Here, we will simulate another 3,110,000 cases using flat rates for new values of  $ng=25, 34$  and  $44$ , then we will apply the same approach which we used in section 5.3.1 for the 95% CI and also the equation  $\mu \pm 2.58 \times \sigma$  for the 99% CI. After that we can compare the CI with the new simulation. We are more interested in the area from *n.i.midway* to the end of the range of values for  $ng$ . These values are from 12.5 (we consider 12) to 25 when  $ng=25$ , and from 17 and 22 when  $ng=34$  and  $44$  respectively.

The reason for choosing the new values of  $ng$  as 25, 34 and 44 is to examine the 95% and 99% CI from the  $\mu$  and  $\sigma^2$  equations together with the new simulated 3,110,000 cases to see whether they have the same trend and pattern or not.

Figures 5.19, 5.20 and 5.21 show  $WCR$  plotted against  $N_{HB}$  for 3,110,000 cases with  $ng=25, 34$  and  $44$  respectively. The red squares represent the 95% CI and the green triangles represent the 99% CI using the general equation for the mean and variance. We can see from the figures that the 95% and 99% CI from  $\mu$  and  $\sigma^2$  equations have the same general trend and pattern with the new simulated 3,110,000 cases, for each  $ng$ .



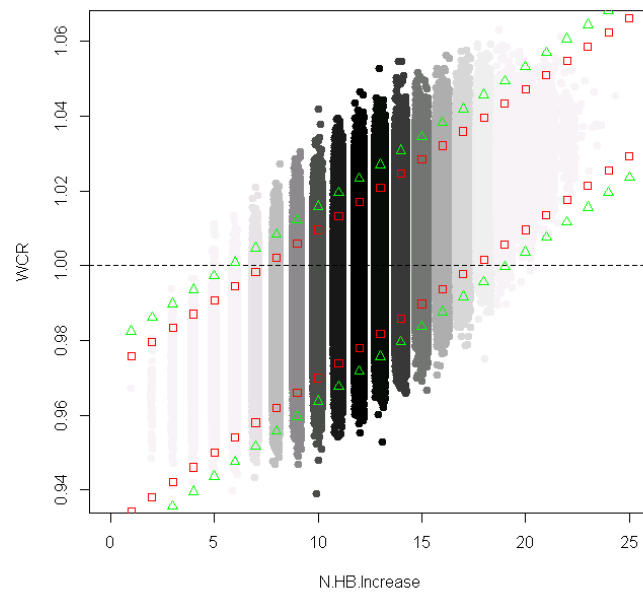


Figure 5.19:  $WCR$  against  $N_{HB}$  for 3,110,000 cases with  $ng=25$ . The red squares represent the 95% CI and the green triangles represent the 99% CI using the general equations to get the mean and standard deviation of  $WCR$ . The black gradient reflects the intensity of the number of simulations. This uses 10,000 cases each week for 6 years.

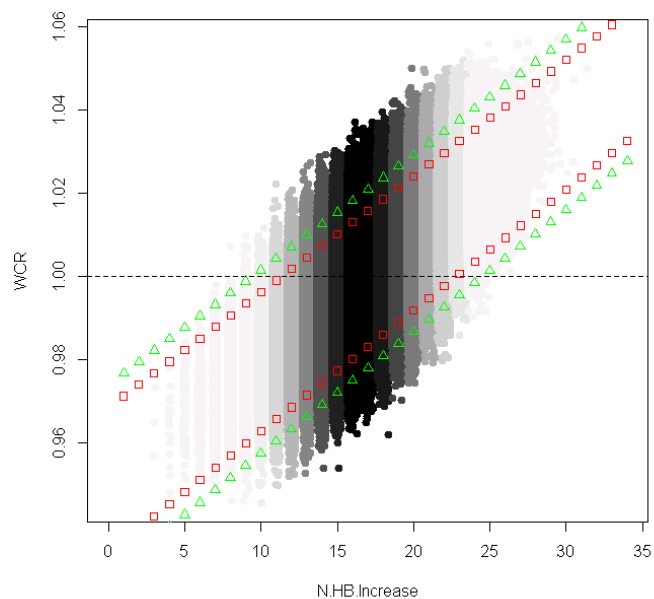


Figure 5.20:  $WCR$  against  $N_{HB}$  for 3,110,000 cases with  $ng=34$ . The red squares represent the 95% CI and the green triangles represent the 99% CI using the general equations to get the mean and standard deviation of  $WCR$ . The black gradient reflects the intensity of the number of simulations. This uses 10,000 cases each week for 6 years.

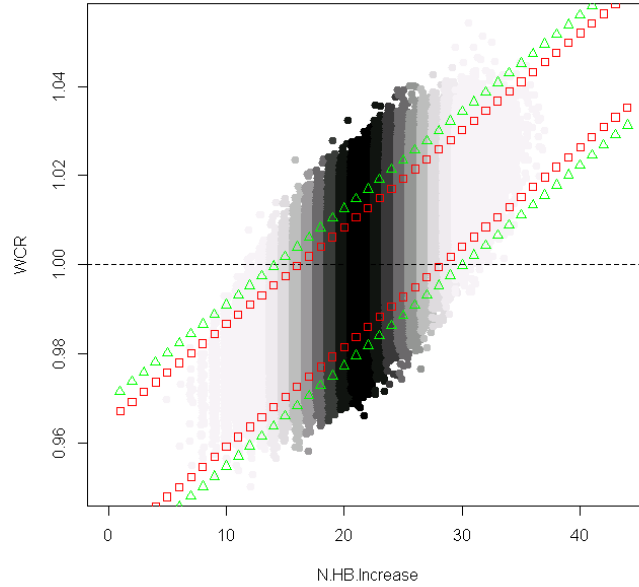


Figure 5.21:  $WCR$  against  $N_{HB}$  for 3,110,000 cases with  $ng=44$ . The red squares represent the 95% CI and the green triangles represent the 99% CI using the general equations to get the mean and standard deviation of  $WCR$ . The black gradient reflects the intensity of the number of simulations. This uses 10,000 cases each week for 6 years.

We can see from figures 5.19, 5.20 and 5.21, that the confidence intervals using  $\mu$  and  $\sigma$  from equations 5.4.7 and 5.4.10 almost cover the majority of simulated cases which gives a good sign of using these equations to help of setting some rejection region in the top right as mentioned earlier in this chapter.

## 5.6 Conclusion

During this chapter, we have taken several important steps and reached important results, as follows. Firstly, we created the joint distribution of  $WCR$  and  $N_{HB}$  with different numbers of spatial locations and different population structures in sections 5.1 and 5.2. Secondly, we tried to link the mean and standard deviation of  $WCR$  with a given  $N_{HB}$  using a normal distribution but it did not give good prediction of either the mean or the standard deviation in section 5.3. Thirdly, we linked the mean and variance of  $WCR$  with a given  $N_{HB}$  using a Gamma distribution which gave good prediction of the variance in section 5.4. Finally, we tested the previous promising equation using 95% and 99% CI with different numbers of spatial locations and they gave the right coverage of the data.

From this chapter, under the situation which we have looked at with number of HBs from 14 to 50, it looks as if we can predict the mean and the variance of  $WCR$ , so we can

conclude with two promising equations for  $\mu$  (equation 5.4.7) and  $\sigma^2$  (equation 5.4.10) to link  $WCR$  to  $N_{HB}$  in the flat rate situation for any reasonable  $ng$  from 14 to 50+. These equations can then be used to find a critical region for deciding if there has been an increase in the rates, which is covered in the next chapter.

Chapter 6 will investigate the possibility of using these promising models in an early detection system, by identifying a rejection region for a test of  $(WCR, N_{HB})$ .

## Chapter 6

# Detection system

In chapter 3, we discussed some statistical and mathematical methods for early warning of influenza. Many of these methods depend on getting data on flu cases, without taking into account the geographical areas in which cases occur. When the researcher has such extra details about the data it is beneficial to make use of these. Here in this research, we wish to use data that are not limited only to the number of influenza cases but the geographic locations of those influenza cases, integrating spatial data with (ILI) consultation data from primary care centres.

In the previous chapter 5, in section 5.4, we ended with general equations for the mean and variance of  $WCR$  with a given  $N_{HB}$ . We will use those equations in the  $WCR$  system in this chapter using a hypothesis testing approach.

Singh et al. in 2010 [142] tested the value of an observation of  $(WCR, N_{HB})$  by assessing the value of the joint pmf for  $(WCR, N_{HB})$  from the historical data. Our approach develops a rejection region for a test of  $(WCR, N_{HB})$ . This extends the work of Singh et al. in 2010 [142].

In chapter 5, we appear to have promising results which will enable us to predict the mean and variance of the  $WCR$  when the rate is constant for any  $N_{HB}$  and for any reasonable  $ng$  (number of HBs) from 14 to 50 and maybe more. These equations can then be used to find a critical region for deciding if there has been an increase in the rates. We have derived a system for the joint distribution of  $WCR$  and  $N_{HB}$  for any  $ng$ , and now we are going to test this out.

Firstly, we will draw a general view about the  $WCR$  system. Then we will present how this system works, manually step by step, starting with simple steps then try to make these steps general. Then we will see the possibility of using a suitable approach with  $N_{HB}$  to specify the start and the end of the region for the detection system to focus on.

After that we will see a useful approach to set cut-offs for the rejection region of  $WCR$  and  $N_{HB}$ . This can be done by calculating a quantile for  $WCR$  in each  $N_{HB}$ . Those quantiles use probabilities, so we will find a useful approach to set those probabilities.

To calculate a quantile of  $WCR$  in each  $N_{HB}$ , we have two possible approaches:

1. Using simulated cases in each  $N_{HB}$  to calculate a quantile of  $WCR$ .
2. Using  $\mu$  and  $\sigma$  from equations 5.4.7 and 5.4.10 in each  $N_{HB}$  to calculate a quantile of  $WCR$ .

We will present these two approaches and test their validity here.

## 6.1 General view of the system

In the previous chapter, in section 5.3.1, in figure 5.11, we can see that the  $Y$  axis represents the  $WCR$ , which are ratios in the range (0.95, 1.05), while the  $X$  axis represents  $N_{HB}$ , which are integer values in the range (1, 40). We are trying to find an area in this plot using some conditions to include values with  $WCR$  higher than 1 and high  $N_{HB}$ . We consider the range of  $\frac{N_{HB}}{2} + 1$  to the maximum  $N_{HB}$ , by setting cut-off points for  $WCR$  at each  $N_{HB}$ . We prefer not to cover the whole area of  $\frac{N_{HB}}{2} + 1$  upwards as we must take into account the sensitivity level of those cut-offs compared to the total number of simulations. We are trying to identify an area which can be used as the rejection region in a hypothesis test of the levels of  $WCR$  and  $N_{HB}$ , to detect the start of an epidemic.

Figure 6.1 shows the general ideas which can be applied for any size of  $N_{HB}$ , and in order to understand the graph, we will explain some general terms:

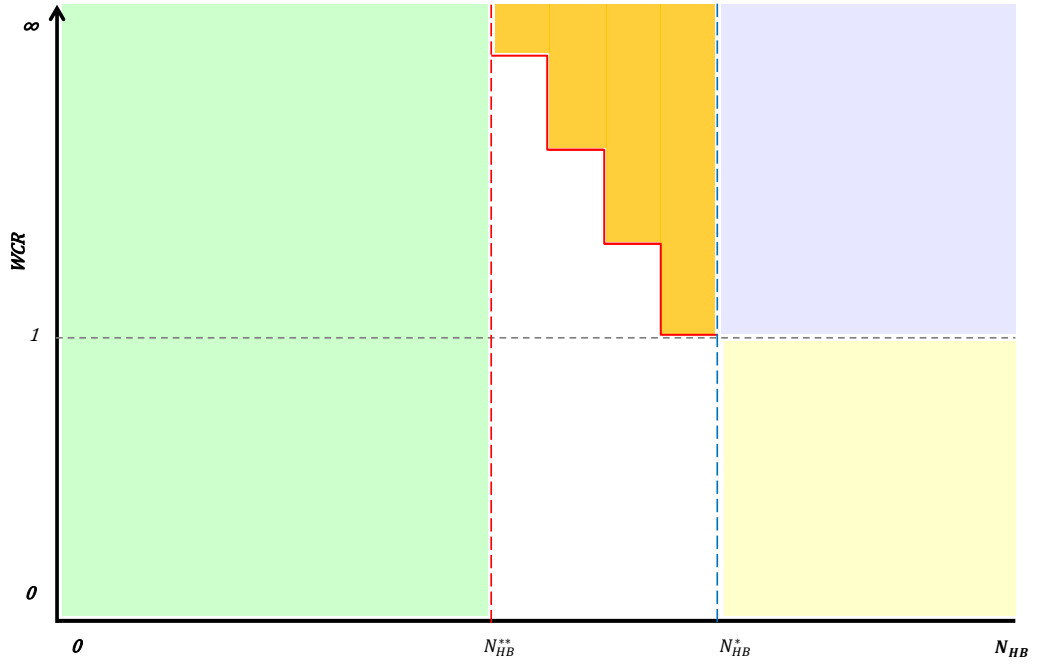


Figure 6.1: General approach to identifying a rejection region from the distribution of  $WCR$  and  $N_{HB}$ , which can apply to any case of  $N_{HB}$ .

- The target significance level ( $\alpha$ ) is shown by the purple area plus the orange area. The probability of the joint distribution of  $WCR$  and  $N_{HB}$  being in the purple and orange areas could be estimated as the number of simulations in the purple and orange areas divided by the total number of simulations.

We are going to use a Binomial distribution in conjunction with the normal distribution for  $WCR$  and the equations from chapter 5, section 5.4, to calculate the probability of being in the orange and purple region without the need for simulation.

- The area where  $WCR$  is less than 1 and  $N_{HB}$  is higher than  $N_{HB}^*$ , shown in the graph by yellow, is very unlikely to apply in any case of an epidemic as  $WCR$  will tend to be very high in this situation as fewer than half of the health boards have an increase.
- The area where  $N_{HB}$  is less than  $N_{HB}^{**}$ , shown in the graph by light green, should never signal as the middle of the range of  $N_{HB}$  has not been reached.

As the  $X$  axis contains only integer numbers ( $1, 2, 3, \dots, N_{HB}$ ), we will use the Binomial distribution to calculate the probability of different values of  $N_{HB}$ , so the processing will be as follows:

1. Choosing the significance level ( $\alpha$ ), say 5 % (the most common value of  $\alpha$ ). Later in chapter 9, we also used  $\alpha$  equal to 1%.
2. Finding the cut-off  $N_{HB}^*$  by solving  $Pr(N_{HB} \geq N_{HB}^*) = \alpha_1$ ; where  $0 \leq \alpha_1 \leq \alpha$ . Solving this probability can give the minimum  $N_{HB}$  with all  $WCR$  higher than 1. The achieved value  $\alpha_1$  will be close to  $\alpha$ , but not equal to  $\alpha$ , because we are using a discrete distribution. (The dotted vertical blue line in figure 6.1 shows the position of  $N_{HB}^*$ ).
3. Finding  $N_{HB}^{**}$  which we consider here as  $\frac{N_{HB}^*}{2} + 1$ . (The dotted vertical red line in figure 6.1 shows the position of  $N_{HB}^{**}$ ).
4. Calculating the decreasing area (the orange area in figure 6.1) by:

$$\sum_{i=N_{HB}^{**}}^{(N_{HB}^*-1)} Pr(N_{HB} = i) \times Pr(WCR \geq \tau_i | N_{HB} = i) \quad (6.1.1)$$

where the term  $\tau_i$  in equation 6.1.1 can be calculated by finding the quantile of  $WCR$  for each  $N_{HB}$  in the range  $N_{HB}^{**}$  to  $(N_{HB}^* - 1)$  using decreasing probabilities, because the means of  $WCR$  are increasing as  $N_{HB}$  increases, so we are trying to achieve the condition of each quantile for  $N_{HB(t)}$  being lower or equal to the quantile for  $N_{HB(t-1)}$ . This is important because of the increasing nature of the joint distribution of  $(WCR, N_{HB})$  and we need these probabilities to address this increase.



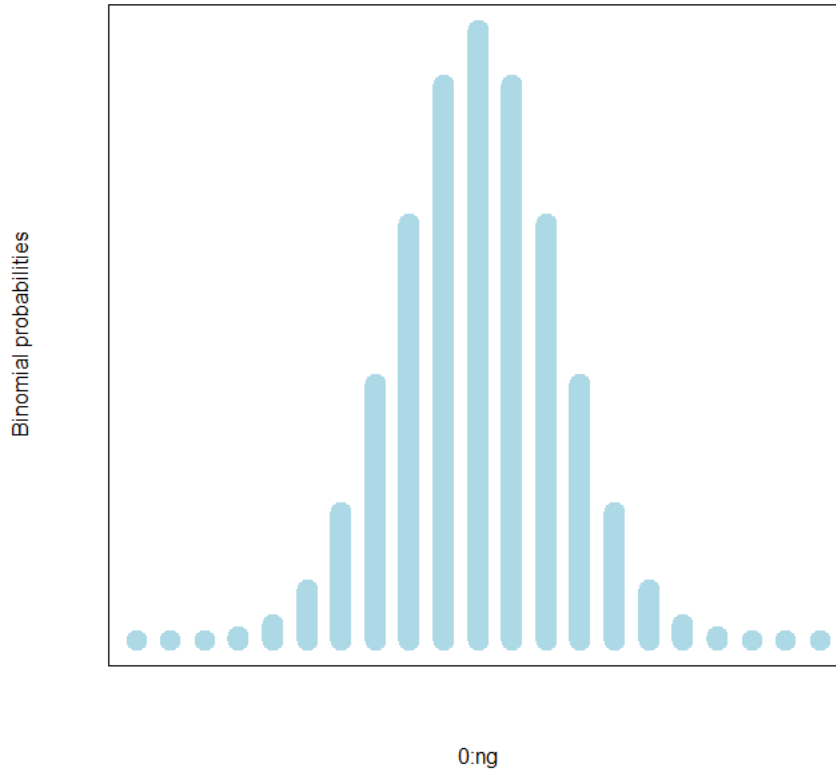


Figure 6.2: Size of Binomial probabilities for a given  $ng$ , with success probability 0.5, because we assumed that  $WCR$  are symmetric in section 5.2.

The first part of this equation 6.1.1 ( $Pr(N_{HB} = i)$ ) comes from Binomial probabilities of each  $N_{HB}$  as in figure 6.2. The purple area in figure 6.1 represents the probability of the number  $N_{HB}$  with an increase in the rate being greater than or equal to  $N_{HB}^*$ , as all  $WCR$  for all  $N_{HB}$  are equal or higher than 1 in this area. The total area beyond the quantile in the purple and orange area will give the theoretical probability close to the significance level ( $\alpha$ ).

### 6.1.1 Example

In this example, we consider the case of  $HBs = 40$  and as we did in section 5.2, we will simulate 10,000 cases for each week with flat rates to get  $10,000 \times 52$  (number of weeks in year)  $\times 6$  (number of years) = 3,120,000 cases. To calculate  $WCR$ , we cannot calculate it in the first week as we do not have a previous week, so the total number of simulations will be  $3,120,000 - 10,000 = 3,110,000$  cases. Those cases are shown in figure 6.3 by the black and gray colour. The next step is to find  $N_{HB}^*$ , which reflects the minimum  $N_{HB}$  with all

$WCR$  higher than 1. It is 31 in this case, shown in the plot by the red dashed vertical line. The second step is finding  $N_{HB}^{**}$ , which equals  $\frac{40}{2} + 1 = 21$ , shown by the orange dashed vertical line. The third step is calculating the quantile of each  $N_{HB}$  from  $N_{HB}^{**}$  to  $(N_{HB}^* - 1)$  using equation 6.1.1 with significance level  $\alpha = 5\%$ . To get decreasing quantiles from  $N_{HB}^{**}$  to  $(N_{HB}^* - 1)$  we will use manual choosing of decreasing probabilities which equal (0.99999, 0.9999, 0.999, 0.991, 0.78, 0.3, 0.1, 0.04, 0.01, 0.001) to find the second part with  $\tau_i$  in equation 6.1.1. The area from  $N_{HB}^*$  to the end of  $N_{HB}$  is considered to have probabilities =100% as all  $WCR > 1$  as there is no point to exclude any  $WCR$  when all of them are higher than 1. These probabilities values are chosen manually with trying to achieve two terms together: 1. each quantile for  $N_{HB(t)}$  being lower or equal to the quantile for  $N_{HB(t-1)}$ , 2. achieving the exact value of our chosen  $\alpha$ . The shaded area in orange (figure 6.3) represents the theoretical probability, equal to the significance level ( $\alpha$ )  $\approx 5.00\%$  from the total number of simulations. We used here simulated cases and we did not come yet to use equation of  $\mu$  and  $\sigma$  from the last chapter.

From the means and standard deviations for the cases of HBs= (14, 30, 50) used in the previous chapter and as we found that each  $N_{HB}$  is normally distributed, we scale these  $\mu$  and  $\sigma$  from the midway to the end of HBs to be equivalent to this case with HBs= 40, then we can use each  $\mu$  and  $\sigma$  to find the cut-offs and can apply those cut-offs, in the current case in which HBs= 40. **R** provides the function `qnorm` [11,16] which can give a quantile from the normal distribution using  $\mu$  and  $\sigma$  from HBs= (14, 30, 50). The red steps in figure 6.3 represent cut-offs of  $WCR$  for HBs= 40 calculated using the case of  $\mu$  and  $\sigma$  from HBs= 14. Cut-offs of HBs= 40 calculated using the case for HBs= 30 are presented by purple lines, while the green lines represent cut-offs for HBs= 40 calculated using the case of HBs= 50. The percentages of observations above those cut-offs from the total number of simulations are presented in table 6.1. Those can be compared to the target significance level of 5%.

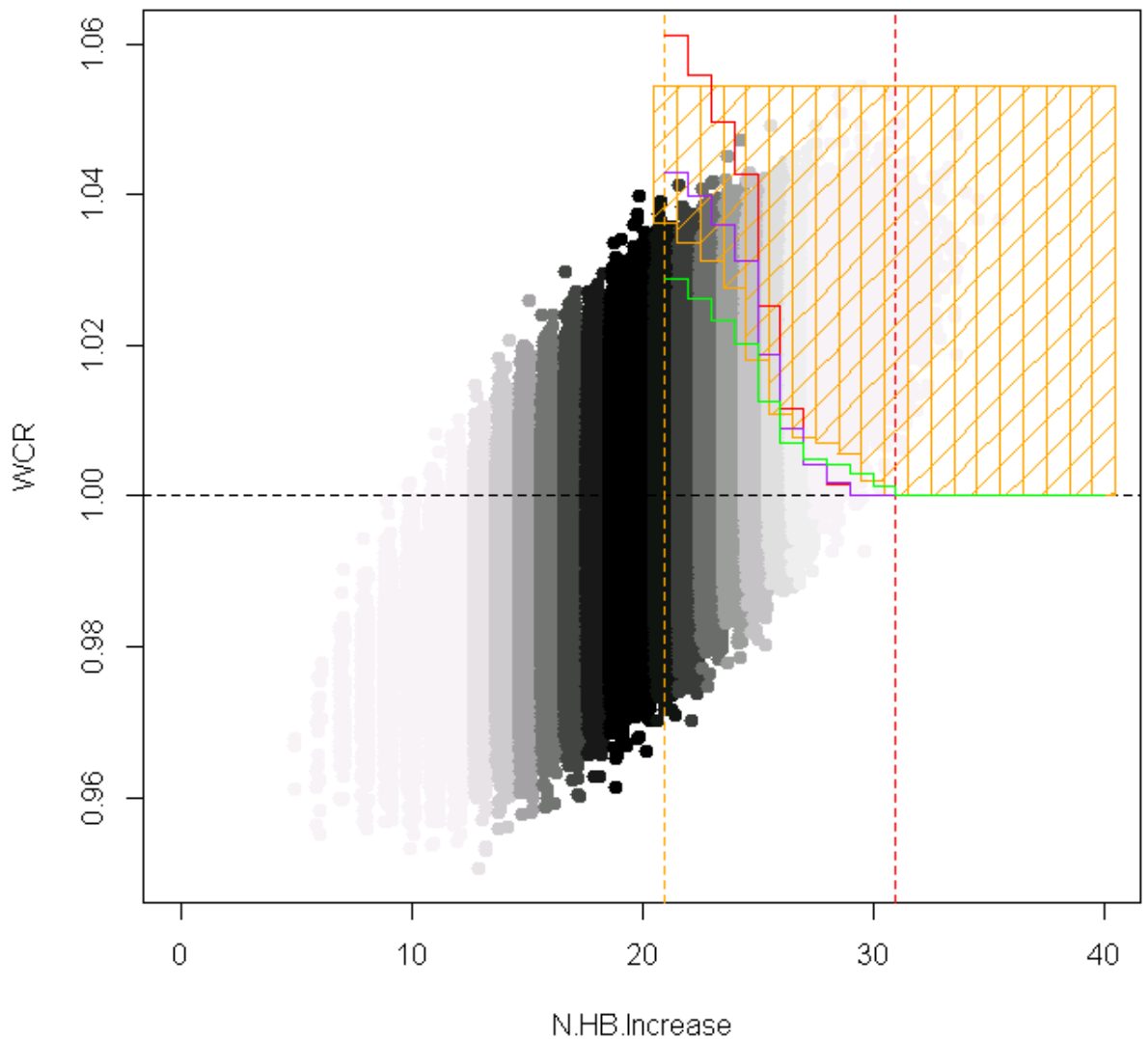


Figure 6.3:  $WCR$  against  $N_{HB}$  for flat rates with 40 HBs and population of 50 million. The black gradient reflects the intensity of the number of simulations. This uses 10,000 cases each week for 6 years. Red steps show the cut-offs from the  $\mu$  and  $\sigma$  calculated from the case of 14 HBs and purple steps represent the cut-offs from HBs= 30 and green steps represent the cut-offs from HBs= 50 after scaling them from the midway to the end of HBs to be equivalent to this case with HBs= 40.

3,110,000 simulations	
HBs	Percentage of simulations above the cut-offs
14	4.55%
30	5.45%
<b>40</b>	<b>5.00%</b>
50	7.00%

Table 6.1: Percentage of simulations above the cut-offs for  $WCR$ , using different HBs, from the total number of simulations. Here, limits are calculated from 3,110,000 simulated cases. The bold text corresponds to the actual number of HBs (40).

We can see from table 6.1 that the percentages of simulations above the cut-offs, using calculations for different HBs, from the total number of simulations, are between 4.55% and 7% and this limited variation gives a hint about a possible way to calculate quantiles for any number of HBs. We used 95% CI ( $z$  - intervals) from HBs= 14, 30 and 50 after scaling to correspond to the case of HBs= 40 to calculate the cut-offs.

## 6.2 Notes and modifications to generalize the system

Previously we used only simulations to determine  $N_{HB}^*$  (the minimum number of increases such that  $WCR > 1$  for all number of increases above  $N_{HB}^*$ ). Now we consider two modifications described below to try to get more general results. They are:

**Note 1** Above we found  $N_{HB}^*$  by solving  $Pr(N_{HB} \geq N_{HB}^*) = \alpha_1$ , where  $0 \leq \alpha_1 \leq \alpha$ , assuming that solving this probability equation can give the minimum  $N_{HB}$  with all  $WCR$  higher than 1. When we calculated this probability, we used the number of simulations in each  $N_{HB}$  to find the first  $N_{HB}$  with all  $WCR$  higher than 1. To have a fair system in various situations, we need to assume some more general considerations. We now use the Binomial distribution to do this calculation, so that it can work more generally, for  $N_{HB}$  between  $N_{HB}^{**}$  to  $N_{HB}^* - 1$ . Steps to do these calculations are explained in the next section 6.3.

**Note 2** We calculated the quantile for each  $N_{HB}$  from  $N_{HB}^{**}$  to  $(N_{HB}^* - 1)$  (the orange area in figure 6.1). Because calculating a quantile must consider a probability or weight ( $k$ ) for each quantile, we used manual selection of those  $k$ , trying to achieve at the same time: 1. each quantile for  $N_{HB(t)}$  being lower or equal to the quantile for  $N_{HB(t-1)}$ , and 2. achieving the exact value of our chosen  $\alpha$ . We calculated the quantile from the simulation using manual selection of weights in the probability calculation in formula 6.1.1 (we used manual scaling here). We need to find an

automatic approach to identify these probabilities or weights. Another alternative way of using simulation with each quantile calculation assumes that we already have a model for  $\mu$  and  $\sigma$ , and then we will use another more general calculation without basing it on the manual probabilities.

### 6.3 More general terms in the system

Note 1 above uses a Binomial distribution but does not use an automatic method of calculating quantile probabilities  $k$ . We will consider new assumptions in this section to give a more flexible and general system. Relevant quantities are:

- *n.area*, the number of locations (previously called HBs)= 14, 30, ..., 50.
- *prob.binom*, the Binomial probability mass function (pmf) on domain  $(0, n.area)$  with probability of 0.5 (as we assumed *WCR* has a symmetric probability density function (pdf) for each  $N_{HB}$ ). We already explained in section 5.2 how to consider *WCR* from  $N_{HB}$  from 0 to  $\frac{N_{HB}}{2} - 1$  as symmetric with the cases from  $\frac{N_{HB}}{2} + 1$  to the end of  $N_{HB}$ . For example, with HBs= 14, if we have values of *WCR* of (0.85, 0.88, 0.95) for  $N_{HB}$  of 5, we find we have values of *WCR* of (1.15, 1.12, 1.05) for  $N_{HB}$  of 9.
- *cum.binom*, the cumulative sum of *prob.binom*.
- $(\alpha)$ , the size of the area from the simulated cases from  $N_{HB}^{**}$  to  $N_{HB}$ , shown as the orange area plus the purple area in figure 6.1. This is a fixed value and we consider it, as before, to equal 5%.
- $(\alpha.nhb)$  which we used to find the theoretical area between  $N_{HB}^{**}$  to  $(N_{HB}^* - 1)$ , shown as the purple area in figure 6.1.
- $(\alpha - \alpha.nhb)$  which is calculated from the simulated cases (shown as the orange area in figure 6.1).
- $N_{HB}^*$  which gives the theoretical level above which we expect to set the  $N_{HB}$  with all *WCR* higher than 1. It can be approximately calculated as  $\min(cum.binom \geq (1 - \alpha.nhb))$  (dotted vertical red line in figure 6.1). These previous steps covered the suggested modification of note 1.
- $N_{HB}^{**}$  which gives the midpoint of  $N_{HB}$ , which equals  $\frac{N_{HB}}{2} + 1$ . If  $N_{HB}$  is an odd number, we use instead the first integer number before  $\frac{N_{HB}}{2}$  (dotted vertical orange line in figure 6.1).
- In the previous chapter in section 5.4, we suggested equation 5.4.7 for the mean of *WCR* and equation 5.4.10 for the variance of *WCR*. Aiming to make the system

generalized and valid without repeating the simulation, we calculate `qnorm` in R with the  $\mu$  and  $\sigma$  from this model (equations 5.4.7, these quantiles have to be combined with Binomial probabilities, and 5.4.10) instead of using `quantile` in R based on the simulation cases (the orange area in figure 6.1).

The second note is finding an automatic way to calculate the cut-off, using the R `quantile` function with  $\mu$  and  $\sigma$  and specified probability  $k$ .

Looking at equation 6.1.1,  $\tau_i$  represents  $k$  for each  $N_{HB}$  from  $N_{HB}^{**}$  to  $(N_{HB}^* - 1)$ . We will add the rest of the area from  $N_{HB}^*$  to the end of  $N_{HB}$  and we assume the total will equal our  $\alpha$ . The whole equation will be as below:

$$\sum_{i=N_{HB}^{**}}^{(N_{HB}^*-1)} Pr(N_{HB} = i) \times Pr(WCR \geq \tau_i | N_{HB} = i) + (1 - \rho) \quad (6.3.2)$$

where  $\rho = cum.binom$  from 0 to  $N_{HB}^* - 1$ . We will consider more than one way to calculate the  $k_i$  in equation 6.3.2, and our aim is now to get 5% or 1% depending on our  $\alpha$ . As in example 6.1.1, we can set those  $k_i$  manually to get our  $\alpha$ , but we will show some functions that can be used to calculate  $k_i$  automatically. The overall curve of the probabilities starts high (say, 85%), and it depends on our predefined level of  $\alpha$ , then starts to decrease until the end. The current task is to search for a function with this gradient in the curve. Our reason for moving from the manual setting to the automatic setting (using a function) is so that we can generalize the system to any number of HBs.

To see  $k_i$  clearly, we will present the R commands for `quantile` and `qnorm` to calculate the quantile of  $WCR$  at each  $N_{HB}$ , below:

```
> qnorm(mean= mu , sd= sd, lower.tail= TRUE, log.p = FALSE, p=k)

> quantile(subset(simulation, N(HB))$WCR, p=k)
```

Those two R codes were used to set the cut-offs of  $WCR$  in each  $N_{HB}$  to specify the rejection region.

In `qnorm`, `mu` and `sd` are the mean and standard deviation using the equations 5.4.7 and 5.4.10, while `quantile` is using simulated cases of  $WCR$  for each  $N_{HB}$ . The value  $k_i$  is `k` in the above command.

We consider below some different possible functions to use for calculation of  $k$ .

### 6.3.1 Linear Function ( $f_1$ )

In order to find a general way to calculate the probabilities in equation 6.1.1, we will look to find a suitable decreasing function. The first attempt will be the simplest way to give

the highest probability at  $N_{HB}^{**}$  and the lowest probability at  $N_{HB}^* - 1$ . Figure 6.4 shows this kind of linear function in the red squares, but this is rather unrealistic as ideally we would like a curve because the mean of  $WCR$  is slightly increasing and we are looking to get some curve probabilities that relate to this increase in  $WCR$ .

### 6.3.2 Ellipse Equation ( $f_2$ )

Another possible way to find a general method to calculate the probabilities in equation 6.1.1 for each quantile is to use a concave decreasing function with a gentle decrease.

The ellipse equation [15,98] is a well-known function with a flexible setting of its curve and a decreasing shape in the top right quadrant. The general equation of an ellipse centered on  $(0,0)$  is:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \quad (6.3.3)$$

where  $a$  here is the length from 0 to  $a$  in the  $X$ - axis, and  $b$  is the maximum value in the  $Y$ - axis. As we may sometimes need to shift up the curve to finish at  $(N_{HB}^* - 1)$  with a value of  $r$ , we rewrite equation 6.3.3 as:

$$y = \sqrt{\frac{(\theta^2 \times s^2) - (\theta^2 \times x^2)}{s^2}} + \gamma \quad (6.3.4)$$

where  $\theta + \gamma$  represents the maximum value in the  $Y$ - axis. It reflects the maximum value which should be close to 1 to use it as  $k$  in `qnorm` at  $N_{HB}^{**}$ , and  $\gamma$  represents the minimum value in the  $Y$ - axis. It reflects the minimum value which should be close to 0 to use it as  $k$  in `qnorm` at  $N_{HB}^* - 1$ ,  $s$  is the length of  $X$ - axis, it is the length from  $N_{HB}^{**}$  to  $N_{HB}^* - 1$ , and  $x$  represents each point on this length of  $X$ - axis. It represents each  $N_{HB}$  from  $N_{HB}^{**}$  to  $N_{HB}^* - 1$ .

Our target is to use and modify the top right curve in figure 6.5, using equation 6.3.4 to get suitable decreasing probabilities ( $k$ ).

The purple triangles in figure 6.4 present the shape of this ellipse function.

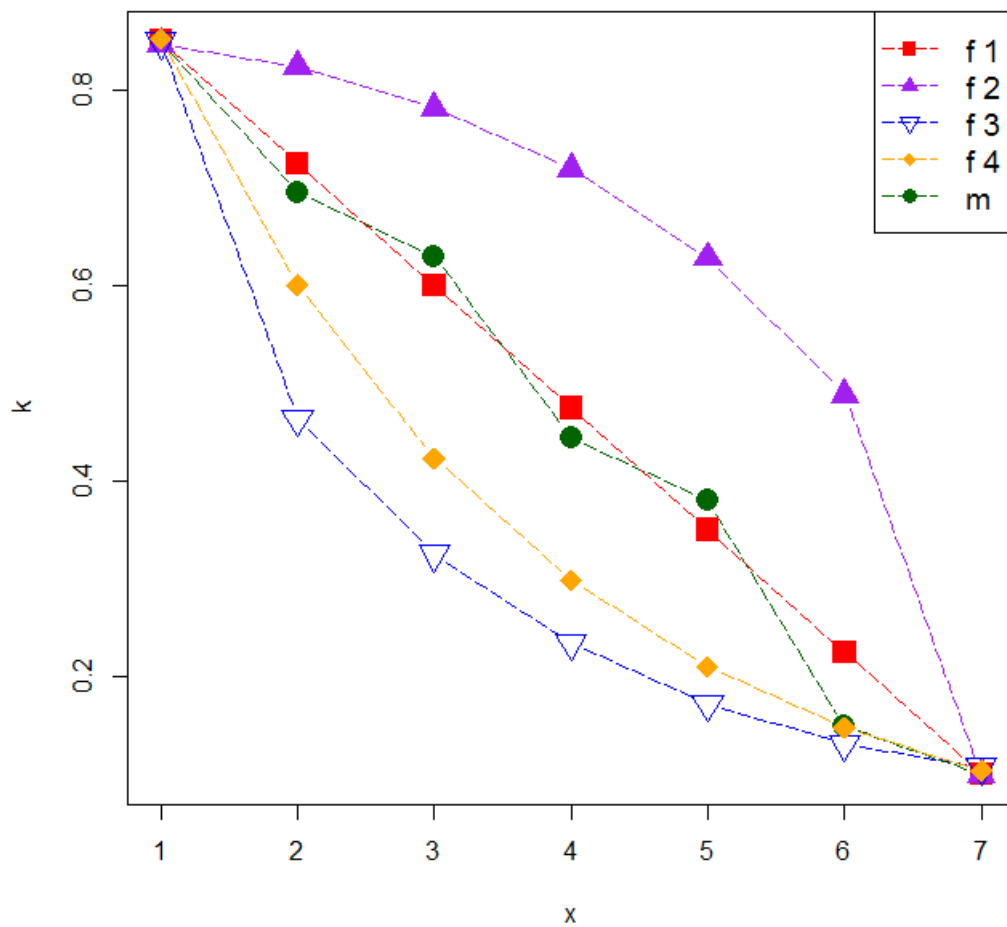


Figure 6.4: Quantile probabilities ( $k$ ) at each  $N_{HB}$  are shown by the red squares for  $f_1$  (section 6.3.1), and by purple triangles for  $f_2$  (section 6.3.2). Blue triangles represent  $f_3$  (section 6.3.3) and the orange diamonds represent  $f_4$  (section 6.3.4). Dark green circles represent the manual values ( $m$ ) of quantile probabilities (section 6.1). This is based on our example in the following section 6.3.5 which considers  $HBs = 40$ . Here 1 in the  $X$ -axis represents  $N_{HB}^{**}$ , and, 7 represents  $N_{HB}^*$ .



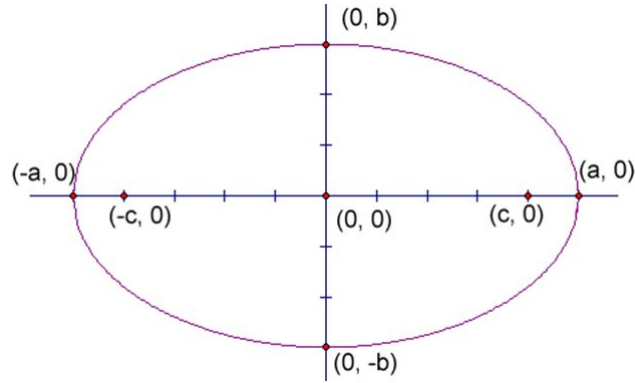


Figure 6.5: The general shape of an ellipse centered on  $(0,0)$ .

### 6.3.3 Concave Ellipse Equation ( $f_3$ )

In the ellipse function in section 6.3.2, equation 6.3.3, we used the formula in its natural form which is concave. Next we tried a different concave curve as in equation 6.3.5, to give a result corresponding to the predefined level of  $\alpha$ , using:

$$y = -\left(\sqrt{\frac{(\theta^2 \times s^2) - (\theta^2 \times (x - s)^2)}{s^2}}\right) + \gamma \quad (6.3.5)$$

where  $\gamma$  reflects the maximum value which should be close to 1 to use it as  $k$  in  $\mathbf{qnorm}$  at  $N_{HB}^{**}$  and  $(\gamma - \theta)$  reflects the minimum value which should be close to 0 to use it as  $k$  in  $\mathbf{qnorm}$  at  $N_{HB}^* - 1$ ,  $s$  is the length from  $N_{HB}^{**}$  to  $N_{HB}^* - 1$ , and  $x$  represents each  $N_{HB}$  from  $N_{HB}^{**}$  to  $N_{HB}^* - 1$ .

Our target is to use and modify the top right curve in figure 6.5 to get same shape of the bottom left curve using equation 6.3.5.

Blue triangles in figure 6.4 present the shape of this inverse ellipse function.

### 6.3.4 Exponential function ( $f_4$ )

When we think of the shape of an exponential pdf [117], this gives an opportunity to make a curve which is concave and decreases from the highest point on the  $Y$ -axis, using the following equation:

$$y = \gamma \times e^{-\theta \times x} \quad 0 \leq x \leq s \quad (6.3.6)$$

where  $\gamma$  reflects the maximum value which should be close to 1 to use it as  $k$  in  $\mathbf{qnorm}$  at  $N_{HB}^{**}$ , while  $\theta$  sets the rate of decrease of the curve,  $s$  is the length from  $N_{HB}^{**}$  to  $N_{HB}^* - 1$ , and  $x$  represents each  $N_{HB}$  from  $N_{HB}^{**}$  to  $N_{HB}^* - 1$ . The shape of this curve is shown in figure 6.4 by the orange diamonds.

We can apply the different approaches in sections 6.3.1, 6.3.2, 6.3.3 and 6.3.4 in equation 6.3.2, or apply the manual values to satisfy the predefined level of  $\alpha$  (say 5%, as before).

To calculate the parameters in each function, we specified a value for  $k$  (maximum  $k$ ) at  $N_{HB}^{**}$ , the specified value for  $k$  (minimum  $k$ ) at  $N_{HB}^*$ , which gives two equations (maximum and minimum) then solved them. In the Ellipse and Exponential cases, we did this by iterative processes to achieve equation 6.3.2.

### 6.3.5 Example

Here, we use the same data which we already used in the last example 6.1.1, but with two changes: 1. using the cumulative sum of the Binomial to set  $N_{HB}^*$  which became 27 instead of 31 in the last example. 2. we will use an automatic setting  $f_1$  as in section 6.3.1,  $f_2$  as in section 6.3.2 ( $\theta = 0.85$ ,  $\gamma = 0.10$ ),  $f_3$  as in section 6.3.3 ( $\theta = 0.75$ ,  $\gamma = 0.85$ ) and  $f_4$  as in section 6.3.4 ( $\theta = 0.35$ ,  $\gamma = 1.21$ ) to calculate the probabilities used each time in R function `qnorm`. In figure 6.6, the plot gives the quantile for each  $N_{HB}$  in the case of the manual setting and the value from `qnorm` for each  $N_{HB}$  in the case of the automatic setting ( $f_1$ ,  $f_2$ ,  $f_3$  and  $f_4$ ).

The manual results are rather oscillating up and down, and the curve-based approaches are smoother.

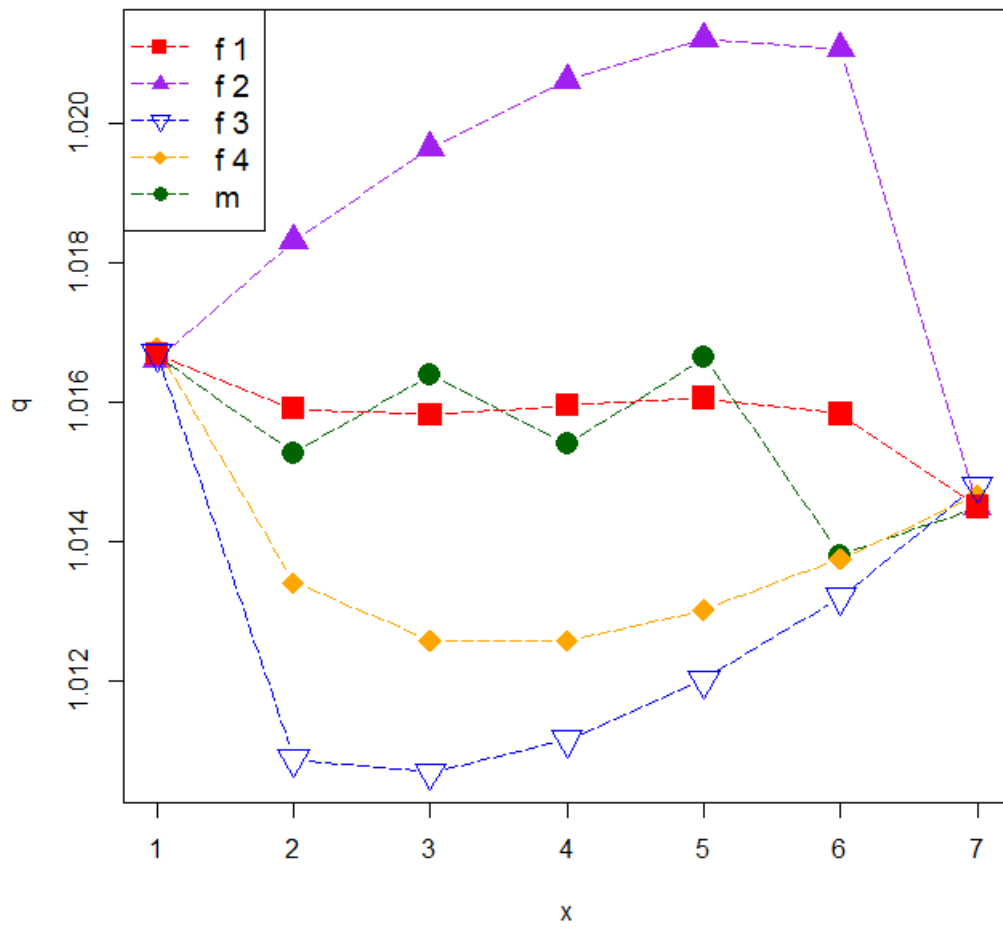


Figure 6.6: Quantiles (cut-offs) from  $N_{HB}^{**}$  to  $N_{HB}^* - 1$  are represented by the red squares for  $f_1$  in section 6.3.1, while purple triangles represent  $f_2$  in section 6.3.2. Blue triangles represent  $f_3$  in section 6.3.3, and the orange diamonds represent the quantile from  $f_4$  in section 6.3.4. Dark green circles show the manual values to calculate quantiles ( $m$ ) in section 6.1.

We can compare the percentage of coverage using each approach and the total number of simulations (3,110,000) in table 6.2.

percentage of coverage	
	$\alpha = 5\%$
$f_1$ (Linear)	0.0497
$f_2$ (Ellipse)	0.0128
$f_3$ (Concave Ellipse)	0.1368
$f_4$ (Exponential)	0.0402
Manual	0.0505

Table 6.2: Percentage of coverage in the rejection region, using different approaches and the total number of simulations (3,110,000).

From table 6.2, we can see that after applying different functions for the probabilities ( $f_1$ ,  $f_2$ ,  $f_3$ ,  $f_4$  and manual choice) with the real cases of simulation, they covered around our predefined significance level of  $\alpha$  (5%) except for  $f_2$  and  $f_3$ . We exclude  $f_3$  from now on as the coverage is too large. Using  $f_2$  the coverage is rather low. This may seem obvious but we discuss each function separately and review them in line with the first considerations we have set for the system in section 6.1. The fourth step states that the quantiles of  $WCR$  for each  $N_{HB}$  from  $N_{HB}^{**}$  to  $(N_{HB}^* - 1)$  should satisfy a pattern of decreasing probability, with the condition that each quantile of  $WCR$  for  $N_{HB(t)}$  is recommended to be lower or the same as the quantile for  $N_{HB(t-1)}$ . This condition will lead us to exclude use of  $f_2$  and  $f_4$ . This leaves case  $f_1$  (which is not perfect, but is acceptable).

We can see that after trying 4 different approaches to calculate the probabilities of quantile of  $WCR$  that the linear function for  $k_i$  which is used to calculate cut-offs from a normal distribution, combined with a Binomial probability, gives acceptable coverage at 5% compared to manual computation from simulations.

## 6.4 Making the system more general

We have investigated and tested the system with 14 locations with relatively high rates, which vary over location, and now we want to test the system with lower rates.

In the previous chapter in section 5.4, we talk about modelling of the mean and the standard deviation or variance of  $WCR$ , so we conclude this section with equation 5.4.7 for the mean and another equation 5.4.10 for the variance. Now, we check for the possibility of adopting those means and variances in the calculation of cut-offs for  $WCR$  at each  $N_{HB}$  using `qnorm`.

The weekly cases ratio at week  $w$  as defined in equation 3.6.1 is:

$$WCR = \frac{Rate_{(w)}}{Rate_{(w-1)}}. \quad (6.4.7)$$

When we see this equation and consider again our mean and variance models in equations 5.4.7 and 5.4.10, we did not take into account that those means and variances are for a ratio ( $WCR$ ) and this may directly affect the mean and standard deviation equations. We will check the first step of our 3,110,000 simulations in section 5.2 to see if it is still usable. In table 6.3, we show key values relating to the data set using number of locations from 1 to 14 HBs as it is the basis of all the other data sets.

In table 6.3, Location gives the number of locations (HBs). Min ILI count gives the minimum ILI weekly count in some specific location from the Scottish data from 2009 to 2015. Max ILI count gives the maximum ILI weekly count in some specific location from the Scottish data from 2009 to 2015.  $\frac{Min+Max}{2}$  gives the average of  $Min$  and  $Max$  ILI count for each location, and  $\frac{Min+Max}{10}$  gives the  $min$  plus  $max$  of the ILI count divided by 10 for each location. Population is the median of total weekly population in each location. Here we will list some comments about this table:

- $\frac{\sum of \frac{Min+Max}{2}}{\sum Population}$  equals 0.000246871 if the rate was the same in all locations, and 0.000247 means 247 cases per 100,000, which is high compared to the actual baseline of influenza activity (0-50) cases [56, 64, 142].
- Some locations like 1, 2, 3, 5, 11, 12 and 14 registered 0 as minimum ILI counts and their maximum ILI counts are also low compared to the other locations, which cause extreme variation between different locations. For example, the rate used in location 12 equals  $\frac{15}{27,125,219.9/14} \times 100,000 = 0.8$  per 100,000 population, while the rate used in location 6 equals  $\frac{1572}{27,125,219.9/14} \times 100,000 = 80$  per 100,000 population. It is clear there is a large difference between 0.8 and 80, and this is caused by assuming equal populations in all locations or HBs.

Location	Min ILI count	Max ILI count	$\frac{Min+Max}{2}$	$\frac{Min+Max}{10}$	Population
1	0	412	206	41.22	27,125,219.9/14
2	0	237	118.5	23.78	27,125,219.9/14
3	0	426	213	42.66	27,125,219.9/14
4	1	527	264.1	52.83	27,125,219.9/14
5	0	721	360.5	72.14	27,125,219.9/14
6	5	3139	1572.15	314.43	27,125,219.9/14
7	1.35	1650	826.03	165.21	27,125,219.9/14
8	0.75	1565	782.94	156.59	27,125,219.9/14
9	0.75	1458	729.77	145.95	27,125,219.9/14
10	2.06	1933	967.88	193.58	27,125,219.9/14
11	0	77	38.5	7.72	27,125,219.9/14
12	0	30	15	3.05	27,125,219.9/14
13	2.34	1131	567.07	113.41	27,125,219.9/14
14	0	67	33.5	6.71	27,125,219.9/14
$\Sigma$	–	–	6696.41	1339.283883	27,125,219.9

Table 6.3: Details of key values for 1 to 14 HBs weekly total and their populations (more details in the text).

To consider these issues in further detail, as  $WCR = \frac{Rate(w)}{Rate(w-1)}$  where in our work  $Rate(w)$  is the number of cases in week  $w$  divided by the population size in week  $w$  (which is assumed constant in all locations). To make the explanation clearer (as explained in section 4.2), we will say  $WCR \approx \frac{ili(w)}{ili(w-1)}$  and  $ili \sim Poisson(\theta)$ ,  $Var(ili) = \theta$  and  $E[ili] = \theta$ . Now, we know that  $WCR$  is a ratio, so we will now calculate the mean and variance of the ratio. As  $ili_{(w-1)}$  and  $ili_{(w)}$  reflect numbers of reported cases in different weeks, they are independent variables, assumed independent as the underlying rate is constant and we do not assume serial correlation. It is worth to mention that it is often in infectious disease modeling is assumed that case in week  $w$  depend on case in week  $w - 1$  but we are assuming independent rates here.

Approximation of the means and variance of ratios of independent variables can be calculated as below:

$$E\left[\frac{ili_{(w)}}{ili_{(w-1)}}\right] \approx \frac{E(ili_{(w)})}{E(ili_{(w-1)})}$$

or

$$E\left[\frac{ili_{(w)}}{ili_{(w-1)}}\right] \approx \frac{E(ili_{(w)})}{E(ili_{(w-1)})} - \frac{cov(ili_{(w)}, ili_{(w-1)})}{[E(ili_{(w-1)})]^2} + \frac{Var(ili_{(w-1)})E(ili_{(w)})}{[E(ili_{(w-1)})]^3}$$

using the bivariate first order Taylor expansion from [46, 147] respectively.

As  $ili_{(w)}$  and  $ili_{(w-1)}$  are independent Poisson ( $\theta$ ) random variables in the steady state case, (we know in influenza seasons, case in week  $w$  depended to case in  $w - 1$ , but we assuming steady state here), so  $cov(ili_{(w)}, ili_{(w-1)}) = 0$ , and

$$\begin{aligned} E\left[\frac{ili_{(w)}}{ili_{(w-1)}}\right] &\approx \frac{\theta}{\theta} + 0 + \frac{\theta^2}{\theta^3} \\ E\left[\frac{ili_{(w)}}{ili_{(w-1)}}\right] &\approx \frac{\theta}{\theta} + \frac{1}{\theta} \\ E\left[\frac{ili_{(w)}}{ili_{(w-1)}}\right] &\approx 1 + \frac{1}{\theta} \approx 1 \end{aligned} \quad (6.4.8)$$

which means that the mean of any *WCR* ratio is approximately 1. For the variance, using the bivariate second order Taylor expansion [46, 147] of *WCR*,

$$\begin{aligned} Var\left(\frac{ili_{(w)}}{ili_{(w-1)}}\right) &\approx \frac{(E(ili_{(w)}))^2}{(E(ili_{(w-1)}))^2} \times \left[ \frac{var(ili_{(w)})}{(E(ili_{(w)}))^2} - 2 \times \frac{cov(ili_{(w)}, ili_{(w-1)})}{E(ili_{(w)}) \times E(ili_{(w-1)})} + \frac{var(ili_{(w-1)})}{(E(ili_{(w-1)}))^2} \right] \\ Var\left(\frac{ili_{(w)}}{ili_{(w-1)}}\right) &\approx \frac{E(ili_{(w)})^2}{(E(ili_{(w-1)}))^2} \times \left[ \frac{var(ili_{(w)})}{(E(ili_{(w)}))^2} + \frac{var(ili_{(w-1)})}{(E(ili_{(w-1)}))^2} \right] \\ Var\left(\frac{ili_{(w)}}{ili_{(w-1)}}\right) &\approx \frac{\theta^2}{\theta^2} \times \left[ \frac{\theta}{\theta^2} + \frac{\theta}{\theta^2} \right] \approx \frac{2\theta}{\theta^2} \approx \frac{2}{\theta}. \end{aligned} \quad (6.4.9)$$

We can show the effect of equations 6.4.8 and 6.4.9 by presenting the following example. Let  $\theta_1 = 20, \theta_2 = 30, \theta_3 = 40$  and  $\theta_4 = 50$ . We will use the Poisson pmf with mean  $\theta$  to simulate 10,000 cases. With  $\theta_1 = 20$ , we will simulate 10,000 cases from the Poisson pmf with mean of 20 and consider that as the ILI rate in week  $w$  (let's call it  $C_i$ ), then simulate another 10,000 with the same mean and consider that as the ILI rate in week  $w-1$  (let's call it  $C_{i-1}$ ). We can calculate the *WCR* as  $\frac{C_i}{C_{i-1}}$ , then calculate the mean and variance of *WCR*. We can do this with all the considered values of  $\theta$ , as presented in table 6.4.

$\theta_1 = 20$	$\theta_2 = 30$
E ( $WCR$ )=1.05121	E ( $WCR$ )=1.0343
Var ( $WCR$ )=0.1302	Var ( $WCR$ )=0.0731
$2/\theta_1=0.1146$	$2/\theta_2=0.0724$
$\theta_3 = 40$	$\theta_4 = 50$
E ( $WCR$ )=1.0342	E ( $WCR$ )=1.0228
Var ( $WCR$ )=0.0586	Var ( $WCR$ )=0.0403
$2/\theta_3=0.0543$	$2/\theta_4=0.0432$

Table 6.4: Example to show how the ratio  $WCR$  is distributed in 10,000 simulations from a Poisson ( $\theta$ ) pmf.

In table 6.4, we can see the mean of  $WCR$  using  $\theta_1, \theta_2, \theta_3$  and  $\theta_4$  is around  $1 + \frac{1}{\theta_x}$ , while the variance of  $WCR$  using  $\theta_1, \theta_2, \theta_3$  and  $\theta_4$  is close to  $\frac{2}{\theta_x}$  where  $x = 1, 2, 3, 4$ . This example illustrates the approximate mean and variance of  $WCR$  as shown in equations 6.4.8 and 6.4.9. Where  $\theta$  is large the variance is small and where  $\theta$  is small the variance is larger.

We will investigate the system coverage when using different baseline values of  $\theta$  which are lower to give more realistic numbers of cases. We started with  $\theta = \frac{Min+Max}{2} = \theta_A$ , say. We now consider different  $\theta$ , written as  $\frac{\theta_A}{r}$ , where  $r$  is a scaling factor. Using  $r = 1$  gives  $\theta_A$  and we call this  $WCR_{v1}$ . Using  $r = 5$  gives  $\theta = \frac{Min+Max}{10}$  and we call this  $WCR_{v2}$ . Using  $r = 10$  gives  $\theta = \frac{Min+Max}{20}$  and we call this  $WCR_{v3}$ .

All we are doing is making the underlying rates, as we are using constant populations, smaller so there will be a greater chance of zeroes in some of the  $N_{HB}$  and also it will be more discrete as the Poisson means will be smaller and we will have a greater chance of exactly the same values two weeks in a row, giving  $N_{HB}$  where there is neither an increase nor a decrease.

Figure 6.7 shows the system in the steady state case and with  $r = 5$ . The aim from this figure to compare the cut-offs using simulated cases to the cut-offs from the normal distribution by using the mean of  $WCR$  in equation 5.4.7 and the variance in equation 5.4.10.



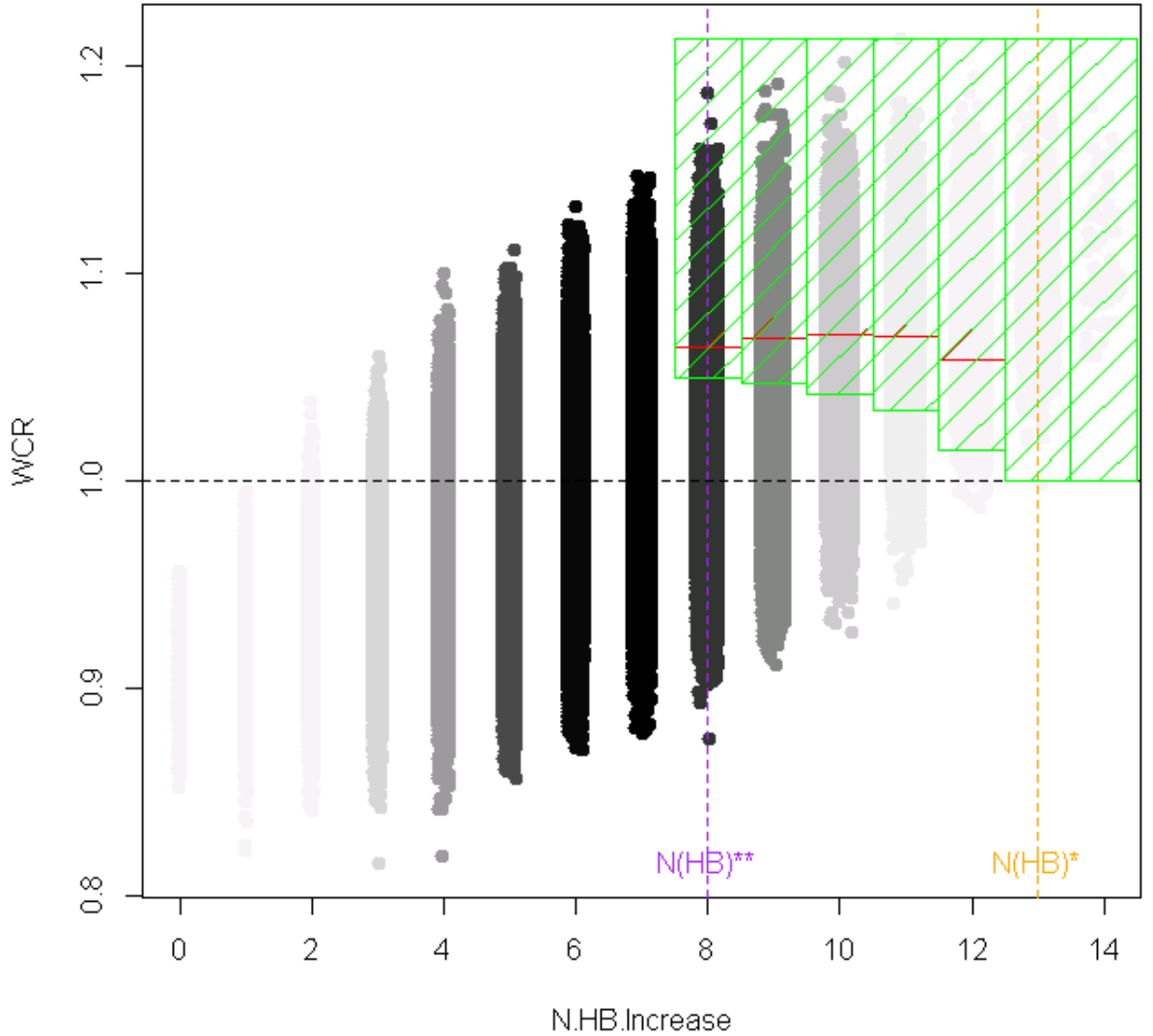


Figure 6.7:  $WCR$  against  $N_{HB}$  for steady rates using  $r = 5$  with 14 HBs and a 5 million population.  $N_{HB}^*$  is presented in the dotted vertical orange line while the dotted vertical purple line represents  $N_{HB}^{**}$ . The red shaped area columns came from calculating the quantile (cut-offs) using  $\mu$  and  $\sigma$  from many simulated cases with a flat rate. They represent around 5% of the most extreme observations under the steady state. The green shaped area columns came from calculating  $qnorm$  (cut-offs) using  $\mu$  and  $\sigma$  from equations 5.4.7 and 5.4.10. They represent around 9% of the most extreme observations under the steady state. Darkness of the black colour reflects the number of simulated cases.

In figures 6.7, B.1, B.2, B.3, B.4 and B.5 in the Appendix and also in table 6.5, we used the exact same ways to set  $N_{HB}^{**}$ ,  $N_{HB}^*$ , and the same quantile probabilities ( $k$ ) to compare

percentage of coverage					
HB= 14	simulation	equation	HB= 34	simulation	equation
$r = 1$	0.05	0.05	$r = 1$	0.05	0.04
$r = 5$	0.05	0.09	$r = 5$	0.05	0.16
$r = 10$	0.05	0.10	$r = 10$	0.05	0.17

Table 6.5: Percentage of coverage of the rejection region using the total number of simulations (3,110,000), where  $r$  is the scaling factor used. “simulation” means the cut-offs were based on the actual simulated cases, while “equation” means that the cut-offs were calculated based on mean and variance equations 5.4.7 and 5.4.10.

the cut-offs using simulated cases to the cut-offs from the normal distribution by using the mean of  $WCR$  in equation 5.4.7 and the variance in equation 5.4.10. These show both the small case of HBs= 14 and the big case of HBs=34, and we can see that is not accurate in general to assume  $\mu$  and variance from equations 5.4.7 and 5.4.10 to calculate the cut-off in the value of  $WCR$  for each  $N_{HB}$  (see table 6.5). This is because of the large variations in the variance when the baseline of simulation ( $\theta$ ) changes as shown in table 6.4. So we will test the performance of the system using the  $WCR$  based on doing many simulations with each value of the number of HBs.

## 6.5 Conclusion

Now at the end of this chapter, we make some comments and suggestions which are very important, as below:

- In section 6.1, we used a general attempt to set the start and the end of the area of interest in  $N_{HB}$  (defining  $N_{HB}^{**}$  and  $N_{HB}^*$ ) which we expected to work in general circumstances. This works only in specific circumstances, so we moved to a more general attempt.
- In section 6.2 and 6.3, we show how to set an automatic function to calculate the quantile probabilities at each  $N_{HB}$ , either using simulation or equations, by using 5 different approaches.
- Postulating an automatic way of defining the detection area under any circumstances only works when the scaling factor  $r = 1$  or is close to 1. We consider this issue in section 6.4 and suggest scaling the value of  $\theta$ , but as presented in table 6.5, this did not work well. This is likely to be because the variance of  $WCR$  will then increase.
- In section 6.4, we compared cut-offs in  $WCR$  using equations for the mean and variance of  $WCR$  (using command `qnorm` in R), with cut-offs from simulated cases

for each  $N_{HB}$  (using command `quantile` in `R`) and we did not find the cut-offs calculated from equations to be similar to the cut-offs calculated from simulated cases, so we will test further the performance of *WCR* using simulated cut-offs rather than using equations.

In chapter 7, we now set up different situations to be considered in the simulation study, concerning patterns of GP consultations for ILI over time and define some system performance measures which we will use to compare our *WCR* system with other detection systems.

## Chapter 7

# Simulation study

In the previous chapter 6 in section 6.4, we conclude that we cannot use equations for the mean and variance of  $WCR$  (assumed to be Normal  $(\mu, \sigma^2)$ ) to set cut-offs for the  $WCR$  system, so we will use simulation to calculate quantiles of  $WCR$  in each  $N_{HB}$ .

In this seventh chapter, we want to simulate a new environment to reflect the real data which we already have. The aim from this simulation is to create different cases in a virtual world that simulate what is actually likely to happen. This will enable us to evaluate how the detection systems will perform in reality. As we already have over 5 years of historical data from August 2009 to the end of July 2015, we aim to simulate many cases from those years.

Various other studies have used simulation to study such detection systems. One study about measures against transmission of pandemic H1N1 influenza in Japan in 2009 simulated the expansion of pandemic H1N1 influenza in a realistic society called the virtual Chuo Line which models an area to the west of metropolitan Tokyo. It estimated the coefficient of transmission from outbreaks of pandemic H1N1 influenza among school children. Isolation at home, school closure, post-exposure prophylaxis and mass vaccinations of school children are included as evaluation measures in this simulation study, and it concluded that isolation at home and school closure together with post-exposure prophylaxis significantly reduce the total number of cases in society and can alleviate the expansion of pandemic H1N1 influenza, even when there is a delay in the availability of vaccine [191].

Another simulation study in the United States used a large-scale stochastic simulation model to probe the dispersal of a pandemic strain of influenza virus and recommended that in a highly mobile population, restricting travel after an outbreak could delay a little the time course of the outbreak without affecting the final number of cases [54]. Another

study simulated time series data to study the impact of air pollution on health using a Generalized Additive Model (GAM) based on historical data [43].

From the historical data in many studies, the trend for influenza usually increases in the winter, as in the United States between 1959 to 1999 [133] and as monitored in England and Wales in 1975 to 1990 [111], and also in Norway [78]. Data from Australia, France and the United States from 1972 to 1997 also confirm the trends in the winter [163]. So, we can see the increasing trend in winter, but this is not enough because the influenza rates may have different shapes of increasing trends, which we should consider. When we see the historical data from 2009 to 2015 in Scotland, we find that consultation rates tend to be seasonal, but a parametric model with sine and cosine terms does not fully capture the seasonal pattern (as presented in section 2.4), and so GLMs with sine and cosine terms to model the seasonal component are not the best solution. For this reason GAM models for data are more flexible than linear regression as GAMs can model non-linear trends more easily. Therefore we focus on GAMs in our simulation study.

## 7.1 Plan of the simulation study

As we cannot guarantee what will happen in reality, we can use what we can imagine from the historical shapes of a trend in cases, and simulate many cases from these, so this can cover real cases in the future in practice. The main and major target from the simulation study is to assess which system will work with high accuracy to raise the alarm in winter by applying the early detection systems with real data. We will consider some early detection systems to test them with simulated cases, i.e. the following methods:

- Weekly Cases Ratio (WCR) Method (section 3.6).
- Cumulative Summation (CUSUM) Method (section 3.4).
- Moving Epidemic Method (MEM) (section 3.3).

We have actual ILI cases from all of Scotland and also the aggregated 14 health boards (HBs) for weekly consultations in all GPs from the middle of August 2009 to the middle of August 2015.

Now, we are considering many scenarios in our simulation because the real data which we already have may or may not represent different patterns of data which may occur. Also, with the real data we cannot exactly know when the winter increase in rates begins, as different methods detect the start of an epidemic at different times. Another reason for our simulation is that the real data may have trends which we may not take into account, so we can build in a factor to the data by doing the simulation, which means there will

be a fixed pattern in which we know when the winter peak will happen and we can assess the accuracy of the different methods in detecting this. Also in the simulation we can consider different shapes and curves which may occur (for example, two peaks, a very sharp triangle, and flat increasing and decreasing shapes in the winter), as we want to know if the shape influences the results of the winter peak.

The first scenario considered using the historical data for Scotland from 2009/10 to 2013/14 to detect the alarm in winter 2014/15. This scenario was chosen as it already happened in the past, so it is realistic to apply it in this simulation study. It is possible also to have a repeated scenario every year, so we decided to pick one of the past years and repeat it 5 times, which is assumed as the second scenario in our simulation study. The third and fourth scenarios are considering increasing and decreasing trends which may happen also. We pick one of the past years and repeat it 5 times with an increasing or decreasing factor. One of our aims for using these two shapes is to see which system will be affected when we test them. As we are dealing with real consultations here, we chose a GAM to fit the pattern every year.

As we know from the data that influenza shows a very rapid fluctuation in the rises and falls, the GAM again is not sufficient to cover all patterns because the GAM is not symmetric and so can have a steeper increase than decline. In this case we go back to sine and cosine terms as they are symmetric about the peak, and these are used in many published papers to simulate influenza cases [30, 44, 73, 141], which we call the Serfling model.

The fifth, sixth and seventh scenario are using this Serfling model, but changing its parameters to give different shapes.

We still need to think about an outbreak case (like a sudden jump) which is not covered in the above scenarios, so the eighth, ninth and tenth scenarios will be based on a linear model which can cover many cases.

As MEM requires 5 years of data to apply it, we will use 5 years data to detect the following year.

Now, we will explain each scenario below:

## 7.2 Different Shape Cases

In this section, we want to try different shapes in our simulation as we cannot guarantee what will happen in the future. We choose different shapes and patterns from the historical ILI data which we already have, so we can list various scenarios below:

### 7.2.1 Simulate from a GAM from all 5 years historical data

Here, we used the GAM to fit what happened in the previous five years (2009/10-2013/14), then applied the detection system to test the early alarm using the simulated cases. We started by simulating one time series from a Poisson distribution with mean equal to the predicted values ( $\hat{\mu}_t$ ) from the following formula:

```
gam(round(ILI.Corr,0)) ~ s(week,bs="cc") + factor(year)
+ offset(log(Pop)), family=poisson, data=result)
```

where `ILI.Corr` gives the aggregated weekly cases of ILI in week  $t$ , `s(week,bs="cc")` gives the `spline` function [26,62] of `gam`, `Pop` represents the Scottish population in week  $t$ , and `result` is the data frame containing all the historical ILI data. `Year` is used as a factor rather than using it as linear because we are dealing with influenza cases which have increases and decreases over time.

To imagine situations that mimic the reality that we have, we need a model to describe the current reality and then extrapolate to predict what could happen. So we used a GAM, already described in the second chapter 2, section 2.5.

A simulation of weekly cases from this GAM from the five previous years (historical data) is presented in figure 7.1, as an illustration.

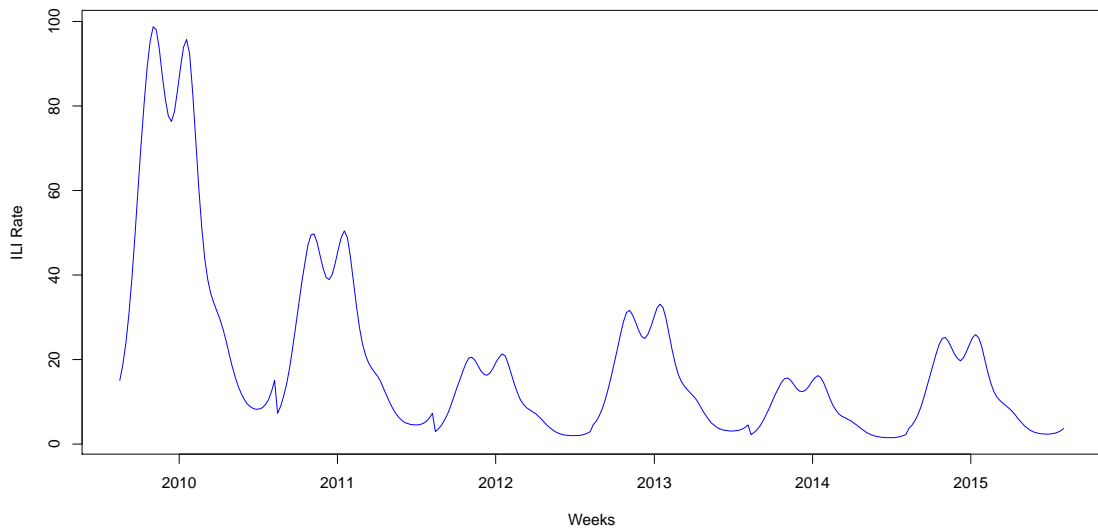


Figure 7.1: Simulated weekly cases from the historical ILI data (2009-2015).

Now, we present the results of how MEM and CUSUM work based on the simulated cases from the historical data. We used the winters of 2009/10 to 2013/14 to test the detection in winter 2014/15. Figure 7.2 shows (in the top plots) how MEM and CUSUM methods work with the ILI data and the lower plots present the results for the ARI data, and show when the alarm is raised. MEM clearly did not work well with the ILI data as the epidemic started without an alarm. When we investigated the reason for this, we found that the ILI rate in the first week was 16.64 per 100,000 population and the trend was increasing until week 6 before going to another peak in week 15. Also MEM delayed 3 weeks in the ARI data, and the reason for this was that we considered the 2009-2010 outbreak in the system and this raised the pre-epidemic threshold.

CUSUM raises the alarm in week 2 for the ILI data while its alarm is raised in week 8 for the ARI data. Both of these are faster than MEM but we should say we cannot guarantee whether they are false alarms or not. We will cover this in the next chapter.



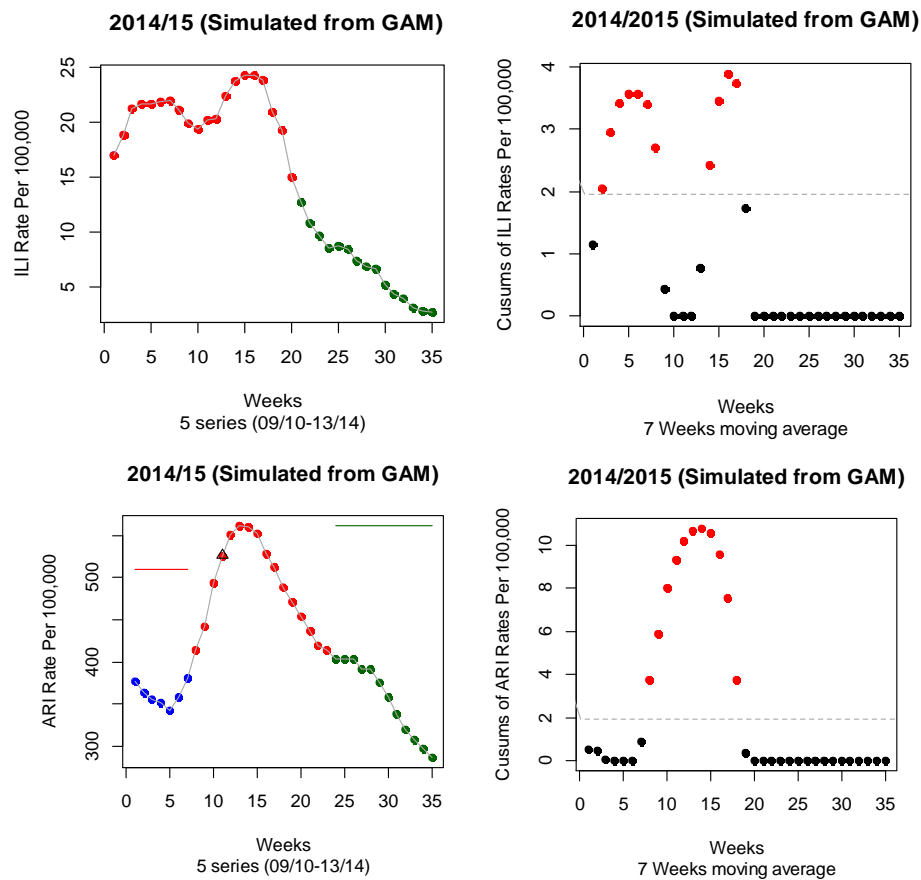


Figure 7.2: MEM and CUSUM methods' performance with simulated cases from the historical ILI and ARI data; MEM plots (left) include the pre-epidemic period (blue), epidemic period (red), and post-epidemic (green), while the triangle gives the alert week and the red line presents the pre-epidemic threshold, while the post-epidemic threshold is presented by the green line. A 5 year series (2009/10 to 2013/14) is used in MEM. CUSUM plots (right) include the normal weeks (black) and the epidemic weeks (red), while the dotted line presents the threshold based on the fact that 95% of the area of a normal distribution is within 1.96 standard deviations of the mean.

From this example, we can see the effect of including the outbreaks from the historical data on system performance. Outbreaks directly affect any system using long historical data as a basis of the system.

### 7.2.2 Simulate from a GAM from only the last year

In this section, we are doing the same fitting of a GAM for 2013/14 and repeating the fitted GAM 4 times to have 5 similar series, then simulating one time series from a Poisson

distribution with the mean equal to the predicted values ( $\hat{\mu}_t$ ) from the previous GAM in the formula in section 7.2.1. Simulating weekly cases from the GAM for the previous year (2013/2014) and repeating this five times is shown in figure 7.3. This gives a repeating trend.

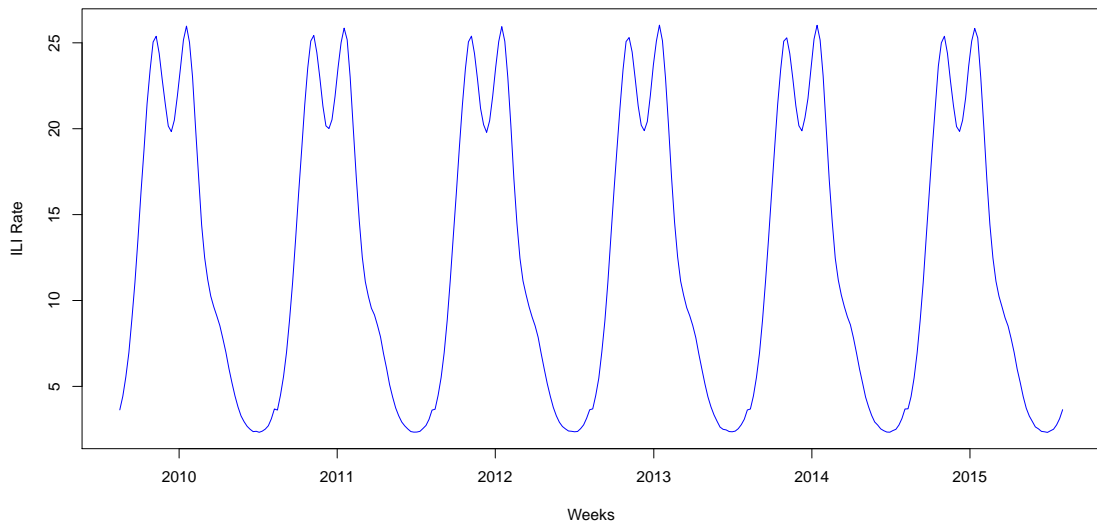


Figure 7.3: Simulated weekly cases from the historical data (repeating 2013/14 scenario 5 times).

### 7.2.3 Repeating with an increasing trend

Other possible cases to face in reality are increasing and decreasing trends. We will consider these two shapes to test which system will be affected more in those scenarios. We did not expect an effect in the *WCR* system as this system did not use much data from the past. Here, we are using a similar scenario to the one in section 7.2.2, but we are using an increasing trend. The simple way to do this is by choosing one of the years which have low rates (say 2014/2015) and repeating it 6 times with an increasing factor each year (we can use any number as the main idea is just to get an increasing trend; we used 50 counts cases here (which means that between 2009/10 and 2013/14, the ILI case rate per 100,000 increased by approximately 50). Then we simulate one time series from a Poisson distribution with the mean equal to the predicted values ( $\hat{\mu}_t$ ) from the previous GAM in the formula in section 7.2.1. Figure 7.4 illustrates this scenario.

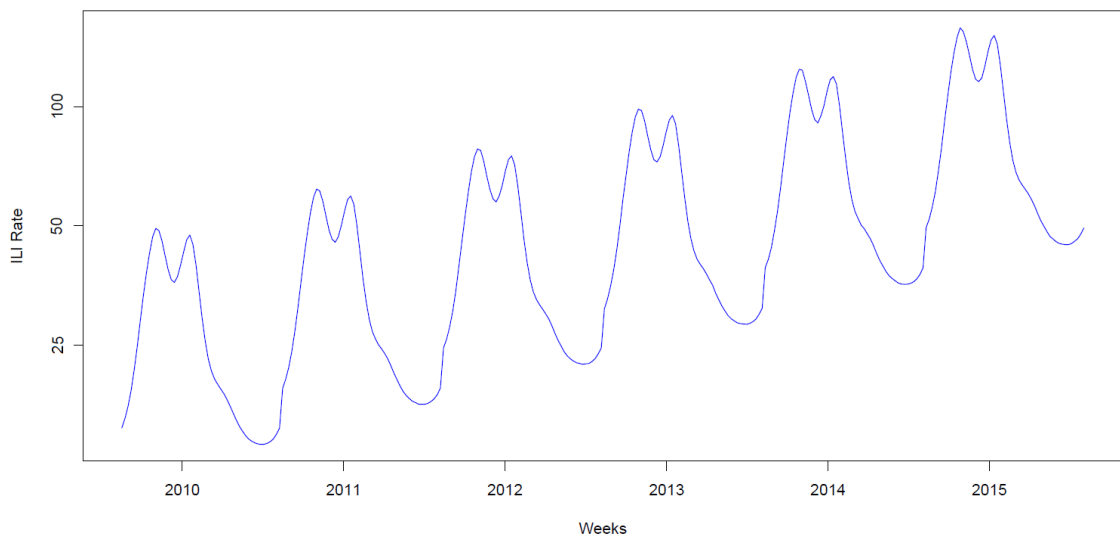


Figure 7.4: Simulated weekly cases from the historical data (repeating 2013/14 scenario 5 times with an increasing trend).

#### 7.2.4 Repeating with a decreasing trend

In this section, we are doing the same fitting of a GAM for 2013/14 and repeating the fitted GAM 4 times similarly to the previous situation in section 7.2.3 but with a decreasing trend, to give 5 scenarios. Again the simple way to do this is by choosing one of the years which have high rates (say 2009/2010) and repeating it 6 times with a decreasing factor each year (again we can use any number as the main idea is just to get a decreasing trend; we used 50 counts cases here (which means that between 2009/10 and 2013/14, the ILI case rate per 100,000 decreased by approximately 50), then simulating one time series from a Poisson distribution with the mean equal to the predicted values ( $\hat{\mu}_t$ ) from the previous GAM formula in section 7.2.1. Figure 7.5 illustrates this scenario.

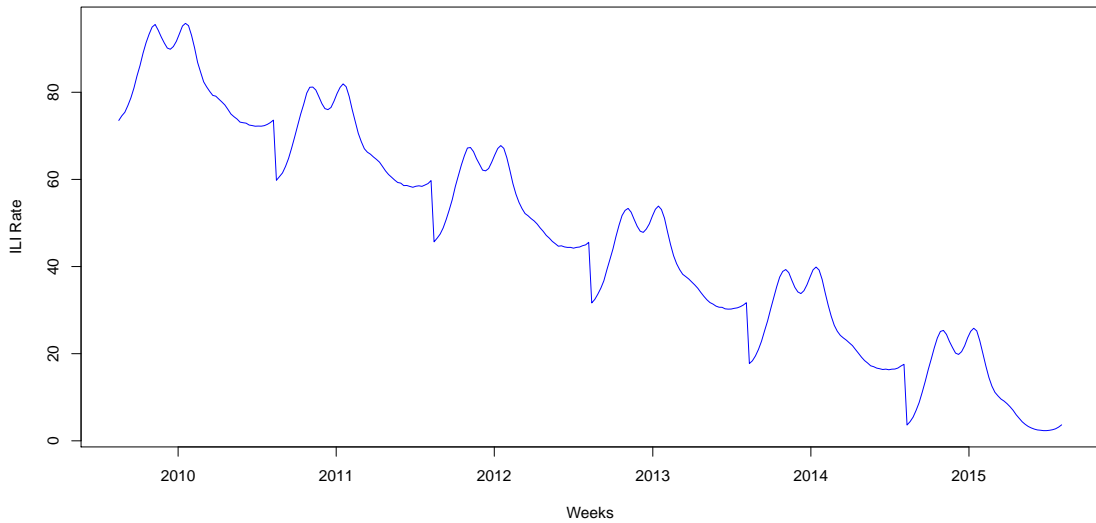


Figure 7.5: Simulated weekly cases from the historical data (repeating the 2009/10 scenario 5 times with a decreasing trend. There is a discrete jump at the end of each year because we selected year 2009/10 and this year has an end rate higher than its start rate (as the idea is to get a decreasing trend).

### 7.2.5 Serfling model

One of the old established methods to simulate influenza cases is the Serfling method. The Center for Disease Control and Prevention (CDC) have used a cyclic regression model to describe the starting of influenza activity (epidemic) and mortality caused by influenza in their influenza surveillance programmes since the mid-1960s [113]. This method for determining the epidemic threshold bears the name of Serfling as he did the original model formulation [113]. The Serfling method uses five years data of weekly influenza cases to fit a cyclic regression model including terms for an intercept, linear tendency, and a pair of symmetric terms to catch the following sinusoidal pattern of seasonal influenza. The Serfling model is formulated as a simple linear regression [113]:

$$y_t = \alpha + \beta_1 \sin\left(2\pi\left(\frac{t}{52}\right)\right) + \beta_2 \cos\left(2\pi\left(\frac{t}{52}\right)\right) \quad (7.2.1)$$

where  $y_t$  represents observed influenza cases in week  $t$  of the whole five year period, and  $\alpha$  and  $\beta_1, \beta_2$  are regression coefficients to be estimated. Because the model can estimate the epidemic and non-epidemic seasonal baseline, we can consider the weeks which are more likely to have influenza activities, which is the winter in Scotland, from September to April.

We consider here 3 different shapes by changing  $\beta_1$  and  $\beta_2$  in equation 7.2.1 to have a one

peak scenario, a two peaks scenario (a low peak then increasing trend) and another two peaks scenario (a high peak then a low peak).

- Using the Serfling model [128] to predict cases with one peak in the year is shown in figure 7.6.

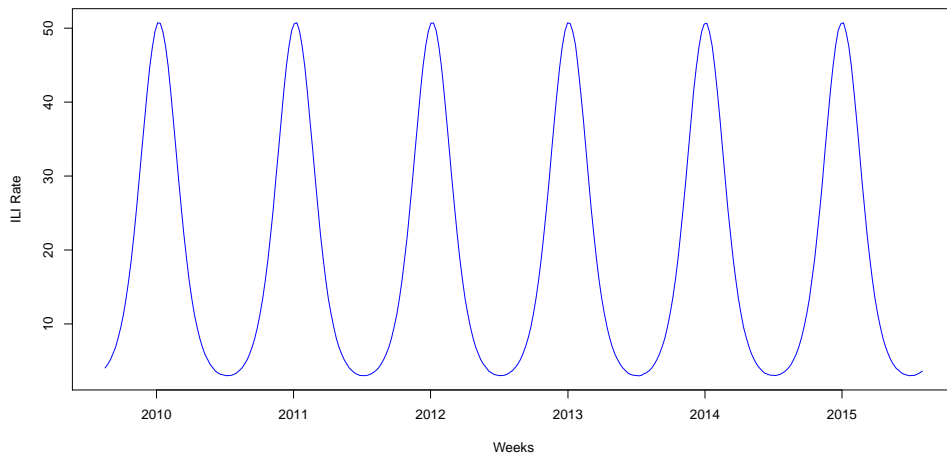


Figure 7.6: Simulated weekly cases from the Serfling model (single peak in winter),  $\alpha = -9, \beta_1 = 1, \beta_2 = 1$ .

- Using the Serfling model [128] to predict cases with two peaks in the year (one high, then one low) is shown in figure 7.7.

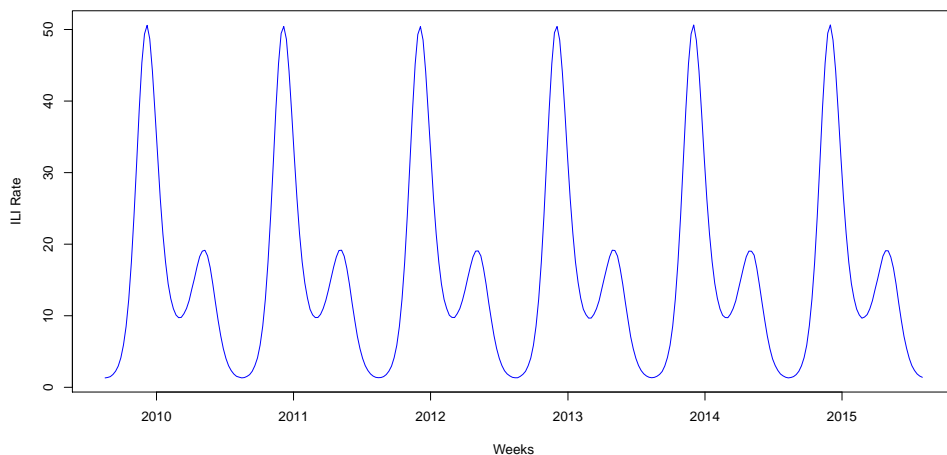


Figure 7.7: Simulated weekly cases from the Serfling model (double peaks with main peak in winter),  $\alpha = -9, \beta_1 = 1, \beta_2 = 0.5, \beta_3 = 0, \beta_4 = 1$ .

- Using the Serfling model [128] to predict cases with peak then increasing trend in

the year is shown in figure 7.8.

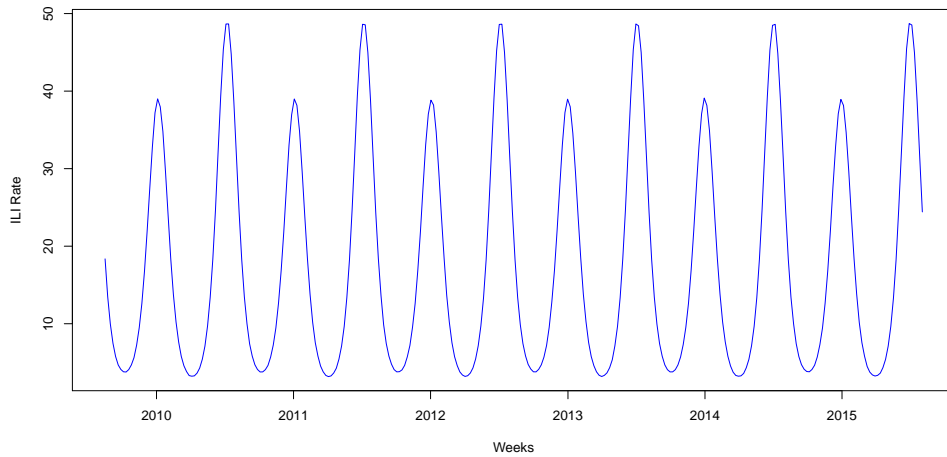


Figure 7.8: Simulated weekly cases from the Serfling model (peak then increasing trend),  $\alpha = -9, \beta_1 = -0.1, \beta_2 = -0.1, \beta_3 = 0.4, \beta_4 = 1.2$ .

### 7.2.6 Linear Model

Another scenario not covered by either the GAM or the Serfling model is when the number of influenza cases changes suddenly (increases or decreases), so we want to consider more sudden scenarios from linear models by using linear step functions rather than GAMs or Serfling models which have much smoother changes.

Here, we show how to model the influenza activities in the winter season. It is similar to the Serfling model but is a linear model which can include some different scenarios like sudden increasing or decreasing. The advantage from these scenarios is that they consider some sudden events which are not part of smooth scenarios. Serfling models and GAMs can consider only the smooth scenarios. The first shape here is one peak in the year with stability at a low level (say around the normal influenza activity), then sudden stability at a high level (say around the influenza season's activity or outbreak) before a sudden decrease to a low level (the normal influenza activity again) (case 1). Low rates reflect the summer season and high rates reflect the winter seasons.

The second scenario of a linear shape is one peak in the year with stability at low level then a gradual deviation to a peak level, then a gradual decrease to a low level (case 2). We can model that by the linear model:

$$Y_t[increasing] = \min Y + (\max Y - \min Y)/L_1 \quad (7.2.2)$$

$$Y_t[decreasing] = \max Y - (\max Y - \min Y)/L_2 \quad (7.2.3)$$

where  $Y_t$  represents the counts of cases in week  $t$  with increasing and decreasing shapes;  $\max Y$  and  $\min Y$  give the maximum and minimum counts of cases; and  $L_1$  and  $L_2$  are the length of the periods for increasing and decreasing weeks.

The last scenario considered from the linear shape is one peak in the year with stability at a low level, then a gradual deviation to a peak level, then a sudden decrease to a low level (case 3). This will use the increasing model 7.2.2 before dropping the number of cases suddenly to the low level.

- Using a linear model to predict cases with one peak in the year (case 1) is shown in figure 7.9.

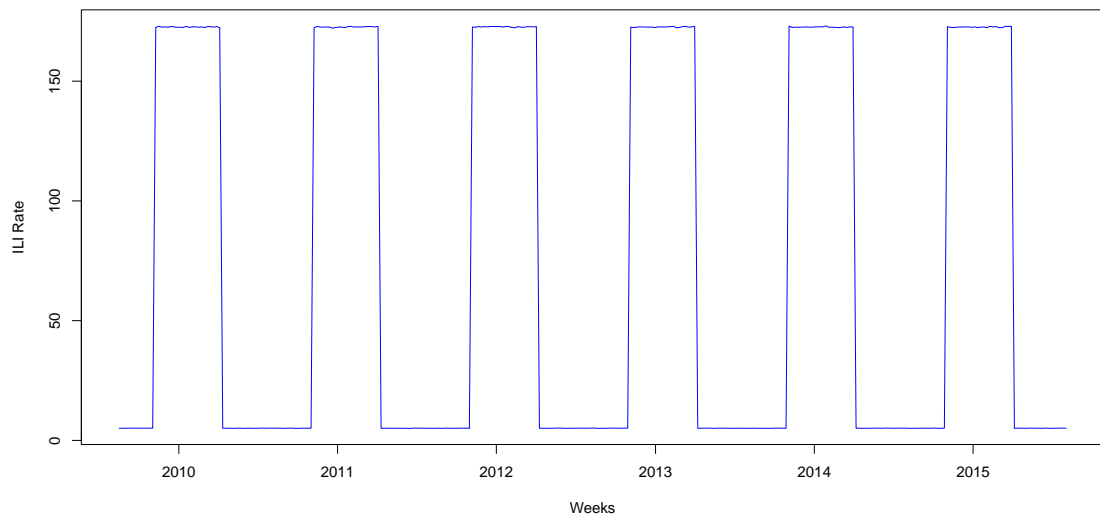


Figure 7.9: Simulated weekly cases from a linear shape (stability at a low level, then sudden stability at a high level, before going back suddenly to a low level).

- Using a linear model to predict cases with one peak in the year (case 2) is shown in figure 7.10.

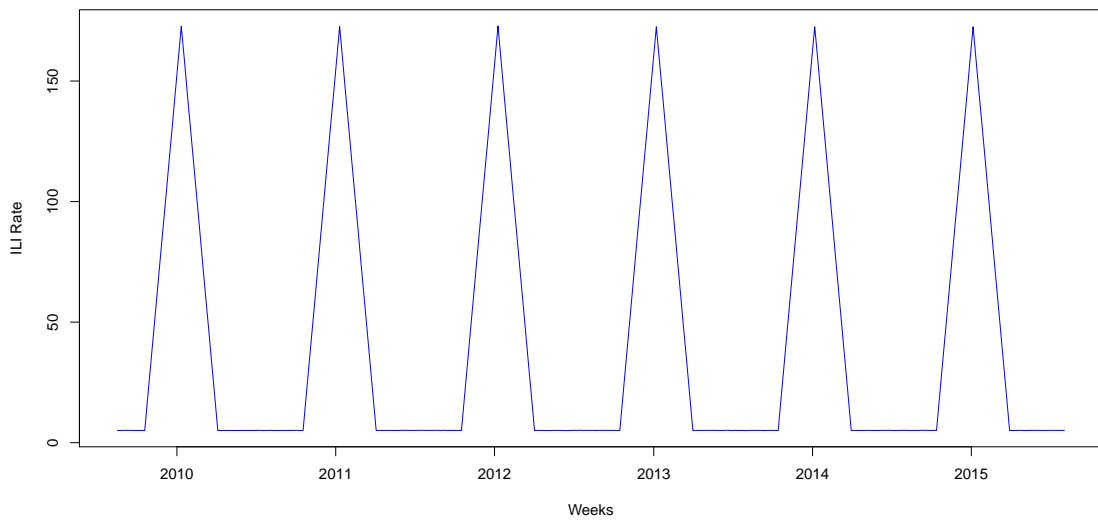


Figure 7.10: Simulated weekly cases from a linear shape (stability at a low level, then a gradual increase to a high level before going back gradually to a low level).

- Using a linear model to predict cases with one peak in the year (case 3) is shown in figure 7.11.

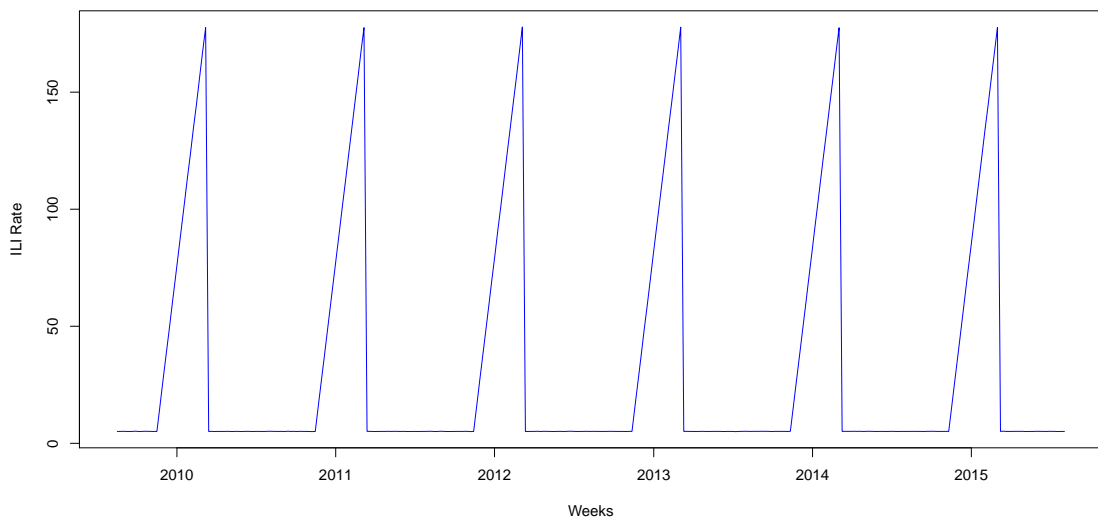


Figure 7.11: Simulated weekly cases from a linear shape (stability at a low level, then a gradual deviation to a peak level, then a sudden decrease to a low level).

There is one possible scenario we did not consider, which is suddenly increasing then gradually decreasing. The reason is that we care about increasing cases and this sudden



increasing scenario is already covered in case 1, although there the sudden increase is stable and is not followed by a gradual decrease (see figure 7.9).

### 7.3 Illustration

Here in this simulation study, we have illustrated what we get from simulating one case in different scenarios. For all of our simulations, the following points apply:

- We repeated the same scenario 5 times (5 years).
- We used the size of the Scottish population as the median population in the period 2009-2014., which is 5,400,000 people.
- We ignored weekends (meaning each week contains 5 days).
- The winter season is considered to last from the first of September to the end of April and the summer season is considered to run from May to the end of August.
- Each week during the year corresponds to the date as presented in table C.1 in the Appendix.

### 7.4 Performance measurements

Now, for the 10 different scenarios in section 7.2, we need to specify the week which is taken as the week when rates increase (which means we start to calculate sensitivity from this week). In the case where influenza activities are linear and they clearly jump in a specific week (the first case in section 7.2.6), we can easily set the week when the rates increase. In the other scenarios, we can set a percentage increase and consider the week when the rates exceed the baseline rate by this amount as the increase week (say, 10%).

We will start testing the sensitivity, specificity and the number of weeks between the time of the alarm and the time of the increase. Those three terms to test performance of the *WCR* system, *CUSUM* and *MEM* are defined below:

- Alert week (*AW*): This reflects the week judged to be the start of increasing rates of influenza (the “increase week”).
- Sensitivity (*SN*): Does the system detect an increase when it occurs (“correct alarm”; *CA*)? A good system should have a high *SN*. *SN* is the percentage of “increase weeks” in which an alarm is raised.
- Specificity (*SP*): Does the system detect an increase when it does not occur (“false alarm”; *FA*)? A good system should register a high *SP*. *SP* is the percentage of weeks when there is not an increase in which the system does not detect an increase.

- Median time to detection (*MED*): How many weeks are there between the alert week and the alarm week issued from the system? A good system should register a low positive *MED*. This was calculated here as the number of the alert week minus the number of the week when the alarm is raised. Negative *MED* means there are delays in the alarm compared to the alert week. Values of *MED* close to zero are desired.

### Example

Figure 7.12 presents an example of specifying the alert week, and calculating *SN*, *SP* and *MED*. The red curve represents the ILI rates, and the blue horizontal dotted line represents an ILI rate of 10% above the first week's rate. The green vertical dotted line represents the "increase week" or alert week *AW* (week 10 in this example). This starts when the ILI rate exceeds a level of 10% above the first week. From this week, we can start to consider any alarm as correct (the sensitivity region, the ILI rate of which is the light blue area). This will continue until the purple vertical dotted line which represents the point when the ILI rates dropped below this higher level. Any alarm before the green vertical dotted line week will be considered as a false alarm (the specificity region, which is the light yellow area). If the alarm is raised in week 10 (alarm number 2), the *MED* equals 0 and this is the best case. If the alarm is raised in week 3 (alarm number 1), the *MED* equals  $10-3=7$  weeks before *AW*. If the alarm is raised in week 18 (alarm number 3), the *MED* equals  $10-18=-8$  which mean there are 8 weeks delay in the alarm.

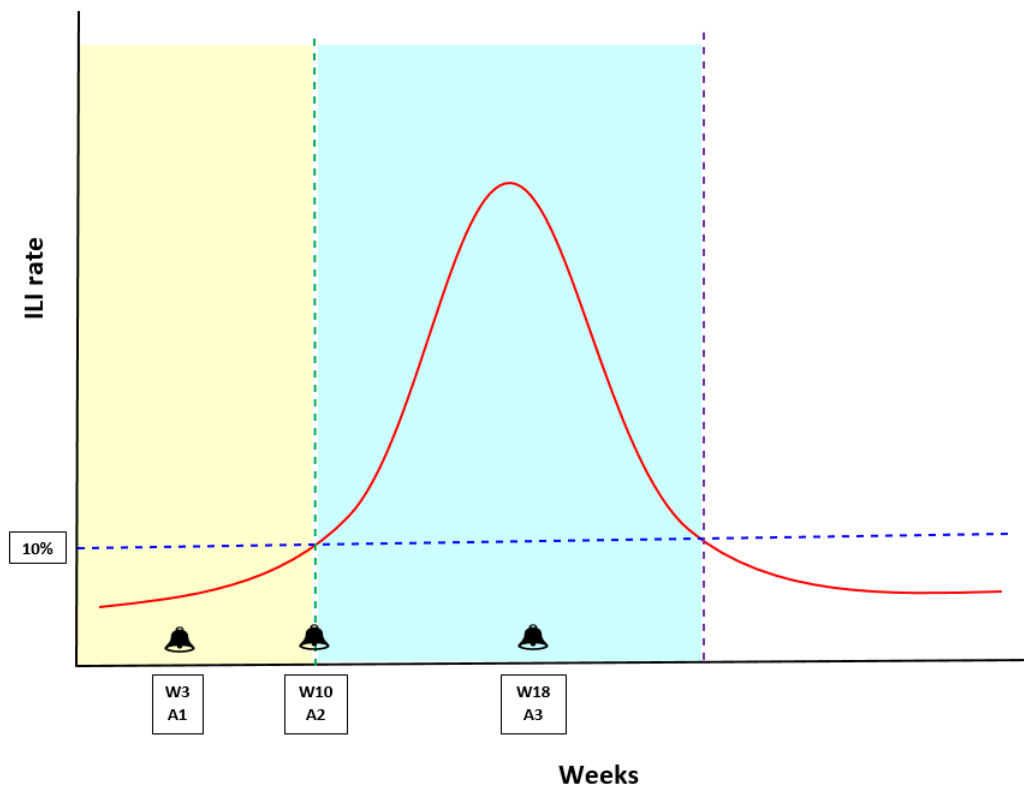


Figure 7.12: Example of specifying the start of rates increasing, and calculating  $SN$ ,  $SP$  and  $MED$ . More details are given in the text above. A1, A2 and A3 represent alarms 1, 2 and 3 respectively.

## 7.5 Conclusion

In this chapter, we presented 10 different shape cases of variation in rates, as we cannot guarantee what will happen in practice. We aim to test performance of  $WCR$ ,  $CUSUM$  and  $MEM$  under those situations of an increase in different circumstances. Some of those scenarios related to historical data and some of them were based on imagination of what might happen (stylised or extreme versions of reality). The mechanism used has been clarified, such as the size of population and the number of weeks and days during this simulation study.

In the next chapter, we will test  $WCR$  with other detection systems and compare their performance in terms of sensitivity, specificity and median time to detection.

## Chapter 8

# Testing *WCR* with other systems

## (1)

We concluded chapter 6 with a version of the *WCR* system. In this chapter, we want to test the performance of this *WCR* system and compare it with the CUSUM method (section 3.4) and MEM (section 3.3). During this testing, we will use different scenarios, listed in chapter 7. Also we will modify the *WCR* system if needed and explain the reason and the means of modification.

### 8.1 Simulation for different scenarios

In chapter 7, section 7.2, we assumed 10 scenarios of patterns of GP ILI consultation rates during the years, namely historical rates, repeating 2014 rates, repeating 2014 rates with an increasing trend, repeating 2014 rates with a decreasing trend, three shapes of linear increasing and decreasing rates and three shapes of Serfling model [133] rates.

We have also simulated 3,110,000 cases from the steady state situation in chapter 5 in section 5.2. One of the targets for this big simulation is to set up cut-off points when the rates are assumed to be a flat rate, which we discussed in the previous chapters. The purpose of this big simulation with steady states is to establish the boundaries of *WCR* and  $N_{HB}$  in lots of situations where there is no change in the rate. We can consider that we have a hypothesis test based on the two hypotheses below:

The null hypothesis  $H_0$ : the observation of  $(WCR, N_{HB})$  comes from the steady state, and the alternative hypothesis  $H_1$ : the observation did not come from the steady state. So, all the simulations from the steady state are used to set cut-offs for the hypothesis test.

In order to investigate the questions in section 5.1, we are going to simulate 100 cases

each week over 6 years for each different scenario, which will give the total of 31,100 simulated cases in each scenario with number of health boards ( $ng$ ) equal to 14 and 34 and using population structure  $P_2$  in table 5.1 (using an equal split which divides the total population by the number of regions).

The aim of doing such simulations to have 10 reasonable situations which can happen (idealised versions of reality). Also the reason for this simulation is to find out where this method will detect a jump in the rates with those scenarios which look reasonable. If we have data which are not from the steady state, we expect extreme values of  $(WCR, N_{HB})$  to occur more than 5% of the time. For example, figure 8.1 represents  $WCR$  against  $N_{HB}$  for historical rates with 14 HBs and a 5 million population. In the steady state, the light blue shaped area covers 5% of the total number of simulations but in the current scenario we expect it will cover more than 5% (it covers around 36%), as this scenario came from historical rates.

In other words, when we see the distribution of  $WCR$  and  $N_{HB}$  in figure 8.1, we find many observations above the cut-offs, more than we expected based on having a steady state. We have many more observations in the rejection region derived from the steady state than we got with the steady state, so we conclude that the rates are not following the steady state.

Since we now have a statistical test of a huge number of situations with constant rates (3,110,000 cases), we can apply it to the different scenarios from section 7.2, so that the increase in simulated cases should not be in the region of the steady state (i.e. lies in the region corresponding to  $H_1$ ) to get a Correct Detection (CD) rate (sensitivity) in the cases of increasing rates. On the other hand, if we do not have an increase in cases (depending on the scenario specified), we expect to be in the region of the steady state ( $H_0$ ). If that happens and we observe the region representing  $H_1$ , it can be considered as a False Alarm (FA).

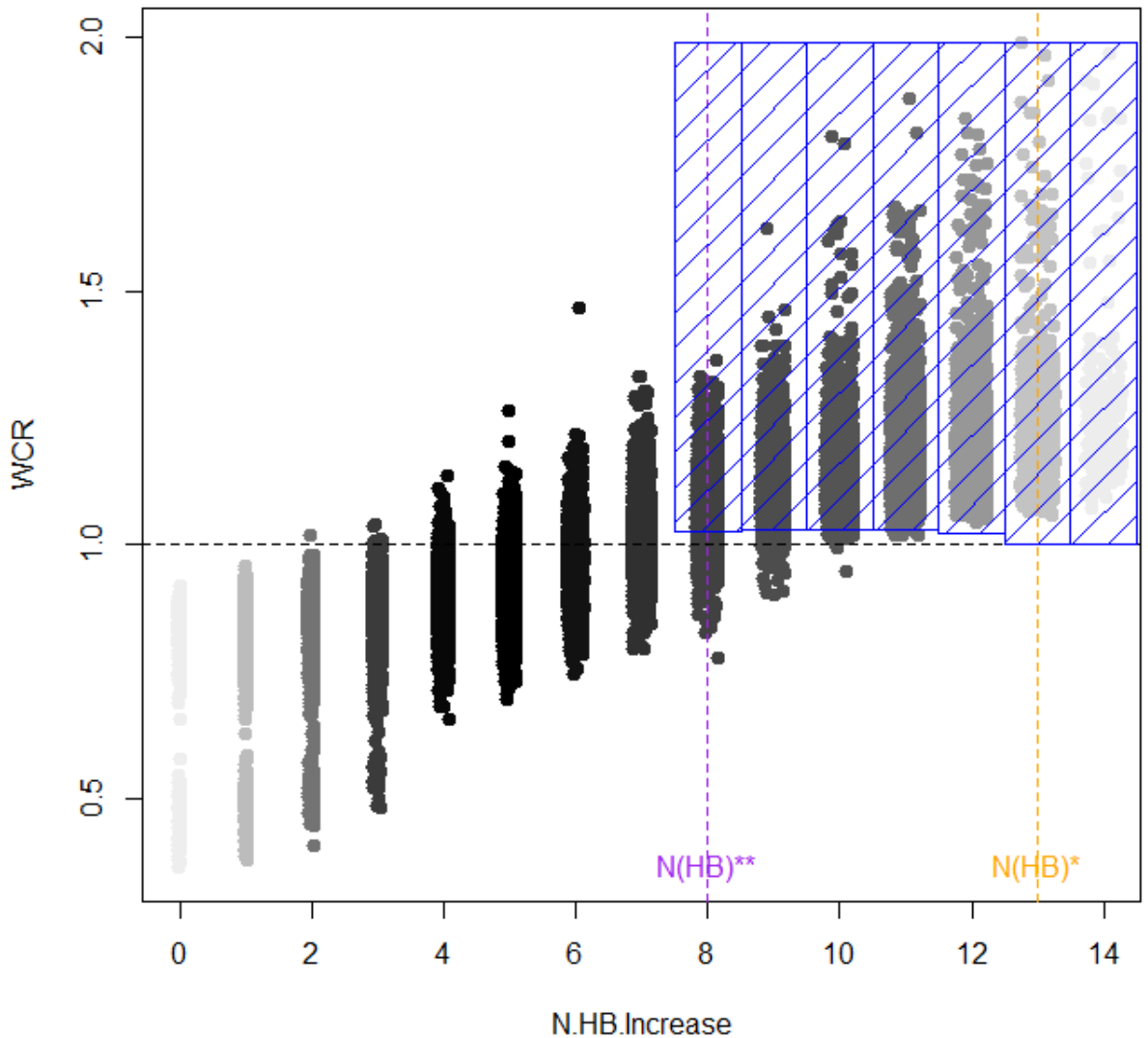


Figure 8.1: *WCR* against  $N_{HB}$  for historical rates with 14 HBs and a 5 million population.  $N_{HB}^*$  is presented in the dotted vertical orange line while the dotted vertical purple line represents  $N_{HB}^{**}$ . The blue shaped area columns came from calculating the quantile using  $\mu$  and  $\sigma$  from the flat rate. They represent the 5% most extreme observations under the steady state. Darkness of the black colour reflects the number of simulated cases.

## 8.2 Testing the *WCR* system (1)

In this section, we are going to test *WCR* as in section 6.4 against CUSUM and MEM with 10 realistic situations presented in section 7.2. We found that the *WCR* system raised the alarm at a very early stage compared to CUSUM and MEM. In this case, we

should not consider those alarms as an advantage of the *WCR* system because there is a possibility for those alarms to be false. To test this point, we will choose linear scenarios because we can then know the exact week of increase in rates.

Figure 8.2 shows a comparison in the week when the alarm is raised between *WCR* (blue triangles), CUSUM (green dots) and MEM (red crosses) in the top figure, and an overview of the scenario used in this test is presented in the bottom figure. This figure represents the data set of  $ng=14$ . Week 10 is the first week when the rate jumps. Weeks 1 to 9 should be considered as steady rates because they were simulated from the steady state, so any alarms before week 10 are false alarms. We can see clearly from the graph that the majority of false alarms are for the *WCR* system (92 false alarms) compared to the systems CUSUM and MEM, using the 100 simulated cases. CUSUM has relatively few false alarms (18 false alarms). MEM registered 40 false alarms. In the *WCR* system, we see that alarms are frequent and unpredictable and not as in the two systems CUSUM and MEM. When the alarm system fails in a situation of stable rates (before a sudden or gradual increase occurs), it is a sign that the system will fail when applied to situations derived from reality, when there is an increase.

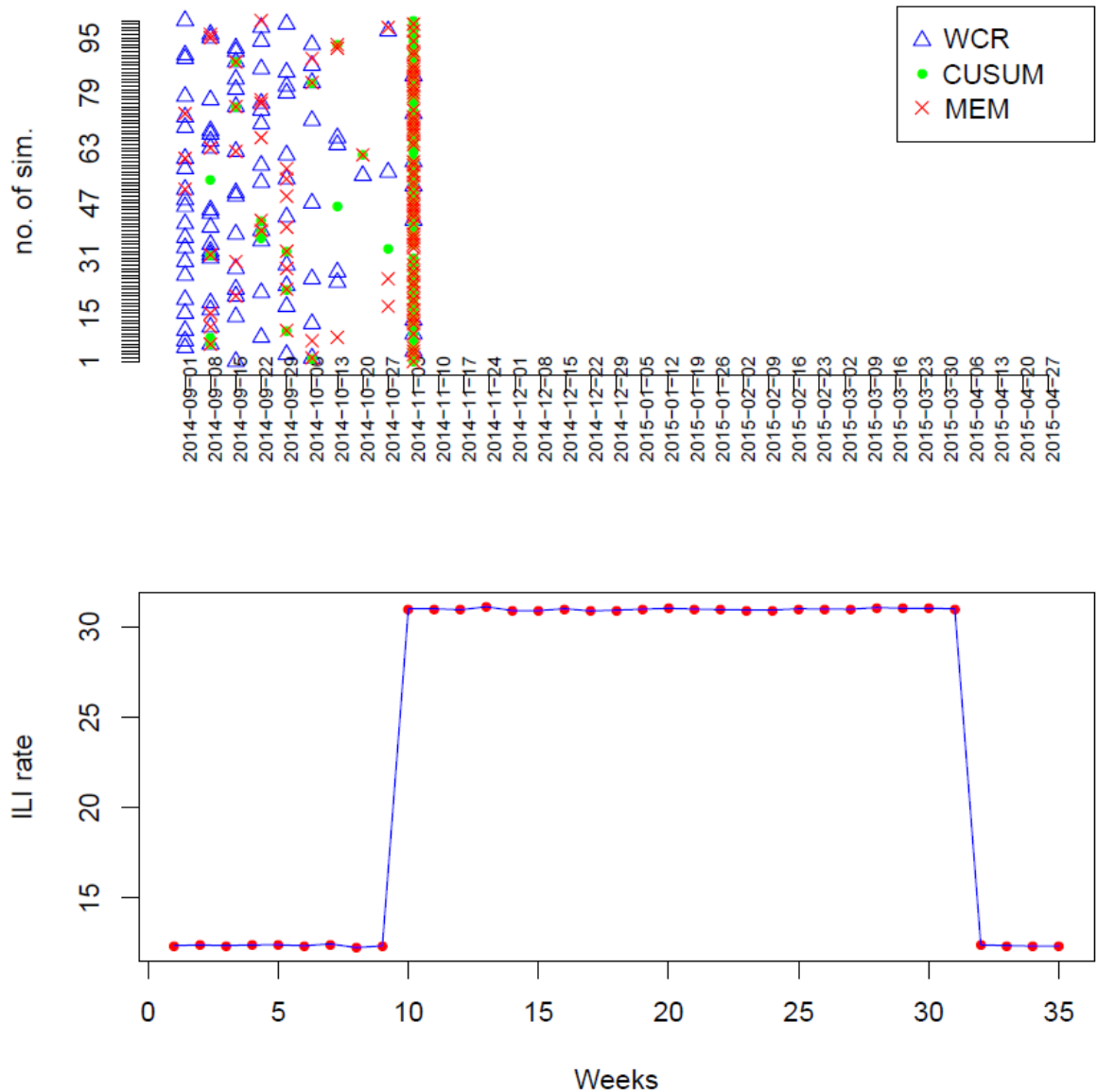


Figure 8.2: The top figure shows a comparison of detection of alert weeks between *WCR* (blue triangles), CUSUM (green dots) and MEM (red crosses). The bottom figure shows the scenario used to clarify the location of the jump in rates; it shows the mean of the 100 simulated cases.

When we investigate the reason for this weakness in the *WCR* system (too many false alarms), we found that it is very common to have a number of geographic regions  $N_{HB}$  in the mid-region ( $\frac{N_{HB}}{2} + 1$ ). We got the same result for either  $ng = 14$  or  $34$ . Now, we will move to see the result when we have another scenario (repeating the 2014 data six times). The difference in this scenario is that we used a realistic reference from 2014 where the rise is gradual and not surprising, and figure 8.3 represents the comparison between the three systems at the time of raising the alarm during the weeks. We are not considering



any one week as an alert week in this figure (so we do not quote numbers of false alarms). *WCR* shows very early detection because the increase in week 2 leads to a significant result (rejecting  $H_0$  in the hypothesis test in section 8.1) which means it is decided that it is not from the steady state.

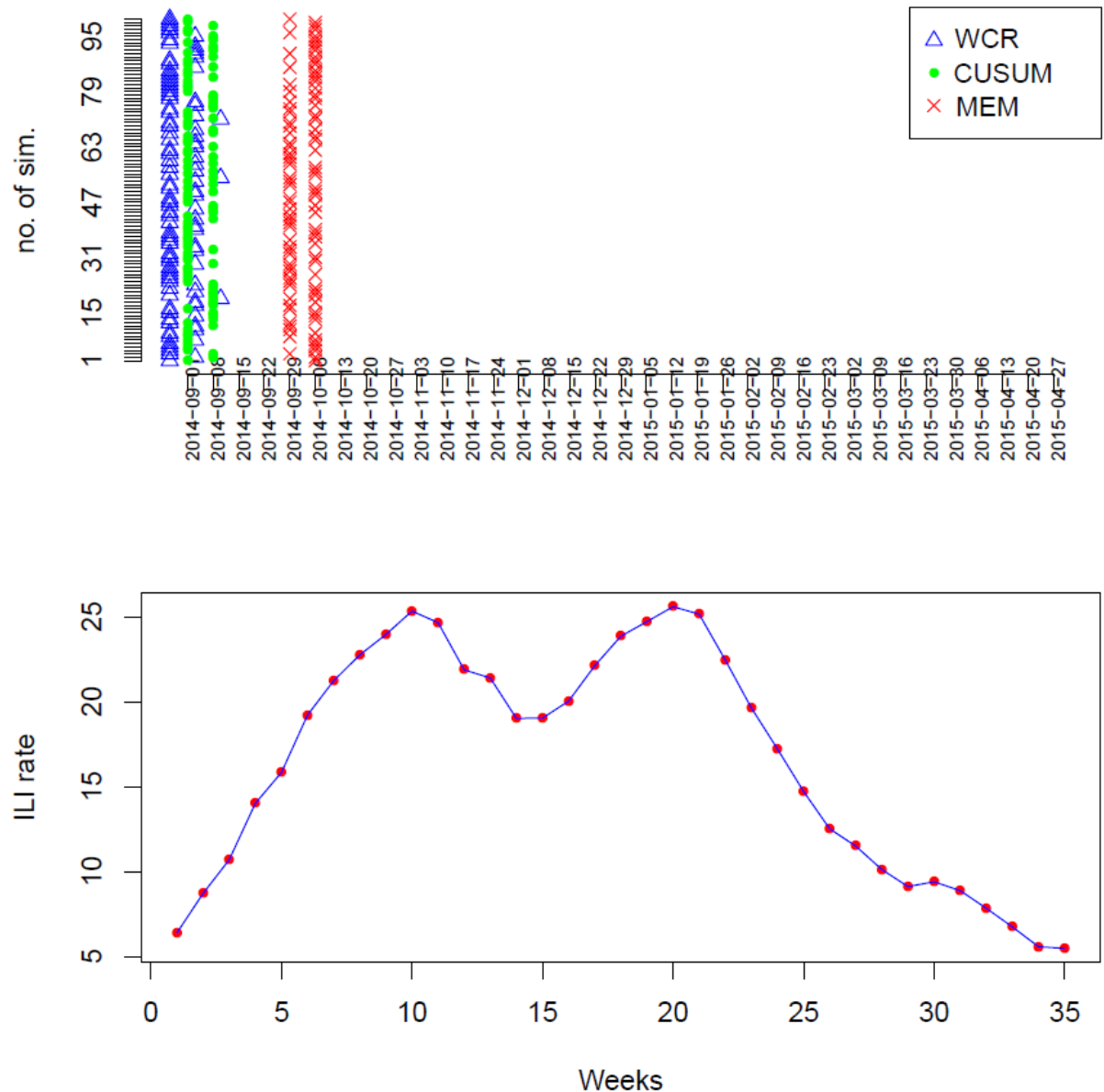


Figure 8.3: The top figure shows a comparison of detection of alert weeks between *WCR* (blue triangles), *CUSUM* (green dots) and *MEM* (red crosses). The bottom figure shows the scenario used to clarify the location of the increase in rates; it shows one of the 100 simulated cases.

Because of getting the alarm in very early weeks in figure 8.3 or raising an alarm while the simulation is from steady rates (figure 8.2), we will try to modify the system so that

it starts after the midpoint of the number of geographical regions ( $ng$ ).

In order for the system to be general for all  $ng$  (no matter how different the number of  $N_{HB}$ ), we will choose the beginning of the system ( $N_{HB}^{**}$ ) to be when the cumulative distribution of the Binomial function is close to 80% and the ending of the system ( $N_{HB}^*$ ) when the cumulative distribution of the Binomial function is close to 99.99%, presented in figure 8.4. We will name this setting  $WCR_{v1}$ , meaning version 1.

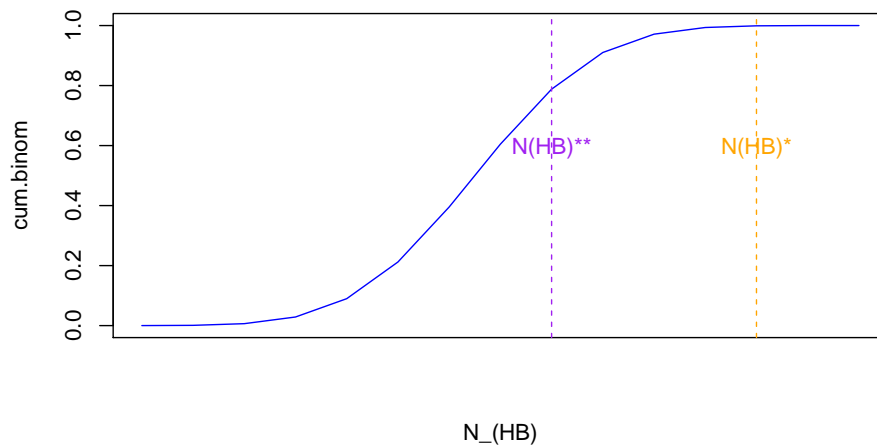


Figure 8.4: The beginning of the system ( $N_{HB}^{**}$ ) when the cumulative distribution of the Binomial function is close to 80%, and the ending of the system ( $N_{HB}^*$ ) when the cumulative distribution of the Binomial function is close to 99.99%.

Now, in this current setting, while the aim is to raise the alarm in one specific week, *WCR* raises many alarms in more than one week, as shown in figure 8.3. From this point we will start comparing the three systems depending on the sensitivity and specificity of the system, and the number of weeks from the week of the alarm raised to the increase week. Those three measures to compare the three systems were defined in the last chapter 7, in section 7.4.

We will continue now to apply the system for all our assumed scenarios presented in section 7.2, and compare the performance between  $WCR_{v1}$ , CUSUM and MEM. We can set the epidemic week (alert week or *AW*) by choosing a specific rate of increase in that week compared to the first week, so all those weeks from week 1 to this specific week of increase will be considered for calculation of specificity, and those weeks after that will be used for calculation of sensitivity. In this approach, we will choose the alert week (week of increase) as the first week when the rate is at least 10% higher than in week 1.

Figure 8.5 presents a comparison of detection of the alert week of  $WCR_{v1}$  (blue triangles),

and CUSUM (green dots). The bottom figure shows the scenario used to clarify the location of the increase in rates (this scenario is based on historical data from 2009 to 2014 and the target season is the winter of 2014-2015 (Sep 2014 - April 2015)). MEM did not detect any alarms. The reason for missing the alarms with MEM is because MEM considers the pre-epidemic threshold as the average rate in the pre-epidemic periods in the past years, and we know from the actual data that there was an outbreak in winter 2009, which causes an increase in the average rates in the pre-epidemic periods and leads in the end to missing the alarms (the *WCR* system raised 66 false alarms, while the CUSUM system raised 14 false alarms).

Table 8.1 shows a comparison between these three systems when using version 1 of *WCR* in section 8.2 ( $WCR_{v1}$ ) and again using 100 simulations. We see that performance is quite similar for both 14 and 34 HBs, and we can give some comments:

- In the historical scenario, CUSUM performance is the best. MEM failed here because its warnings depend on the average rates in the past 5 years. There was an outbreak in 2009, which caused a rise in the average rate of the past years and thus an absence of alarms before the epidemic start.
- In the scenario of repeating 2014 cases six times, we got similar results to the historical scenario with *WCR* but not as good with CUSUM, while MEM showed high sensitivity and specificity with a delay in its alarm.
- In the increasing scenario, the best performance was with the CUSUM system, while *WCR* and MEM systems do not show good sensitivity and specificity.
- In the decreasing scenario, *WCR* and CUSUM do not show perfect detection as the sensitivity and specificity are low with *WCR* and high with CUSUM but with a delay for CUSUM.
- For the shape of the first linear scenario, CUSUM performance is the best, then MEM comes second, while *WCR* is the worst as its sensitivity is very low.
- For the shape of the second and third linear scenarios, and the shape of the first Serfling scenario, CUSUM performance is the best overall in terms of all 3 measures. MEM has better sensitivity and specificity but has a delay in its alarm, while *WCR* is the worst as its sensitivity is very low.
- For the shape of the second Serfling scenario, MEM shows high sensitivity and specificity (but with a delay in alarms), as does CUSUM with 14 HBs but not 34 HBs, while *WCR* shows the worst performance and detects a little early.
- For the shape of the third Serfling scenario, *WCR* and CUSUM show low sensitivity and high specificity with 14 HBs. MEM again here shows high sensitivity and speci-

ficity but a delay in its alarm. With 34 HBs, *WCR* and MEM have low sensitivity. MEM has very poor time to detection with 34 HBs in this scenario.

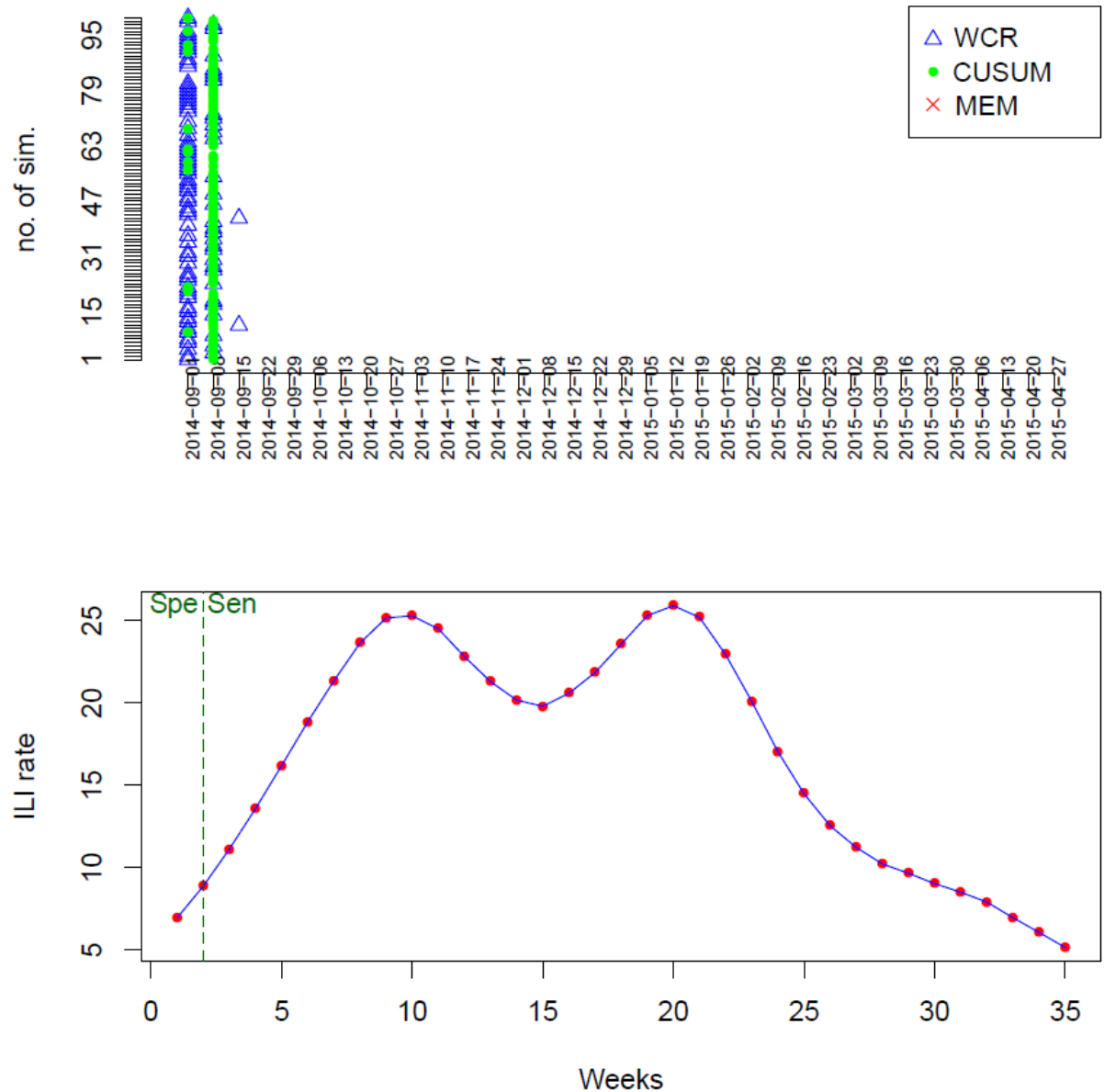


Figure 8.5: Comparison of detection of alert weeks between  $WCR_{v1}$  (blue triangles), CUSUM (green dots). The bottom figure shows the scenario used to clarify the location of the increase in rates (this scenario is based on historical data from 2009 to 2014 and the target season is the winter of 2014-2015 (Sep 2014 - April 2015)); it shows the mean of the 100 simulated cases.

$ng=14$									
	$WCR_{v1}$			CUSUM			MEM		
Kind of Shape	Sen.	Sp.	$MED$	Sen.	Sp.	$MED$	Sen.	Sp.	$MED$
Historical	0.04	0.04	0.96	0.94	0.94	0.05	<b>F</b>	<b>F</b>	<b>F</b>
Repeat 2014	0.05	0.05	0.95	0.62	0.62	0.38	1	1	-3.7
Increasing	0	0.5	1.9	0.94	0.94	0.05	0.01	0.51	1.05
Decreasing	0.02	0.02	0.98	1	1	-4	<b>F</b>	<b>F</b>	<b>F</b>
Linear 1	0.04	0.89	6.6	0.86	0.98	0.8	0.71	0.97	1.7
Linear 2	0.03	0.88	5.2	0.37	0.97	1.2	1	1	-1.7
Linear 3	0.02	0.91	8.6	0.84	0.99	0.92	1	1	-2.4
Serfling 1	0.21	0.21	0.79	0.63	0.63	0.32	1	1	-8.2
Serfling 2	0.46	0.46	0.45	1	1	-6	1	1	-6.2
Serfling 3	0	0.9	4	0	0.9	1.02	1	1	-1.2

$ng=34$									
	$WCR_{v1}$			CUSUM			MEM		
Kind of Shape	Sen.	Sp.	$MED$	Sen.	Sp.	$MED$	Sen.	Sp.	$MED$
Historical	0.02	0.02	0.98	0.96	0.96	0.04	<b>F</b>	<b>F</b>	<b>F</b>
Repeat 2014	0.02	0.02	0.98	0.56	0.56	0.44	1	1	-3.8
Increasing	0	0.67	2.7	1	1	-1	0	0.66	3
Decreasing	0.07	0.07	0.93	1	1	-4.2	<b>F</b>	<b>F</b>	<b>F</b>
Linear 1	0.07	0.9	6.02	0.85	0.98	0.7	0.69	0.96	2.03
Linear 2	0	0.88	5.03	0.33	0.92	1.1	1	1	1.8
Linear 3	0.04	0.91	8.01	0.75	0.98	1.34	1	1	-2.4
Serfling 1	0.18	0.18	0.82	0.76	0.76	0.23	1	1	-8.1
Serfling 2	0	0	1	0.11	0.11	0.89	1	1	-1
Serfling 3	0	0	1	0.03	0.7	0.93	0	1	<b>-30</b>

Table 8.1: Comparing performance of  $WCR_{v1}$  with CUSUM and MEM in different scenarios, Sen. represents sensitivity, Sp. represents specificity,  $MED$  represents median time to detection, F means that the method failed to raise an alarm.

We already mentioned the shape of each scenario in section 7.2. From the first illustration (as in figure 8.5), we find that  $WCR_{v1}$  almost always raises the alarm in the first week, and before we can consider this as an advantage for  $WCR_{v1}$ , we should test it in the steady state without any increase to check if those are false alarms or not. Another point worth mentioning is that MEM always misses the detection when there are outbreaks in the historical data or reported high rates in the past or if the history is a decreasing scenario. To see the performance of  $WCR_{v1}$  when there is no increase at all, we will change the rate

to be a flat rate and see how many alarms are raised with this steady state situation.

Figure 8.6 shows the false alarms (FAs) (because we assumed a flat rate).  $WCR_{v1}$  raised 63 FAs, while CUSUM raised 53 FAs and MEM raised 72 FAs. The lowest FA rate in this case was found with CUSUM, while the worst performance was found with MEM, and  $WCR_{v1}$  was between the two.

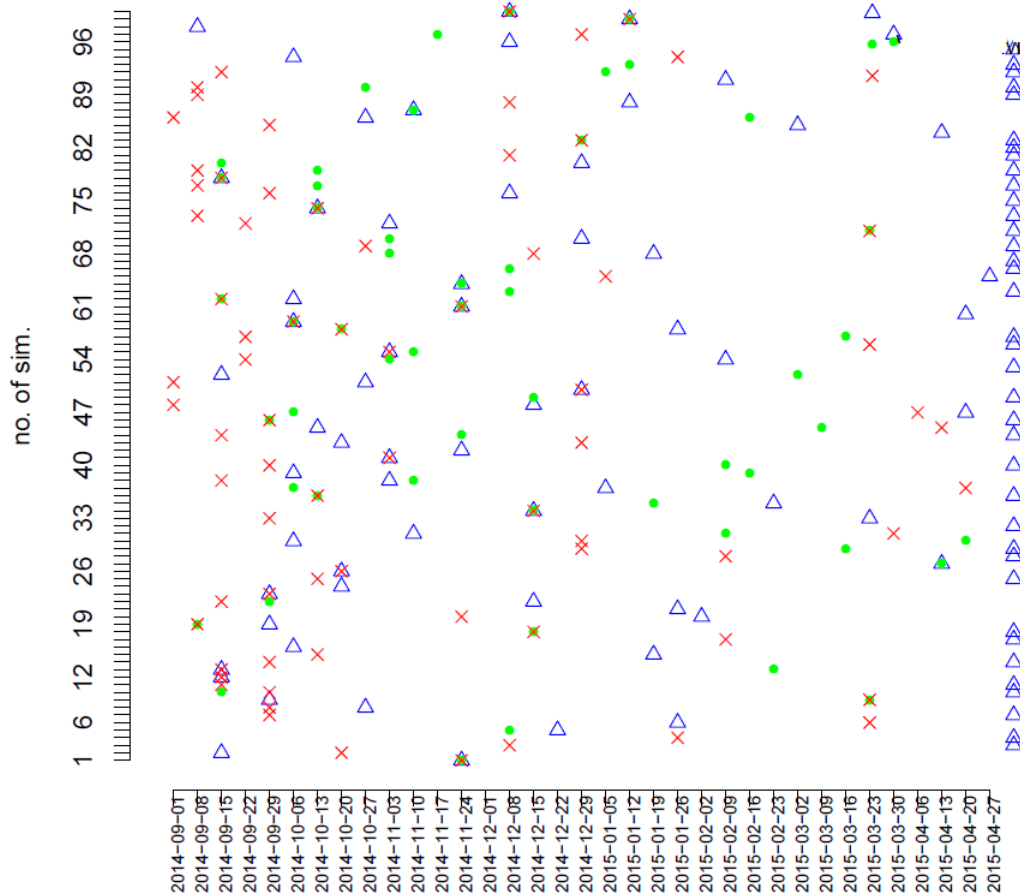


Figure 8.6: Comparison of false alarms (FA) in weeks between  $WCR_{v1}$  (blue triangles), CUSUM (green dots) and MEM (red crosses); in the flat rate case with 100 simulations.

As we found lots of FAs in figure 8.6, we may accept some of those FAs with CUSUM and MEM, since they rely mainly on historical data, but we will think again about some ways to raise the cut-offs in  $WCR_{v1}$  and make some change in its setting to avoid those FAs with the steady rates.

In section 6.4, we talk about the possibility of adoption of the mean and standard deviation of  $WCR$  at each  $N_{HB}$  in the calculation of cut-offs in each  $N_{HB}$  by using `qnorm`, and

we concluded that is not accurate to assume  $\mu$  and  $\sigma$  from equations 5.4.7 and 5.4.10 to calculate the cut-offs at each  $N_{HB}$ , so we will test the WCR based on doing many simulations with different numbers of locations (HBs).

Also, we suggested using different values of the scaling factor  $r$  in the mean of the assumed underlying Poisson distribution for the rates ( $r = 1, 5, 10$ ) for testing the WCR system. We checked the performance when we considered  $r = 1$  ( $WCR_{v1}$ ). Now, we will show the performance when we consider the mean  $\theta = \frac{Min+Max}{10}$ , so  $r = 5$  ( $WCR_{v2}$ ), and  $\theta = \frac{Min+Max}{20}$ , so  $r = 10$  ( $WCR_{v3}$ ).

Using  $WCR_{v1}$ ,  $WCR_{v2}$ , and  $WCR_{v3}$  we can compare the performance when the baseline of hypothesis simulation is lower, as explained in section 5.2.

### 8.2.1 Testing $WCR_{v2}$ and $WCR_{v3}$

Here, in  $WCR_{v2}$  every setting will be the same as for  $WCR_{v1}$  except that the scaling factor  $r$  will equal 5, so cut-offs will change for this reason. Again in this approach, we will choose the alert week (week of increase) as the first week when the rate is at least 10% higher than in week 1. Results of comparison of the performance of the methods is shown in table 8.2.

Then we will continue to apply the system for all our assumed scenarios presented in section 7.2 and compare the performance between  $WCR_{v2}$  and  $WCR_{v3}$  in table 8.2 with  $WCR_{v1}$  and CUSUM and MEM in table 8.1.

We can find the epidemic week by choosing a specific rate of increase compared to the first week (we considered the alert week as the first week where the rate is 10% above the rate in week 1), so all those weeks from week 1 and before this specific week of increase will be considered for specificity and those weeks from the increase week and after that will be for sensitivity. Table 8.2 shows the comparison between using version 2 of WCR ( $WCR_{v2}$ ) and version 3 of WCR ( $WCR_{v3}$ ) in section 8.2. We did not include CUSUM and MEM as we do not expect changes in the result as there is not any change in their setting. Again here we are using 100 simulations. We see that performance is quite similar in most cases for both 14 and 34 HBs, and we can give some comments:

- In the historical, repeating 2014, decreasing, and the first Serfling scenarios, the performances are quite similar between  $WCR_{v2}$  and  $WCR_{v3}$  for either HBs= 14 and 34.
- In the increasing scenario, there are clear similarities in sensitivity and specificity between  $WCR_{v2}$  and  $WCR_{v3}$  with either HBs= 14 and 34 with a delay in their alarms except for  $WCR_{v2}$  with HBs= 14 as there is no delay in *MED*.

- For the shape of the first, the second and the third linear scenario, we see that with HBs= 34 there is lower sensitivity and specificity and also lower *MED* for either  $WCR_{v2}$  and  $WCR_{v3}$  compared to HBs= 14.
- For HBs= 34, both  $WCR_{v2}$  and  $WCR_{v3}$  have sensitivity and specificity zero with 1 week to detection.
- For the shape of the second Serfling with HBs= 14, the sensitivity and specificity are better as is the time to detection.
- For the shape of the third Serfling with HBs= 14, the specificity is better for both  $WCR_{v2}$  and  $WCR_{v3}$ .
- The sensitivity for  $WCR_{v3}$  with the second and third Serfling are not bad with HBs= 14.



$ng=14$						
	$WCR_{v2}$			$WCR_{v3}$		
Kind of Shape	Sen.	Sp.	$MED$	Sen.	Sp.	$MED$
Historical	0.02	0.02	0.98	0.06	0.06	0.94
Repeat 2014	0.02	0.03	0.97	0.06	0.06	0.94
Increasing	0.29	0.65	0.91	0.53	0.87	<b>-6.7</b>
Decreasing	0.07	0.07	0.93	0.07	0.07	0.93
Linear 1	0.36	0.93	3.9	0.84	0.98	0.88
Linear 2	0.12	0.89	2.7	0.51	0.94	0.01
Linear 3	0.39	0.94	4.5	0.78	0.98	1.3
Serfling 1	0.24	0.24	0.74	0.31	0.31	0.66
Serfling 2	0.5	0.5	0.46	0.42	0.42	0.49
Serfling 3	0	0.9	3.9	0	0.9	3.7
$ng=34$						
	$WCR_{v2}$			$WCR_{v3}$		
Kind of Shape	Sen.	Sp.	$MED$	Sen.	Sp.	$MED$
Historical	0.03	0.03	0.97	0.05	0.05	0.95
Repeat 2014	0.04	0.04	0.96	0.04	0.04	0.96
Increasing	0.33	0.89	<b>-10.5</b>	0.26	0.9	<b>-14</b>
Decreasing	0.03	0.03	0.97	0.06	0.06	0.94
Linear 1	0.22	0.91	4.3	0.6	0.96	2.4
Linear 2	0.02	0.88	4.09	0.34	0.92	1.7
Linear 3	0.22	0.93	5.5	0.68	0.97	1.9
Serfling 1	0.17	0.17	0.83	0.16	0.16	0.84
Serfling 2	<b>0</b>	<b>0</b>	1	<b>0</b>	<b>0</b>	1
Serfling 3	<b>0</b>	<b>0</b>	1	<b>0</b>	<b>0</b>	1

Table 8.2: Comparing performance of  $WCR_{v2}$  and  $WCR_{v3}$  in different scenarios, Sen. represents sensitivity, Sp. represents specificity,  $MED$  represents median time to detection.

### 8.2.2 Testing $WCR_{v1,2,3}$ with different setting of the alert week

In the results in tables 8.1 and 8.2, we see that warnings often came late with the MEM system. This cannot be considered as a disadvantage with MEM, because MEM adopts a different strategy for the onset of the epidemic, as we showed in section 3.3.

Now, we will  $WCR_{v1}$ ,  $WCR_{v2}$  and  $WCR_{v3}$  and compare the performance with CUSUM and MEM as we did in section 8.2 and 8.2.1 but now changing the alert week to be the

same as in the MEM setting. Firstly, we will test  $WCR_{v1}$ , using scaling factor  $r=1$ , and table 8.3 shows the comparison between these three systems when using version 1 of *WCR* from section 8.2.

We see that performance in many cases is similar, for both better or worse, with 14 and 34 HBs and we can give some comments:

- MEM outperformed the other methods in general and this is because of the different approach of the MEM method for the alert week setting.
- For the first linear shape, the CUSUM method outperformed in general the MEM method, which came second, while  $WCR_{v1}$  had the worst performance.
- Here we cannot ignore to mention that there is clear poor performance of  $WCR_{v1}$  and CUSUM with HBs= 34. It is still poor with HBs= 14 but worse with HBs= 34 apart from CUSUM with either increasing or decreasing shape. 14 HBs is the baseline used and the data for 34 HBs come from re-sampling these 14 HBs, but the reason for the poor performance is not clear.

$ng=14$									
	$WCR_{v1}$			CUSUM			MEM		
Kind of Shape	Sen.	Sp.	$MED$	Sen.	Sp.	$MED$	Sen.	Sp.	$MED$
Historical	0	0.91	5	0	0.91	4	<b>F</b>	<b>F</b>	<b>F</b>
Repeat 2014	0	0.91	5	0	0.91	4.4	0.71	0.97	0.3
Increasing	0	0.91	3.9	0	0.91	2	0	0.91	3
Decreasing	0	0.91	5	0.99	0.99	0.01	<b>F</b>	<b>F</b>	<b>F</b>
Linear 1	<b>0.07</b>	0.92	6.3	<b>0.78</b>	0.98	1.1	<b>0.66</b>	0.97	2
Linear 2	0	0.91	7.5	0	0.91	3.4	0.78	0.98	0.1
Linear 3	0	0.91	10	0	0.91	3	0.3	0.94	0.7
Serfling 1	0	0.91	9	0	0.91	8	1	1	-0.2
Serfling 2	0	0.91	7	0	0.91	6	0.21	0.93	0.8
Serfling 3	0	0.91	6	0	0.91	3	0.19	0.93	0.8

$ng=34$									
	$WCR_{v1}$			CUSUM			MEM		
Kind of Shape	Sen.	Sp.	$MED$	Sen.	Sp.	$MED$	Sen.	Sp.	$MED$
Historical	0	0	5	0	0	4	<b>F</b>	<b>F</b>	<b>F</b>
Rep. 2014	0	0	5	0	0	4	0.82	0.82	0.18
Increasing	0	0	2	1	1	-0.9	0	0	3
Decreasing	0	0	5	1	1	-0.3	<b>F</b>	<b>F</b>	<b>F</b>
Linear 1	<b>0.07</b>	0.07	6	<b>0.79</b>	0.79	1.1	<b>0.74</b>	0.74	1.4
Linear 2	0	0	8.8	0	0	4.4	0.07	0.07	1.2
Linear 3	0	0	11	0	0	4.5	0.41	0.41	0.6
Serfling 1	0	0	9	0	0	8	1	1	-0.1
Serfling 2	0	0	2	0	0	1.8	1	1	0
Serfling 3	0	0	29	0	0	29	1	1	-2

Table 8.3: Comparing performance of  $WCR_{v1}$  with CUSUM and MEM in different scenarios, Sen. represents sensitivity, Sp. represents specificity,  $MED$  represents median time to detection, F means that the method failed to raise an alarm.

Now, we will continue to apply the system for all our assumed scenarios presented in section 7.2 and compare the performance of  $WCR_{v2}$  and  $WCR_{v3}$  in table 8.4 with  $WCR_{v1}$ , and CUSUM and MEM in table 8.3.

We can find the epidemic week by assuming the alert week to be the same as in the MEM setting, as explained in section 3.3.

$ng=14$						
	$WCR_{v2}$			$WCR_{v3}$		
Kind of Shape	Sen.	Sp.	$MED$	Sen.	Sp.	$MED$
Historical	0	0.91	5	0	0.91	5
Repeat 2014	0	0.91	5	0	0.91	5
Increasing	0.03	0.91	2.9	0.19	0.95	-5.8
Decreasing	0	0.91	5	0	0.91	5
Linear 1	0.34	0.94	4.1	0.75	0.98	1.3
Linear 2	0	0.91	6	0.1	0.92	1.6
Linear 3	0	0.91	5.6	0	0.91	3.9
Serfling 1	0	0.91	8.8	0	0.91	8.7
Serfling 2	0	0.91	7.4	0	0.91	7.5
Serfling 3	0	0.91	5.9	0	0.91	5.4

$ng=34$						
	$WCR_{v2}$			$WCR_{v3}$		
Kind of Shape	Sen.	Sp.	$MED$	Sen.	Sp.	$MED$
Historical	0	0	5	0	0	5
Repeat 2014	0	0	5	0	0	5
Increasing	0.35	0.59	-8	0.34	0.67	-10
Decreasing	0	0	5	0	0	5
Linear 1	0.23	0.23	4.6	0.56	0.56	2
Linear 2	0	0	5.9	0	0	3.7
Linear 3	0	0	8.5	0	0	6
Serfling 1	0	0	8.9	0	0	8.7
Serfling 2	0	0	2	0	0	2
Serfling 3	0	0	29	0	0	29

Table 8.4: Comparing performance of  $WCR_{v2}$  and  $WCR_{v3}$  in different scenarios, Sen. represents sensitivity, Sp. represents specificity,  $MED$  represents median time to detection.

### 8.3 Conclusion

From these tables 8.3 and 8.4, we can conclude that:

- In general, the relationship between sensitivity and  $MED$  is inverse, which means that when the sensitivity is high, we will lose some time from  $MED$ .
- CUSUM and  $WCR$  (in its different versions) usually raise the alarms before MEM, which is an advantage for CUSUM and  $WCR$ .

- MEM usually has high sensitivity and specificity compared to CUSUM and  $WCR$  in this setting.
- If we had an outbreak in the past years, MEM will never raise an alarm.
- In the decreasing scenario, MEM did not raise any alarm.
- When we compare  $WCR_{v1}$  ( $r = 1$ ),  $WCR_{v2}$  ( $r = 5$ ) and  $WCR_{v3}$  ( $r = 10$ ), keeping the alert week (week of increase) as the first week when the rate is at least 10% higher than in week 1, we found that the sensitivity and specificity are higher when the scaling factor  $r$  is higher (i.e. the baseline rates  $\theta$  are lower) while maintaining good performance in  $MED$ . In the same setting, we found high sensitivity and specificity with clear delay in alarms ( $MED$ ) with MEM.
- When we compare  $WCR_{v1}$ ,  $WCR_{v2}$ ,  $WCR_{v3}$  and CUSUM, keeping the alert week to be the same as in the MEM setting (as explained in section 3.3), we found the sensitivity and specificity are in general low, while maintaining good performance in  $MED$ . Again, in this setting, we found higher sensitivity and specificity with clear improvement in delay of alarms ( $MED$ ) with MEM.
- For  $ng = 14$ , sensitivity of CUSUM is high for linear 1 and decreasing scenarios.
- For  $ng = 14$  and  $WCR_{v3}$  sensitivity is high for linear 1 shape.
- Moreover the performance in  $MED$  is not always good i.e. with  $ng = 34$  and all methods apart from MEM the Serfling 3 shape has  $MED = 29$ .
- The linear 3 and Serfling 1 shapes seem to have high  $MED$  sometimes for  $WCR_{v1}$ ,  $WCR_{v2}$ ,  $WCR_{v3}$  and CUSUM.
- When we did this simulation study to compare performance between these three systems, we show  $WCR_{v1}$ ,  $WCR_{v2}$ ,  $WCR_{v3}$ , testing sensitivity and specificity each time, and we did not repeat those tests with CUSUM and MEM. The reason for this is because the change in the system happened in  $WCR$  only.

To conclude this chapter, we can see clear deficiencies in performance of  $WCR$  compared to CUSUM and MEM, and this requires more investigation to find the reasons for these problems.

In section 5.1, we can see that the population in each location was considered to be one million, but this is not realistic. When we did the flat rate simulation in section 5.2, the baseline of simulation was very high, which affects the steady state hypothesis test. Finally, a very important point is that we assumed all  $WCR$  in the area  $N_{HB}$  from  $0, 1, 2, \dots, \frac{ng}{2} - 1$  are symmetric with all  $WCR$  in the area  $N_{HB}$  from  $\frac{ng}{2} + 1, \dots, ng$ , but in reverse (see chapter 5, section 5.2) but this may not be accurate enough.

We will discuss those points in more detail and present some suggested solutions and further results in the next chapter.

## Chapter 9

# Testing *WCR* with other systems (2)

In chapter 5, we discussed a virtual environment to consider as the normal situation and examined possibilities for determining the situation when an alarm should be raised in the *WCR* system. When we tested the system in different scenarios in section 8.2, we found some limitations in our assumptions which may have affected the results. We reconsider these here and make changes to try to generate better results from our extended *WCR* system.

- In chapter 5, section 5.1, we considered the population in each sub-region as equal, around one million. In reality, this is not realistic as we cannot guarantee this size in different countries.
- In chapter 5, section 5.2, we considered the baseline of simulation as the ILI counts in Scotland and as we mentioned in section 6.4, this leads to a rather high rate per 100,000 of population in the steady state case, which affects the hypothesis test used in the *WCR* system. We will consider a different baseline here.
- In chapter 5, section 5.2, we assumed that all *WCR* in the range of  $N_{HB}$  in  $0, 1, 2, \dots, \frac{ng}{2} - 1$  are symmetric with all *WCR* in the range of  $N_{HB}$  in  $\frac{ng}{2} + 1, \dots, ng$ , but in reverse. When we investigate the exact numbers in each  $N_{HB}$ , for example, we can see table 9.1 with  $ng=14$ . It tells us that  $N_{HB}$  before 7 are not equal to  $N_{HB}$  after 7, as they are not symmetric. The reason why they are not symmetric is because if we take  $ng=14$ , for example, then looking at the count in week  $i$  it is greater than in week  $i-1$ , and when the rate is low, for example, 1 per 100,000, then it is easy to have the same number of  $N_{HB}$  in the different weeks, which means no increase in  $N_{HB}$ . This also affects another consideration, which is the success prob-

$N_{HB}$	0	1	2	3	4	5	6
number of cases	50	3493	21505	82260	217621	416212	593224
7	8	9	10	11	12	13	14
650381	543014	344809	164341	57246	13570	1942	132

Table 9.1: Example for the number of cases in each  $N_{HB}$  with  $ng=14$ . The total number of simulated cases is 3,110,000. This is based on steady state simulation.

ability of the Binomial distribution for  $N_{HB}$  considered in section 6.3. We address each of these limitations in the sections below.

### 9.1 Symmetry deception

When we investigate the exact numbers of each  $N_{HB}$  occurring in the simulations with different  $ng$ , we found the frequency of each value of  $N_{HB}$  before the midway point is not the same as the frequency of each value of  $N_{HB}$  after the midway point, which means they are not symmetric. This affected another consideration, which is the probability of the Binomial distribution for  $N_{HB}$  used in section 6.3. When we calculated the Binomial pmf function for  $N_{HB}$ , we considered the success probability = 0.5 as we thought the  $N_{HB}$  were symmetric, but they are not. For example, table 9.2 presents the actual means of  $N_{HB}$  divided by  $ng$  when  $ng=14$  and 34, which estimate the actual success probability in each case.

Binomial probability with fixed rate			
$ng$	1 per 100,000	3 per 100,000	5 per 100,000
14	0.4425	0.4679	0.4753
34	0.4159	0.4536	0.4644
Binomial probability with varying rate			
$ng$	1 per 100,000	3 per 100,000	5 per 100,000
14	0.4417	0.4676	0.4751
34	0.4155	0.4535	0.4646

Table 9.2: Actual mean of  $N_{HB}$  divided by  $ng$  when  $ng=14$  and 34. These came from 3,110,000 simulations of ILI cases in the steady state case.

To examine this problem and its effects, we present an example of the Binomial pmf with probability = 0.5 in figure 9.1 as well as the actual Binomial pmf with probability = 0.4159 for  $ng=34$ . They do not match well. This is a serious issue which affects the setting of  $N_{HB}^{**}$  and  $N_{HB}^*$ , and so affects the hypothesis test in chapter 8, section 8.1. This



means we need to calculate the actual mean of  $N_{HB}$  divided by  $ng$  to use as the success probability in the Binomial distribution for  $N_{HB}$ . All future results in chapter 9 all use this modification.

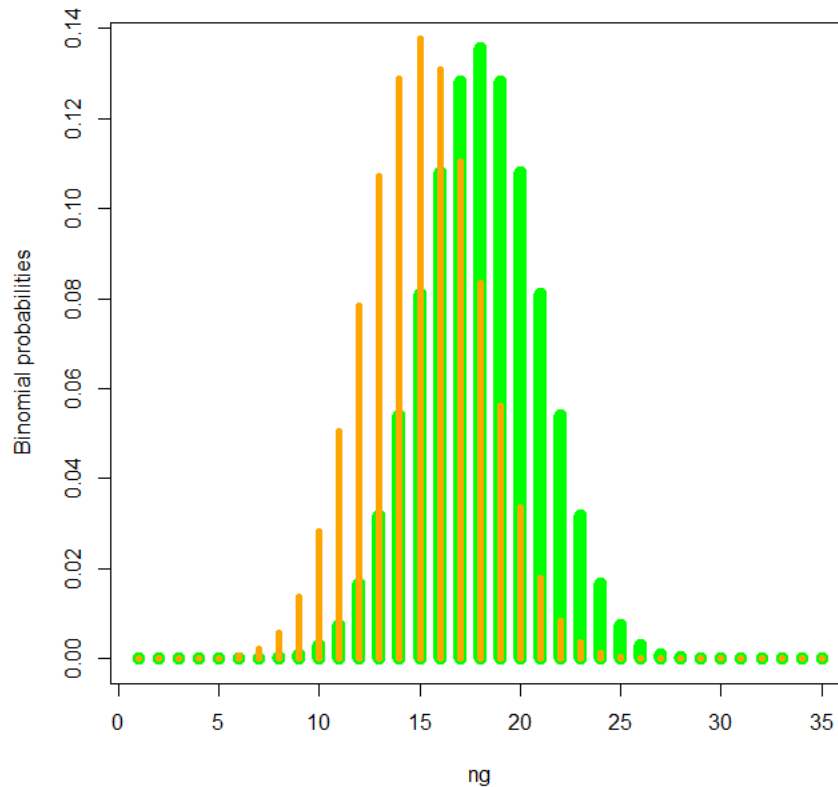


Figure 9.1: Binomial probability mass function with probability = 0.5 shown by the green columns and the actual Binomial pmf with probability = 0.4159 shown by the orange column, for 34 HBs.

## 9.2 Baseline of simulation (hypotheses)

In the fifth chapter, section 5.2, we considered the mid-range  $\frac{\max(ILI) + \min(ILI)}{2}$ , where max and min are maximum and minimum ILI count in  $L$ ,  $L$  is a location, over all the years in each location as the base point of our simulation of ILI counts from a Poisson distribution with that mean value. We did not take into account that in any  $N_{HB}$ , the ILI rates usually decreased to zero in the summer, which means the baseline will be  $\frac{\max(ILI)}{2}$  and so using the mid-range leads to too high a number for the baseline of our simulation. The mid-range was around 500, and when we simulated from this number (as the mean ILI count) and calculated the rate, then calculated the  $WCR$ , we found the mean rate was around 30 cases per 100,000 of population. In fact, this cannot be considered as a

Simulation label	Rate structure	Kind of rate
6 $\Leftrightarrow$ 1	1 per 100,000	Fixed rate
7 $\Leftrightarrow$ 2	3 per 100,000	Fixed rate
8 $\Leftrightarrow$ 3	5 per 100,000	Fixed rate
9 $\Leftrightarrow$ 4	5 per 100,000	Varying rate
10 $\Leftrightarrow$ 5	10 per 100,000	Varying rate

Table 9.3: Illustration of parameters of each approach of simulation.

baseline of the steady state, as this is too high. This point was already shown in table 6.3 in chapter 6.

There are different ways to set the base point of simulation. In chapter 5 in section 5.2.3 and 5.2.4, we explained how to use fixed and varying rates as base points to simulate ILI counts. In the fixed rate case, we use this approach to calculate the baseline counts for simulation: we will consider the rate as fixed in each location and we can consider it as 1, 3 or 5 per 100,000 (we can choose any number but we select those to compare with the other structure below). The same logical consideration to sample ILI counts for 30 and 50 HBs from the Scottish 14 HBs in section 5.2.2 was considered. For the varying rate case, we can consider the base point of simulation as using varying rates in each location, but keeping the average as in the Scottish data. To calculate the rate in the end, we can use the same Scottish population structure as in section 5.2.3. For example, if we consider the rate as 1 per 100,000, with fixed rate, all different locations use the same rate 1 per 100,000, but with varying rate, the rate will move up and down (say 0.8 per 100,000 in  $L1$  and 1.2 in  $L2$ ) between different locations with keeping the average for all locations as 1 per 100,000.

As explained, there are two possible ways to consider the population in each location, which are:

**a** Considering the Scottish population structure ( $P_1$ ).

**b** Considering equal populations in each different location ( $P_2$ ).

Before now, we have used structure  $P_2$ . Here we selected  $P_1$ , the Scottish population structure, as it is more realistic, and the means of  $WCR$  using  $P_1$  and  $P_2$  are similar, so we will consider it here, with another 5 cases labelled 6 to 10 in table 9.3. Cases 1 to 5 considered  $P_2$  and their results (means) are presented in figure 9.2.

Figure 9.3 represents means of  $WCR$  for each  $N_{HB}$  using fixed and varying rates, and figure 9.4 represents variances of  $WCR$  for each  $N_{HB}$  with fixed and varying rates. In

each case, the trend is similar when we use fixed or varying rates, so we will assume one of them (we will choose fixed rates).

From figures 9.2, 9.3 and 9.4, we will investigate modelling of the mean and variance of  $WCR$  to check for the possibility of adopting those means and variances in the calculation of cut-offs for  $WCR$  at each  $N_{HB}$ . Means seem to increase against  $N_{HB}$  while variances seem to be constant.

We used the same notations ( $n_i$ ,  $\mu$ ,  $\text{var}$ ,  $n.i.\text{midway}$ ,  $\text{weights}$ ) as in section 5.4.

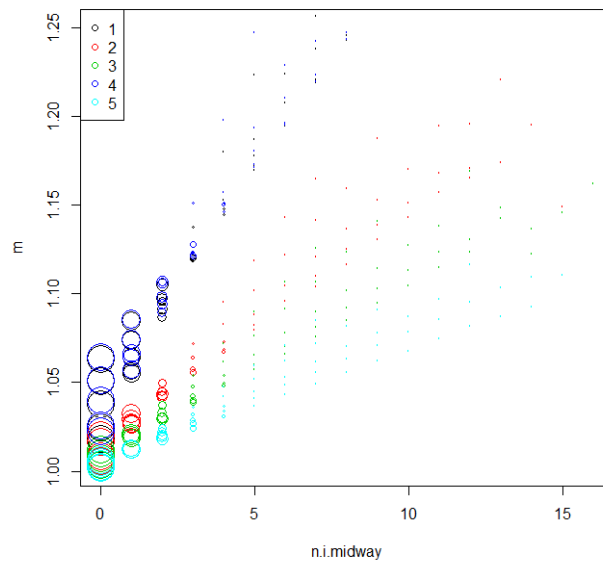


Figure 9.2: Means ( $m$ ) of  $WCR$  against  $N_{HB}$  for fixed rates (labelled 1, 2, 3) and varying rates (labelled 4, 5) in table 9.3 represented as circles based on 3,110,000 simulations of ILI consultations using the steady state case. The size of circles represents the number of simulations for each  $n_i$ .

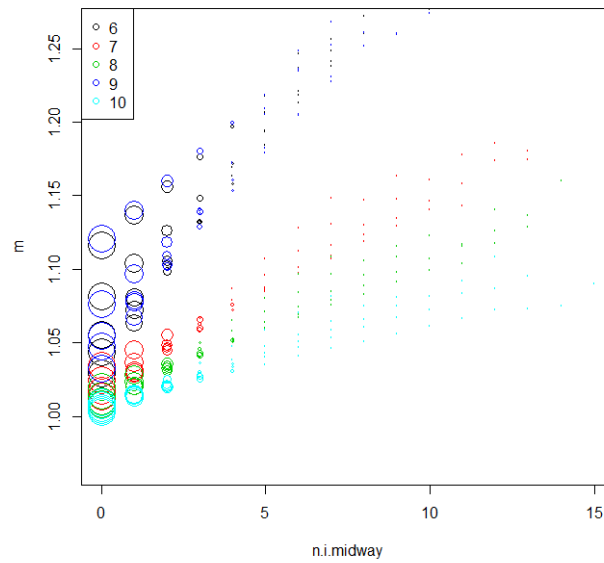


Figure 9.3: Means ( $m$ ) of  $WCR$  against  $N_{HB}$  for fixed rates (labelled 6, 7, 8) and varying rates (labelled 9, 10) in table 9.3 represented as circles based on 3,110,000 simulations of ILI consultations using the steady state case. The size of circles represents the number of simulations for each  $n_i$ .

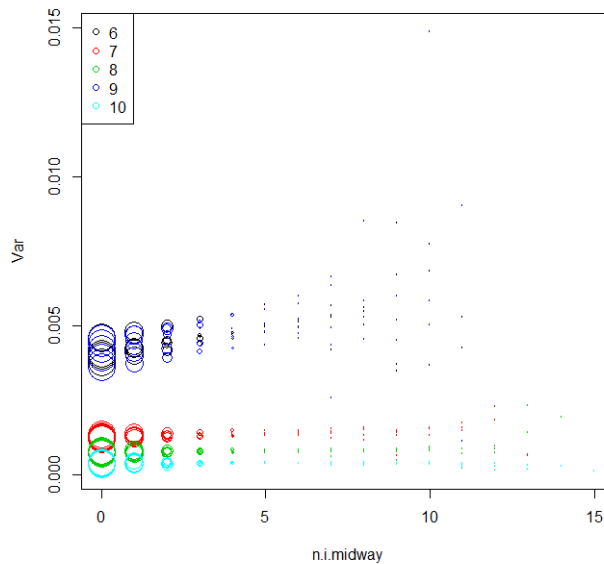


Figure 9.4: Variances ( $Var$ ) of  $WCR$  against  $N_{HB}$  for fixed rates (labelled 6, 7, 8) and varying rates (labelled 9, 10) in table 9.3 represented as circles based on simulation of 3,110,000 ILI consultations using the steady state case. The size of circles represents the number of simulations for each  $n_i$ .

z7= lm(m ~ n.i.midway + rate + ng, weights= weight)			
coefficient	estimate	Std. Error	<i>p</i> -value
Intercept	1.0429080	0.0045276	<2e-16
n.i.midway	0.0157269	0.0011714	<2e-16
rate	-0.0089893	0.0004590	<2e-16
ng	0.0008943	0.0001192	<2e-16

Table 9.4: Short summary of model (z7).

z8= glm(var ~ n.i.midway + rate + ng, family = Gamma, weights= weight)			
coefficient	estimate	Std. Error	<i>p</i> -value
Intercept	-11.2329	4.0867	0.006373
n.i.midway	-9.3168	0.9155	<2e-16
rate	265.4865	1.6027	<2e-16
ng	-0.3855	0.1022	0.000198

Table 9.5: Short summary of model (z8).

### 9.3 Modelling means and variance

In chapter 5, sections 5.3 and 5.4, we discussed possible ways to find a general model for the mean and the standard deviation or variance of  $WCR$  for each  $N_{HB}$ , assuming  $WCR$  had a normal distribution for each  $N_{HB}$ . We will now try to find a model for the means and variances with the new many simulated cases as presented in section 9.2. A summary of the linear model for the means (z7) is presented in table 9.4. This model includes weights (as in section 5.4). We can see from the *p*-values less than 0.05 that the terms in the model are significant. Model (z8) in table 9.5 shows the summary of `glm` for variances.

Details of models (z7) and (z8) are presented in Appendix D.1 and tables 9.4 and 9.5. Figures 9.5 and 9.6 show predicted values of the means and variances. As presented in the plots, the means show an increasing trend from the midway point to the end of  $N_{HB}$ , while the variances show much more constant rates.

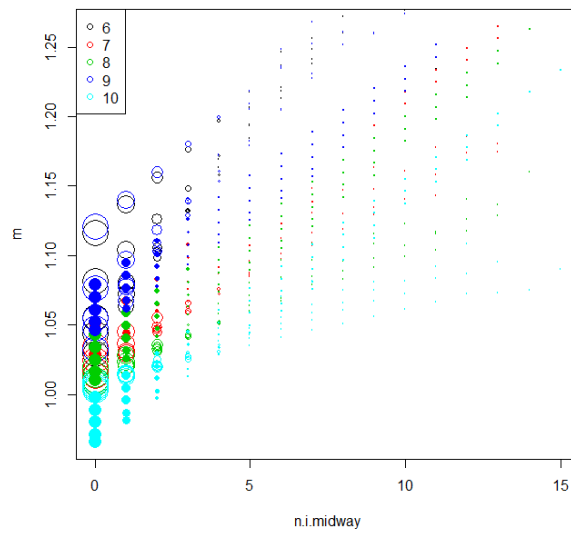


Figure 9.5: Means ( $m$ ) of  $WCR$  against  $N_{HB}$  for fixed rates (labelled 6, 7, 8) and varying rates (labelled 9, 10) in table 9.3 represented as open circles. Coloured solid circles represent the predicted values from (z7) model. The size of circles represents the number of simulations for each  $n_i$ . Despite trying more complex models with interactions and transformations, (z7) model did not fit very well as there is under prediction with high rates like 10 per 100,000.

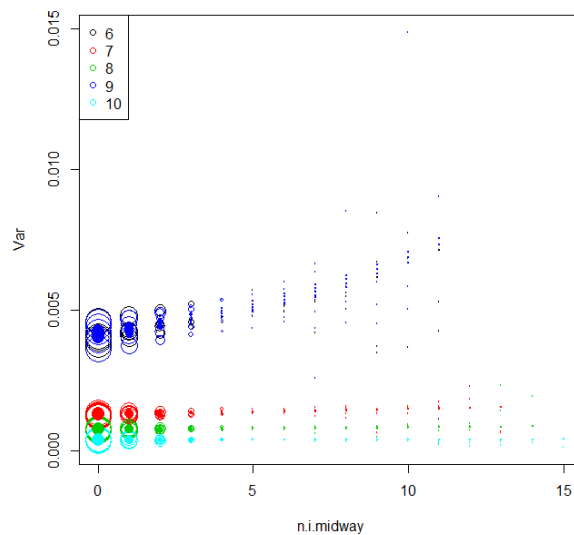


Figure 9.6: Variances ( $Var$ ) of  $WCR$  against  $N_{HB}$  for fixed rates (labelled 6, 7, 8) and varying rates (labelled 9, 10) in table 9.3 represented as open circles. Coloured solid circles represent the predicted values from (z8) model. The size of circles represents the number of simulations for each  $n_i$ .

Using the fitted models (z7), and (z8), we can write the equations for the mean and variance of  $WCR$  for each value of HBs ( $ng$ ) as below:

$$Mean = 1.0429 + 0.0157269 \times (n.i.midway) - 0.0089 \times (Rate) + 0.0009 \times (ng). \quad (9.3.1)$$

$$Variance = \frac{1}{-11.2329 - 9.3168 \times (n.i.midway) + 265.4865 \times (Rate) - 0.3855 \times (ng)}. \quad (9.3.2)$$

### 9.3.1 Testing of mean and variance equation

As explained in chapter 6, section 6.4, there were problems developing accurate equations for the mean and the standard deviation or variance of  $WCR$  to use in the calculation of cut-offs for  $WCR$  with each  $N_{HB}$ . We will consider the performance of the new equations 9.3.1 and 9.3.2 for the mean and variance of  $WCR$  for each  $N_{HB}$ .

We will investigate the system coverage of the rejection region when we consider a fixed rate= 1, 3 or 5 cases per 100,000 population.

Figure 9.7 shows the system in the steady state case and with fixed rate= 1 per 100,000 population when the significance level  $\alpha = 5\%$ . It compares the cut-offs using  $\mu$  and  $\sigma$  from 3,110,000 simulated cases in the flat rate case to the cut-offs using  $\mu$  and  $\sigma$  from equations 9.3.1 and 9.3.2. We can see that the equations 9.3.1 and 9.3.2 for the mean and variance of  $WCR$  are also not usable for cut-off limits because using the  $\mu$  and  $\sigma$  equations did not give the same or similar cut-offs as using the simulated cases.

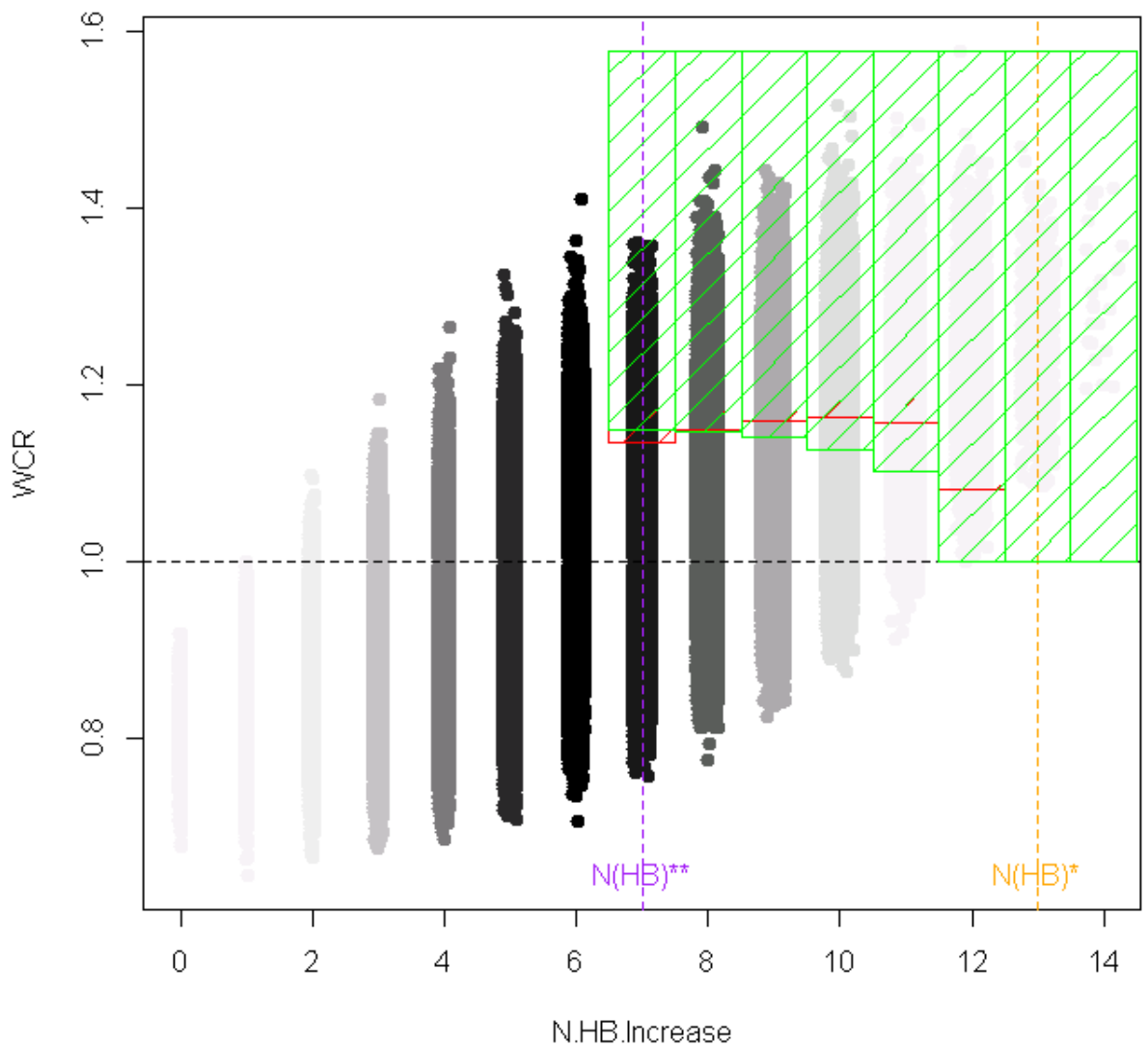


Figure 9.7:  $WCR$  against  $N_{HB}$  for steady rates using fixed rate = 1 per 100,000 population with 14 HBs.  $N_{HB}^*$  is presented in the dotted vertical orange line while the dotted vertical purple line represents  $N_{HB}^{**}$ . The red shaped columns came from calculating the quantiles (cut-offs) using  $\mu$  and  $\sigma$  from many simulated cases (3,110,000) in the flat rate situation. They represent around 5% most extreme observations under the steady state. The green columns came from using `qnorm` to calculate cut-offs using  $\mu$  and  $\sigma$  from equations 9.3.1 and 9.3.2. They represent around 6% most extreme observations under the steady state. Darkness of the black colour reflects the number of simulated cases.

In figures 9.7, D.1, D.2, D.3, D.4 and D.5 in the Appendix and tables 9.6 and 9.7, which show the small case of HBs= 14 and big case of HBs=34, we can see that it is not accurate



percentage of coverage when $\alpha = 5\%$					
HB= 14	simulation	equation	HB= 34	simulation	equation
$fr = 1$	0.05	0.06	$fr = 1$	0.05	0.06
$fr = 3$	0.05	0.03	$fr = 3$	0.05	0.07
$fr = 5$	0.05	0.04	$fr = 5$	0.05	0.007

Table 9.6: Percentage of coverage of the rejection region using the total number of simulations (3,110,000), where  $fr$  means fixed rate, “simulation” means the cut-offs are based on the actual simulated cases (3,110,000), while “equation” means that the cut-offs are calculated based on mean and variance equations 9.3.1 and 9.3.2. The simulated cases did to calculate the cut-off and then these are repeats with further simulations using the same cut-off.

percentage of coverage when $\alpha = 1\%$					
HB= 14	simulation	equation	HB= 34	simulation	equation
$fr = 1$	0.01	0.02	$fr = 1$	0.01	0.02
$fr = 3$	0.016	0.008	$fr = 3$	0.01	0.002
$fr = 5$	0.015	0.02	$fr = 5$	0.01	0.0008

Table 9.7: Percentage of coverage of the rejection region using the total number of simulations (3,110,000), where  $fr$  means fixed rate, “simulation” means the cut-offs are based on the actual simulated cases (3,110,000), while “equation” means that the cut-offs are calculated based on mean and variance equations 9.3.1 and 9.3.2. The first set of simulated cases were used to calculate the cut-off and then these were repeated with further simulations using the same cut-off.

to assume  $\mu$  and  $\sigma$  from equations 9.3.1 and 9.3.2 to calculate the cut-offs for  $WCR$  with each  $N_{HB}$  as the percentage of coverage of the rejection region using this approach does not match well in various cases with the coverage from the simulations. This means that we have failed to find a simple equation to describe the relationship between both the mean and variance of  $WCR$  at each  $N_{HB}$ .

So instead we will test the  $WCR$  system based on doing many simulations with each change of the number of subset regions (HBs).

Fixed rates were used in sections 9.2 and 9.3.1. In figures 9.8 and 9.9 the cut-offs from fixed and varying rates over regions are compared. It can be seen that they are very similar and for this reason no further investigation with varying rates is adopted.

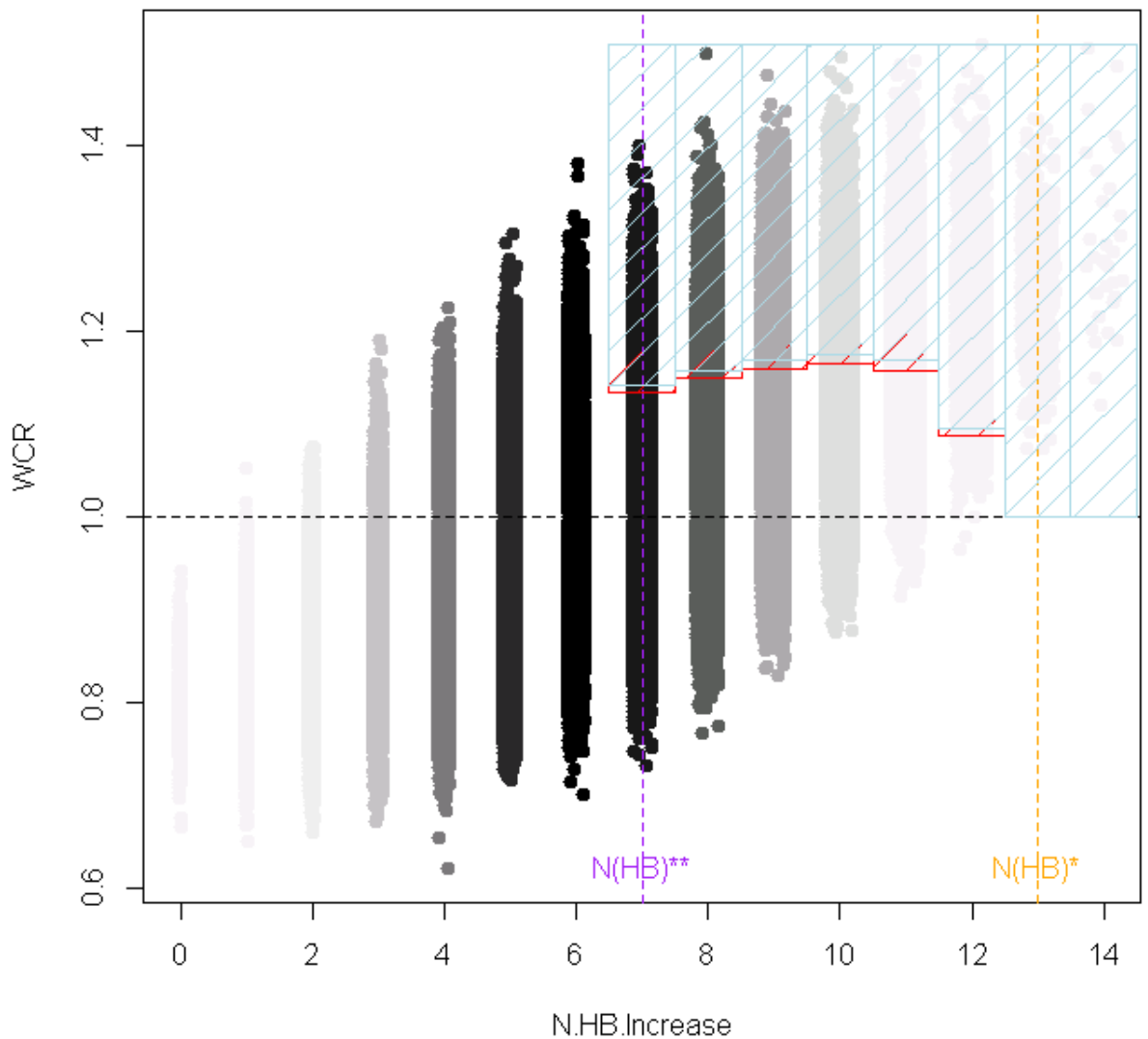


Figure 9.8:  $WCR$  against  $N_{HB}$  for steady rates using fixed rate = 1 per 100,000 population with 14 HBs.  $N_{HB}^*$  is presented in the dotted vertical orange line while the dotted vertical purple line represents  $N_{HB}^{**}$ . The red shaped columns present cut-offs using 3,110,000 simulations from the fixed rates situation, while the light blue columns present cut-offs using simulations from varying rates.

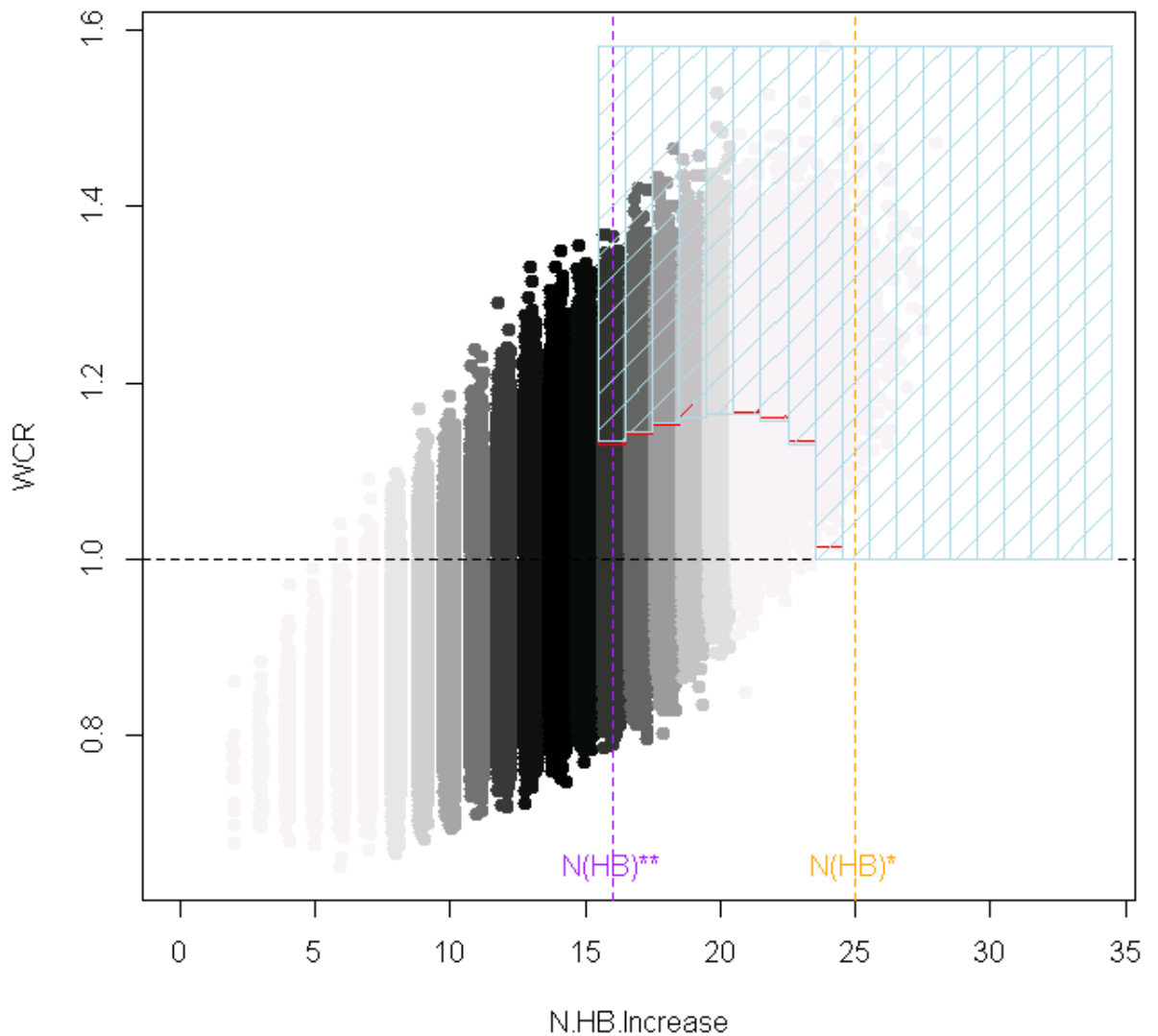


Figure 9.9: *WCR* against  $N_{HB}$  for steady rates using fixed rate = 1 per 100,000 population with 34 HBs.  $N_{HB}^*$  is presented in the dotted vertical orange line while the dotted vertical purple line represents  $N_{HB}^{**}$ . The red shaped columns present cut-offs using 3,110,000 simulated cases from the fixed rates situation, while the light blue columns present cut-offs using simulated cases from varying rates.

## 9.4 Testing the *WCR* system (2)

We previously tested the *WCR* system in 10 different scenarios and compared its performance with CUSUM and MEM in chapter 8. Here we have made some modifications of the *WCR* system in this chapter, so we will again test the *WCR* system with its new

modifications (cut-offs from the simulations and non-symmetric Binomial with probabilities calculated from the simulations) under the hypothesis test with significance levels of 1% and 5%.

We will start as we did in section 8.2 and we will choose the alert week (week of increase) as the first week when the rate is at least 10% higher than in week 1. We use 100 simulations.

Table 9.8 shows the performance of the modified *WCR* system when  $ng=14$  and 34 and  $fr = (1, 3, 5)$  with  $\alpha = 5\%$ , with CUSUM and MEM in different scenarios. From this table, we mention some points below:

- In the historical scenario, CUSUM performance is much the best. MEM fails here because its warnings depend on the average rates in the past 5 years. There was an outbreak in 2009, which caused a rise in the average rate over the past years and thus a loss of alarms after the epidemic start. *WCR* is relatively poor.
- In the scenario of repeating the 2014 cases six times, as for the historical scenario *WCR* is much poorer than CUSUM, while MEM shows the best, high, sensitivity and specificity but with a delay in its alarm.
- In the increasing scenario, the best performance is with CUSUM, while the *WCR* and MEM systems do not show good sensitivity.
- In the decreasing scenario, MEM fails completely. *WCR* and CUSUM do not show perfect detection as the sensitivity and specificity are low for *WCR*, and although they are perfect with CUSUM there is a delay in the alarm for CUSUM.
- With the shape of the first linear scenario, CUSUM performance is the best, then MEM comes second, while *WCR* is the worst as its sensitivity is generally low (except for  $ng=34$  where it does well).
- With the shape of the second and third linear scenarios and the shape of the first Serfling scenario, CUSUM performance is considered the best overall, then MEM comes second as it has perfect sensitivity and specificity, but with a delay in its alarm, while *WCR* is the worst as its sensitivity is very low for  $ng=14$  though it does well with  $ng=34$  (excluding the first Serfling scenario).
- With the shape of the second Serfling scenario, for  $ng=14$ , CUSUM and MEM show high sensitivity and specificity but with delay in alarms, as does MEM with  $ng=34$ , while *WCR* shows the worst performance.
- With the shape of the third Serfling scenario, for  $ng=14$ , *WCR* and CUSUM show low sensitivity and high specificity. MEM again here shows high sensitivity and high specificity but with a delay in its alarm. For  $ng=34$  all methods are poor.

$ng=14, \alpha = 5\%, E.S. = 10\%$															
Kind of Shape	WCR, Fixed rate=1			WCR, Fixed rate=3			WCR, Fixed rate=5			CUSUM			MEM		
	Sen.	Sp.	MED	Sen.	Sp.	MED	Sen.	Sp.	MED	Sen.	Sp.	MED	Sen.	Sp.	MED
Historical	0.1	0.1	0.9	0.1	0.1	0.9	0.16	0.16	0.84	0.95	0.95	0.05	<b>F</b>	<b>F</b>	<b>F</b>
Repeat 2014	0.12	0.12	0.88	0.06	0.05	0.94	0.08	0.08	0.92	0.76	0.76	0.33	1	1	-2.3
Increasing	0.29	0.86	0.13	0.32	0.67	0.5	0.01	0.5	1.67	0.53	0.77	0.47	0	0.5	1.3
Decreasing	0.17	0.17	0.82	0.16	0.16	0.8	0.1	0.1	0.9	1	1	-4.1	<b>F</b>	<b>F</b>	<b>F</b>
Linear 1	0.31	0.94	4.04	0.36	0.93	3.9	0.19	0.91	5.2	0.83	0.98	0.83	0.65	0.99	2.19
Linear 2	0.37	0.91	2.45	0.43	0.92	2.7	0.19	0.88	3.8	0.77	0.97	0.79	1	1	-3.01
Linear 3	0.26	0.93	5.6	0.25	0.93	5.4	0.17	0.92	6.5	0.69	0.97	2.08	1	1	-2.2
Serfling 1	0.56	0.56	0.23	0.28	0.28	0.7	0.15	0.15	0.9	0.74	0.74	0.2		1	-8.2
Serfling 2	0.4	0.4	0.52	0.37	0.37	6	0.44	0.44	0.5	1	1	-0.7	1	1	-6.3
Serfling 3	0	0.9	3.08	0	0.9	3.7	0	0.9	3.9	0.02	0.9	1.01	1	1	-1.2

$ng=34$															
Kind of Shape	WCR, Fixed rate=1			WCR, Fixed rate=3			WCR, Fixed rate=5			CUSUM			MEM		
	Sen.	Sp.	MED	Sen.	Sp.	MED	Sen.	Sp.	MED	Sen.	Sp.	MED	Sen.	Sp.	MED
Historical	0.08	0.08	0.92	0.04	0.04	0.96	0.12	0.12	0.88	0.95	0.95	0.05	<b>F</b>	<b>F</b>	<b>F</b>
Repeat 2014	0.08	0.08	0.92	0.04	0.04	0.96	0.18	0.18	0.82	0.54	0.54	0.96	1	1	-2.4
Increasing	0.23	0.77	-1.2	0.32	0.84	-5.6	0.34	0.84	-5.7	0.98	0.99	-0.06	0	0.76	3
Decreasing	0.13	0.13	0.87	0.08	0.08	0.92	0.12	0.12	0.88	1	1	-4.4	<b>F</b>	<b>F</b>	<b>F</b>
Linear 1	1	1	0	0.91	0.99	0.34	0.47	0.94	3	0.77	0.97	1.1	0.65	0.96	2.2
Linear 2	0.99	0.99	0.01	0.95	0.99	0.3	0.57	0.94	1.7	0.85	0.99	0.62	1	1	-2.97
Linear 3	1	1	0	0.92	0.99	0.7	0.37	0.94	4.7	0.68	0.97	2	1	1	-2.2
Serfling 1	0.58	0.58	0.22	0.27	0.27	0.69	0.29	0.29	0.7	0.71	0.71	0.27	1	1	-8.14
Serfling 2	0	0	1	0	0	1	0	0	1	0.1	0.1	0.9	1	1	-1
Serfling 3	0.01	0.01	0.99	0.01	0.01	0.99	0	0	1	0.04	0.04	0.96	0.01	1	-29.8

Table 9.8: Comparing performance of WCR, when  $ng=14, 34$  and  $f_r = 1, 3, 5$  with  $\alpha = 5\%$ , with CUSUM and MEM in different scenarios, using 100 simulations. Sen. means sensitivity, Sp. means specificity, MED means median time to detection, and E.S. (epidemic start) defines the approach of using the alert week (here the alert week is considered as the first week when the rate is at least 10% higher than in week 1).

- When we compare *WCR* with different  $fr = 1, 3, 5$ , keeping the alert week as the first week with a rate at least 10% higher than in the first week, we found the sensitivity and specificity are sometimes higher (there is no clear trend) when  $fr$  is higher and good performance in *MED* is maintained. *WCR* tends to raise alarms early.
- With MEM, we found generally high sensitivity and specificity, but with clear delay in alarms (*MED*).
- When we compare the performance of *WCR* with  $ng=14$  and 34, we found that *WCR* often performs better with a larger number of areas and this is most evident with sudden rises in the rate.

#### 9.4.1 Testing *WCR* with different alert weeks

Now, we will test *WCR* with different fixed rates  $fr = 1, 3, 5$ , and compare the performance with CUSUM and MEM, as we did earlier in section 9.4 but changing the alert week to be the same as in the MEM setting (as we did in section 8.2.2). Table 9.9 shows the result of testing using this setting when  $ng=14$  and 34.

From table 9.9, we can mention some points below:

- The results are generally poorer than with the previous setting of the alert week.
- In general, the relationship between sensitivity and *MED* is inverse, which means that when the sensitivity is high, we will lose some time from *MED*, so the delay is longer.
- CUSUM and *WCR* (in its different versions) almost always raise the alarm before MEM, which may be counted as an advantage of CUSUM and *WCR*.
- MEM generally has high sensitivity and specificity compared to CUSUM and *WCR* in this setting.
- If we had an outbreak in the past years, MEM will never raise an alarm unless the outbreak data are discounted.
- In the historical and decreasing scenarios, MEM did not raise any alarm.
- When we did this simulation study to compare performance between these three systems, we considered *WCR* with different  $fr = 1, 3, 5$ , testing sensitivity and specificity each time and we did not repeat those tests with CUSUM and MEM. The reason for this is that the change in the system affects *WCR* only.
- When we compare the performance of *WCR* with  $ng=14$  and 34, we found that *WCR* performs better with large  $ng$  with the first linear scenario.

$ng=14, \alpha = 5\%, E.S. = MEM$															
Kind of Shape	WCR, Fixed rate=1			WCR, Fixed rate=3			WCR, Fixed rate=5			CUSUM			MEM		
	Sen.	Sp.	MED	Sen.	Sp.	MED	Sen.	Sp.	MED	Sen.	Sp.	MED	Sen.	Sp.	MED
Historical	0	0.9	2.9	0	0.9	3	0	0.9	3	0	0.9	2.1	<b>F</b>	<b>F</b>	<b>F</b>
Repeat 2014	0	0.9	2.9	0	0.9	3	0	0.9	3	0	0.9	2.3	1	1	-0.3
Increasing	0.31	0.96	-8.9	0.12	0.91	1.5	0	0.9	2.6	0	0.9	1.5	0	0.9	2.3
Decreasing	0	0.9	2.9	0	0.9	2.9	0	0.9	2.9	1	1	-2.1	<b>F</b>	<b>F</b>	<b>F</b>
Linear 1	0.34	0.93	0.39	0.28	0.93	4.3	0.2	0.92	5	0.8	0.98	0.9	0.69	0.97	1.6
Linear 2	0	0.9	5.7	0	0.9	5.6	0	0.9	6	0	0.9	3.7	0.99	0.99	-0.06
Linear 3	0	0.9	8	0	0.9	8.8	0	0.9	9.2	0	0.9	4.5	0.15	0.92	0.9
Serfling 1	0	0.9	8.4	0	0.9	8.7	0	0.9	9.7	0	0.9	8.4	1	1	-0.2
Serfling 2	0	0.9	7.5	0	0.9	7.4	0	0.9	7.5	0	0.9	6.8	0.26	0.93	0.7
Serfling 3	0	0.9	5	0	0.9	5.7	0	0.9	6	0	0.9	3	0.15	0.92	0.9

$ng=34$															
Kind of Shape	WCR, Fixed rate=1			WCR, Fixed rate=3			WCR, Fixed rate=5			CUSUM			MEM		
	Sen.	Sp.	MED	Sen.	Sp.	MED	Sen.	Sp.	MED	Sen.	Sp.	MED	Sen.	Sp.	MED
Historical	0	0.89	2.9	0	0.98	4	0	0.98	2.9	0	0.89	2.2	<b>F</b>	<b>F</b>	<b>F</b>
Repeat 2014	0	0.89	3.9	0.34	0.95	-5.9	0	0.89	2.8	0	0.89	3.5	0.38	0.93	0.6
Increasing	0.3	0.92	-0.2	0.34	0.95	-5.9	0.35	0.95	-7.2	1	0.99	-0.1	0	0.89	3
Decreasing	0	0.89	2.9	0	0.89	2.9	0	0.89	3.9	1	1	-2.4	<b>F</b>	<b>F</b>	<b>F</b>
Linear 1	1	1	0.95	0.95	0.99	0.3	0.45	0.95	3.1	0.8	0.98	1	0.56	0.95	2.4
Linear 2	0	0.89	3	0	0.89	3	0	0.89	6.2	0	0.98	3.6	0.98	0.99	0
Linear 3	0	0.89	3.2	0	0.89	3	0	0.89	7.2	0	0.89	4.3	0.15	0.91	0.9
Serfling 1	0	0.89	9.3	0	0.89	9.8	0	0.89	9.8	0	0.89	9.3	0.91	0.91	0.8
Serfling 2	0	0.89	2	0	0.89	2	0	0.89	2	0	0.89	1.9	1	1	0
Serfling 3	0	0.89	29	0	0.89	0.29	0	0.89	29	0	0.89	29	0.99	0.99	-1.8

Table 9.9: Comparing performance of WCR when  $ng=14, 34$  and  $fr = 1, 3, 5$  with  $\alpha = 5\%$ , with CUSUM and MEM in different scenarios, using 100 simulations. Sen. means sensitivity, Sp. means specificity, MED means median time to detection, and E.S. defines the approach of using the alert week (here the alert week is considered as the same as in the MEM setting).

- When we compare *WCR* with different  $fr = 1, 3, 5$  and CUSUM, keeping the alert week to equal the setting of MEM in section 3.3, we found the sensitivity is generally low though specificity is high. Again, we generally found higher sensitivity and specificity with an improvement in alarms (*MED*) with MEM compared to the previous setting of the alert week.

## 9.5 Improving *WCR* performance

In the previous section 9.4, we found that the *WCR* method has poor performance (meaning low sensitivity), either using the alert week (week of increase) as the first week when the rate is at least 10% higher than in week 1, or the same as in the MEM setting in section 3.3.

We can change the sensitivity by changing the significance level of the test.

Therefore we think about decreasing the significance level  $\alpha$  in our hypothesis test in section 6.1 from 5% to 1% (this means only 1% of simulations from the large simulation in section 5.2 would be considered as not coming from the steady state). We mentioned the *WCR* system coverage in table 9.7 with  $\alpha=1\%$ . The reason for decreasing the significance level  $\alpha$  to 1% and not increasing it (say to 10%) is because we are looking to minimize the rejection region, to avoid false alarms.

Figure 9.10 shows a comparison in performance between *WCR* (blue triangles), CUSUM (green dots) and MEM (red crosses) in the top figure, and an overview of the scenario used in this test is presented in the bottom figure. This figure represents the dataset of  $ng=14$ . Week 10 is the first week of a jump in the rate either using the alert week (week of increase) as the first week when the rate is at least 10% higher than in week 1, or the same as in the MEM setting in section 3.3. Week 1 to week 9 should be considered as steady rates because they were simulated from the steady state, so all the alarms raised before week 10 are false alarms. We can see clearly from the graph that the majority of the false alarms are with systems CUSUM and MEM, based on the 100 simulated cases. In the *WCR* system, we see that around 90% of the alarms are correct and not as in the two systems CUSUM and MEM. *WCR* registered 10 false alarms, while CUSUM registered 24 false alarms and 39 false alarms were registered when using MEM.



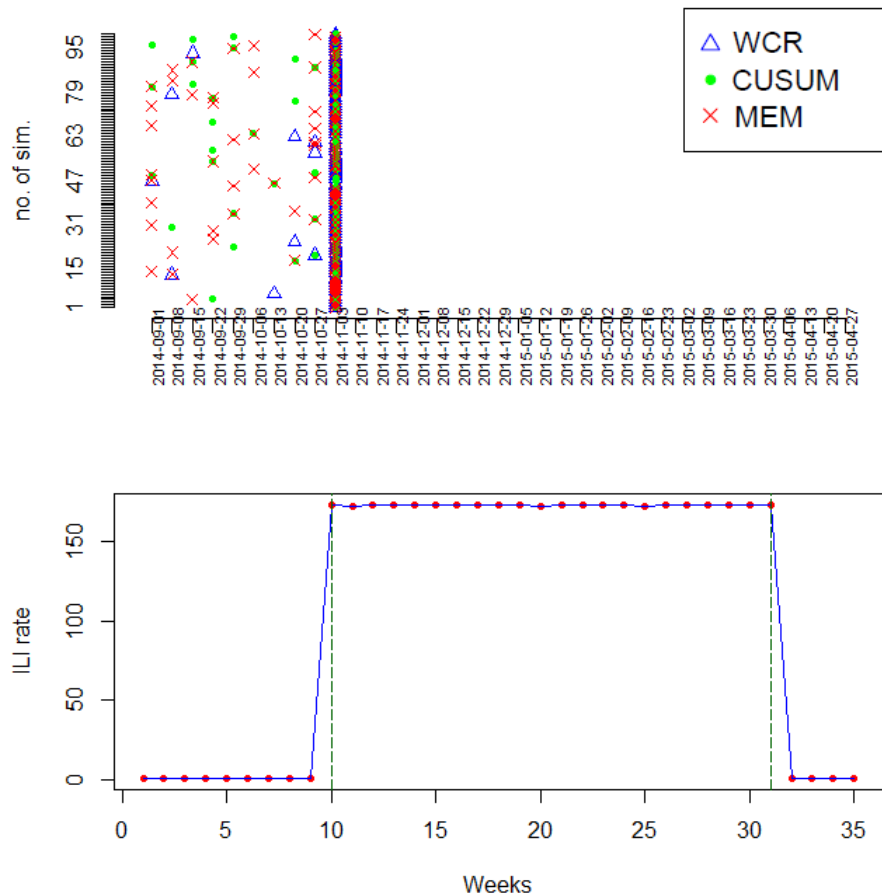


Figure 9.10: The top figure shows a comparison of detection of alert weeks between *WCR* (blue triangles), *CUSUM* (green dots) and *MEM* (red crosses). The bottom figure shows the scenario used to clarify the location of the jump in rates; it shows the mean of the 100 simulated cases. Here  $ng=14$ .

We show another example when  $ng=34$  and try a different scenario in figure 9.11. In this example, we can see the difference between the two ways used to set the epidemic start. When we consider the *MEM* structure (the first green dotted line is the epidemic start), we will consider all *WCR* alarms and the majority of the *CUSUM* alarms as false alarms as they are before the alert week. Also when we using alert week (week of increase) as the first week when the rate is at least 10% higher than in week 1, we will consider all *MEM* alarms as delayed alarms as they are after the alert week. When we consider the alert week (week of increase) as the first week when the rate is at least 10% higher than in week 1, *WCR* did not register any false alarms, while *CUSUM* registered 20 false alarms and all *MEM* alarms are delayed alarms. But when we consider the alert week as the same as in the *MEM* setting in section 3.3, all *WCR* and *CUSUM* alarms are false alarms and *MEM* registered only 3 false alarms.

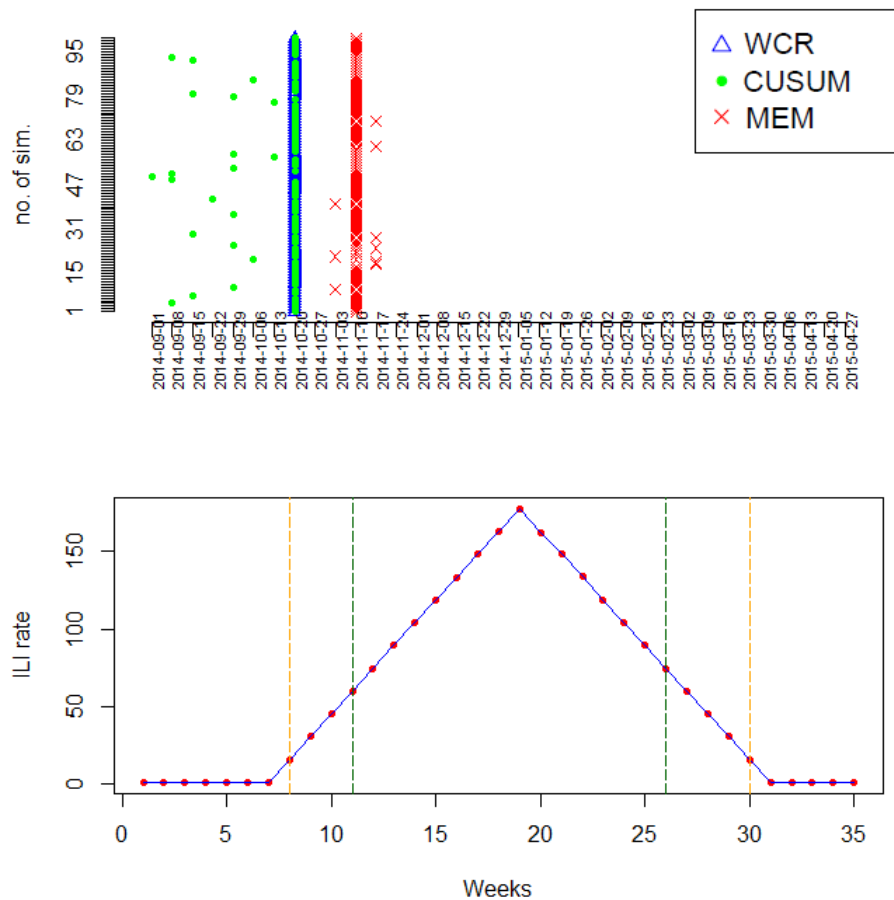


Figure 9.11: The top figure shows a comparison of detection of alert weeks between *WCR* (blue triangles), *CUSUM* (green dots) and *MEM* (red crosses). The bottom figure shows the scenario used to clarify the location of the gradual increase in rates; it shows the mean of the 100 simulated cases. Here  $ng=34$ . Green dotted lines represent the start and the end sensitivity area according to *MEM* setting while orange dotted lines represent the alert week (week of increase) as the first week when the rate is at least 10% higher than in week 1.

We will start as we did at the start of section 9.4 and we can find the week of the epidemic by choosing a specific rate of increase compared to the first week, so all those weeks from week 1 to this specific week of increase will be considered for calculation of sensitivity and those weeks after that will be considered in the calculation of specificity. Firstly, we will consider the alert week as the first week when the rate is at least 10% higher than in week 1.

Table 9.10 shows the performance of the *WCR* method when  $ng=14$  and 34 and  $fr = 1, 3, 5$  with  $\alpha = 1\%$  for *CUSUM* and *MEM* in different scenarios. From this table, we mention some points below:

$ng=14, \alpha = 1\%, E.S. = 10\%$															
Kind of Shape	WCR, Fixed rate=1			WCR, Fixed rate=3			WCR, Fixed rate=5			CUSUM			MEM		
	Sen.	Sp.	MED	Sen.	Sp.	MED	Sen.	Sp.	MED	Sen.	Sp.	MED	Sen.	Sp.	MED
Historical	0.36	0.37	0.12	0.26	0.26	0.73	0.24	0.24	0.73	0.91	0.91	0.09	<b>F</b>	<b>F</b>	<b>F</b>
Repeat 2014	0.27	0.27	0.65	0.23	0.23	0.75	0.26	0.26	0.72	0.68	0.68	0.32	1	1	-2.32
Increasing	0.33	0.97	-21.61	0.4	0.85	-8.4	0.18	0.59	1.18	0.55	0.78	0.45	0	0.5	1.29
Decreasing	0.24	0.24	0.71	0.22	0.22	0.76	0.18	0.18	0.8	1	1	-4.12	<b>F</b>	<b>F</b>	<b>F</b>
Linear 1	0.91	0.99	0.41	0.58	0.95	2.06	0.48	0.94	3.2	0.8	0.99	0.9	0.69	0.99	1.58
Linear 2	0.93	0.99	0.27	0.67	0.96	1.38	0.51	0.93	2.08	0.84	0.99	0.73	1	1	-3.06
Linear 3	0.92	0.99	0.39	0.55	0.96	3.03	0.49	0.95	3.6	0.74	0.99	1.53	1	1	-2.15
Serfling 1	0.71	0.71	-0.55	0.42	0.42	0.49	0.42	0.42	0.47	0.65	0.65	0.34		1	-8.18
Serfling 2	0.76	0.76	-0.05	0.64	0.64	0.16	0.59	0.59	0.17	1	1	-0.53	1	1	-6.26
Serfling 3	0.04	0.9	2.16	0	0.9	3.21	0	0.9	3.2	0.01	0.9	1.02	1	1	-1.15

$ng=34$															
Kind of Shape	WCR, Fixed rate=1			WCR, Fixed rate=3			WCR, Fixed rate=5			CUSUM			MEM		
	Sen.	Sp.	MED	Sen.	Sp.	MED	Sen.	Sp.	MED	Sen.	Sp.	MED	Sen.	Sp.	MED
Historical	0.26	0.26	0.69	0.35	0.35	0.62	0.47	0.47	0.46	0.92	0.92	0.02	<b>F</b>	<b>F</b>	<b>F</b>
Repeat 2014	0.36	0.36	0.57	0.34	0.34	0.6	0.29	0.29	0.67	0.58	0.58	0.42	1	1	-2.3
Increasing	0.3	0.84	-6.35	0.27	0.88	-12	0.24	0.9	-15	0.98	0.99	-0.03	0	0.67	3
Decreasing	0.39	0.39	0.56	0.27	0.27	0.69	0.32	0.32	0.63	1	1	-4.5	<b>F</b>	<b>F</b>	<b>F</b>
Linear 1	1	1	0	1	1	0	0.81	0.89	1	0.77	0.97	1.08	0.66	0.96	1.7
Linear 2	1	1	0	1	1	0	0.9	0.99	0.5	0.79	0.97	0.82	1	1	-3
Linear 3	1	1	0	1	1	0	0.75	0.98	1.8	0.7	0.99	1.9	1	1	-2.2
Serfling 1	0.76	0.76	-0.81	0.56	0.56	0.32	0.38	0.38	0.6	0.76	0.76	0.24	1	1	-8.2
Serfling 2	0	0	1	0	0	1	0	0	1	0.11	0.11	0.89	1	1	-1
Serfling 3	0.03	0.03	0.97	0.01	0.01	0.99	0.03	0.97	0.97	0.01	0.01	0.99	0	1	-30

Table 9.10: Comparing performance of WCR, when  $ng=14, 34$  and  $fr = 1, 3, 5$  with  $\alpha = 1\%$ , with CUSUM and MEM in different scenarios, using 100 simulations. Sen. means sensitivity, Sp. means specificity, MED means median time to detection, and E.S. defines the approach of using the alert week (here the alert week is considered as the first week when the rate is at least 10% higher than in week 1).

- In the historical scenario, CUSUM performance is much the best. MEM fails here because its warnings depend on the average rates in the past 5 years. There was an outbreak in 2009, which caused a rise in the average of the past years and thus a loss of alerts before the epidemic start. *WCR* is relatively poor.
- In the scenario of repeating the 2014 cases six times, as for the historical scenario *WCR* is much poorer than CUSUM, while MEM shows the best, high sensitivity and specificity but with a delay in its alarm.
- MEM does not seem to have delay in the increasing situation, either for  $ng=14$  or  $ng=34$  *WCR* has delays in this situation except for  $fr=5$ ,  $ng=14$ .
- In the decreasing scenario, MEM fails completely. *WCR* and CUSUM do not show perfect detection as the sensitivity and specificity are low in *WCR*, and although they are perfect with CUSUM there is a delay in the alarm for CUSUM.
- With the shape of the third Serfling scenario for  $ng=14$ , *WCR* and CUSUM show low sensitivity and high specificity. This is also true for *WCR* with  $fr=5$  for  $ng=34$ . For the other *WCR* with  $ng=34$  and CUSUM with  $ng=34$  both sensitivity and specificity are low. MEM again here shows high specificity for  $ng=14$  and  $ng=34$  but with delay in its alarm. It has high sensitivity for  $ng=14$  only.
- With the shape of the second and third linear scenarios for  $ng=14$  and  $ng=34$  and for the first Serfling scenario with  $ng=14$ , *WCR* with  $fr=1$  is better than CUSUM. For the first Serfling scenario with  $ng=34$  they are the same in sensitivity and specificity but *WCR* alarms with a slight delay whilst CUSUM alarms slightly late. In all cases MEM has performed sensitivity and specificity, but with a delay in its alarms. So overall *WCR* with  $fr=1$  is the best in this case.
- With the shape of the second Serfling scenario, MEM shows high sensitivity and specificity but with delay in alarms, as does CUSUM with  $ng=14$ , while *WCR* shows the worst performance.
- With the shape of the third Serfling scenario, for  $ng=14$ , *WCR* and CUSUM show low sensitivity and high specificity. This is also true for *WCR* with  $fr=5$  and  $ng=34$ . For other *WCR* with  $ng=34$  and CUSUM with  $ng=34$  both sensitivity and specificity are low. MEM again here shows high specificity but with a delay in its alarm. It has high sensitivity for  $ng=14$  only.
- When we compare *WCR* with different  $fr=1, 3, 5$ , keeping the alert week as the first week with a rate at least 10% higher than in the first week, we found the sensitivity and specificity are higher when  $fr$  is 1 and MEM is generally good. *WCR* has delays in this situation except for  $fr=5$ ,  $ng=14$ .

- With MEM, we found generally high sensitivity and specificity, but with clear delay in alarms (*MED*).
- When we compare the performance of *WCR* with  $ng=14$  and  $34$ , we found that *WCR* performs generally, but not always, slightly better with the larger number of areas and this is most evident in the linear cases.

We can see that using  $\alpha = 1\%$  (table 9.10) has led to a marked improvement in performance of *WCR*, especially in case  $ng=14$ , compared to the results in section 9.4 where  $\alpha = 5\%$  was used (table 9.8).

We can test *WCR* in different cases of  $fr = 1, 3, 5$  and compare the performance with CUSUM and MEM, as we did earlier in section 9.4.1, considering the alert week to be the same as in the MEM setting in section 3.3.

From table 9.11, we mention some points below:

- Again the results are generally poorer than with the previous approach to defining the alert week.
- In general, the relationship between sensitivity and *MED* is inverse, which means that when the sensitivity is high, we will lose some time from *MED*, so the delay is longer.
- CUSUM and *WCR* (in its different versions) almost always raise the alarm before MEM, which may be counted as an advantage of CUSUM and *WCR*. *WCR* does show delay in the increasing scenario.
- MEM often has high sensitivity and specificity compared to CUSUM and *WCR* in this setting.
- If we had an outbreak in the past years, MEM will never raise an alarm unless the outbreak data are discounted.
- In the historical and decreasing scenarios, MEM did not raise any alarm.
- When we did this simulation study to compare performance between these three systems, we considered *WCR* with different  $fr = 1, 3, 5$ , testing sensitivity and specificity each time and we did not repeat those tests with CUSUM and MEM. The reason for this is that the change in the system affects *WCR* mainly and there are only minor effects with CUSUM and MEM.
- When we compare the performance of *WCR* with  $ng=14$  and  $34$ , we found that the *WCR* performs slightly better with large  $ng$  with the first linear scenario.
- When we compare *WCR* with different  $fr = 1, 3, 5$  and CUSUM, keeping the alert

$ng=14, \alpha = 1\%, E.S. = MEM$															
Kind of Shape	<i>WCR</i> , Fixed rate=1			<i>WCR</i> , Fixed rate=3			<i>WCR</i> , Fixed rate=5			<i>CUSUM</i>			<i>MEM</i>		
	Sen.	Sp.	<i>MED</i>	Sen.	Sp.	<i>MED</i>	Sen.	Sp.	<i>MED</i>	Sen.	Sp.	<i>MED</i>	Sen.	Sp.	<i>MED</i>
Historical	0.02	0.9	2.6	0.01	0.9	2.7	0.01	0.9	2.8	0.05	0.9	2.2	<b>F</b>	<b>F</b>	<b>F</b>
Repeat 2014	0.03	0.9	2.2	0.01	0.9	2.6	0	0.9	2.7	0	0.9	2.4	1	1	-0.4
Increasing	0.26	0.98	-18	0.17	0.95	-10	0.03	0.9	2	0	0.9	1.5	0	0.9	2.4
Decreasing	0.02	0.9	2.6	0	0.9	2.7	0	0.9	2.7	1	1	-2.1	<b>F</b>	<b>F</b>	<b>F</b>
Linear 1	0.91	0.99	0.6	0.7	0.97	1.6	0.65	0.96	2	0.8	0.98	0.9	0.70	0.97	1.7
Linear 2	0	0.9	3.1	0	0.9	4.6	0	0.9	5.5	0	0.9	3.6	0.99	0.99	-0.05
Linear 3	0	0.9	3.9	0	0.9	5.6	0	0.9	7.3	0	0.9	4.3	0.15	0.92	1
Serfling 1	0	0.9	7.5	0	0.9	8.5	0	0.9	8.5	0	0.9	8.6	1	1	-0.3
Serfling 2	0	0.9	6.9	0	0.9	6.5	0	0.9	7.3	0	0.9	6.7	0.27	0.93	0.8
Serfling 3	0	0.9	4.2	0	0.9	5	0	0.9	5.4	0	0.9	3.3	0.15	0.92	0.9

$ng=34$															
Kind of Shape	<i>WCR</i> , Fixed rate=1			<i>WCR</i> , Fixed rate=3			<i>WCR</i> , Fixed rate=5			<i>CUSUM</i>			<i>MEM</i>		
	Sen.	Sp.	<i>MED</i>	Sen.	Sp.	<i>MED</i>	Sen.	Sp.	<i>MED</i>	Sen.	Sp.	<i>MED</i>	Sen.	Sp.	<i>MED</i>
Historical	0	0.89	3.8	0	0.98	3.6	0	0.98	2.7	0	0.89	2.9	<b>F</b>	<b>F</b>	<b>F</b>
Repeat 2014	0	0.89	2.5	0.35	0.95	2.5	0	0.89	2.6	0	0.89	2.5	1	1	-0.3
Increasing	0.32	0.95	-6	0.28	0.96	-11	0.38	0.98	-15.5	1	1	-0.1	0	0.89	2.9
Decreasing	0	0.89	2.5	0.01	0.89	2.6	0	0.87	3.6	1	1	-2.6	<b>F</b>	<b>F</b>	<b>F</b>
Linear 1	1	1	0	0.95	0.99	0.3	0.83	0.99	0.8	0.8	0.98	1.1	0.56	0.95	2.3
Linear 2	0	0.89	2.9	0	0.89	3.8	0	0.89	3.5	0	0.98	3.9	0.05	0.89	1
Linear 3	0	0.89	3.1	0	0.89	3	0	0.89	4.2	0	0.89	4.5	0.23	0.91	0.8
Serfling 1	0	0.89	8.5	0	0.89	9.4	0	0.89	9.5	0	0.89	9.2	0.21	0.91	0.8
Serfling 2	0	0.89	2	0	0.89	1.8	0	0.89	1.9	0	0.89	2	1	1	0
Serfling 3	0	0.89	29	0	0.89	0.29	0	0.89	29	0	0.89	29	0.99	0.99	-1.9

Table 9.11: Comparing performance of *WCR* when  $ng=14, 34$  and  $fr = 1, 3, 5$  with  $\alpha = 1\%$ , with *CUSUM* and *MEM* in different scenarios, using 100 simulations. Sen. means sensitivity, Sp. means specificity, *MED* means median time to detection, and E.S. defines the approach of using the alert week (here the alert week is considered as the same as in the *MEM* setting).

week the same as for the setting of MEM in section 3.3, we found the sensitivity is generally low though specificity is high. Again, we generally found higher sensitivity and specificity with delay in alarms (*MED*) with MEM when the alert week is taken as the first week with a rate 10% above the rate in week 1.

Using this approach to defining the alert week, there is no longer a marked improvement in *WCR* using  $\alpha = 1\%$  (table 9.11) compared to the results in section 9.4.1 where  $\alpha = 5\%$  (table 9.9), except for the first linear case.

## 9.6 Conclusion

In this ninth chapter of this thesis, we made some big changes in the *WCR* system, trying to improve its performance. These changes are summarised below:

- We tried a new baseline of simulation using fixed and varying rates as baselines when we simulated 3,110,000 cases. This affected the hypothesis test of the steady state.
- In this chapter, we investigated the symmetry of *WCR* in each  $N_{HB}$  from 0 to  $\frac{N_{HB}}{2} - 1$  compared to these in each  $N_{HB}$  from  $\frac{N_{HB}}{2} - 1$  to the end of  $N_{HB}$ . We confirm in section 9.1 that they are not symmetric and we considered this in the same section.
- We tried to get equations for the mean and variance of *WCR*, aiming to calculate cut-off points for *WCR* using  $\mu$  and  $\sigma$ , but results of these equations were not as helpful as we expected.
- We tested the performance of the *WCR* system with its new setting using different significance levels for the rejection region, as we used 1% and 5%. Using the significance level of 1% seems to be more useful for *WCR* in general, especially with a linear increase in the rate.

The *WCR* method does not look to be a good candidate to replace CUSUM or MEM, despite being based upon a bivariate distribution. In the next final chapter, we will discuss some conclusions and possible further work related to this thesis.

## Chapter 10

# Conclusions and further work

In this final chapter, we will talk about some results and make some suggestions for future work for anyone who is interested in early detection systems for pandemic and seasonal influenza.

In this thesis, chapter 1 contains background information about influenza, while the next chapter 2 is about description of data. In chapter 3, statistical methods for the detection of outbreaks are reviewed. Applying and modifying some well-known detection systems for influenza is presented in chapter 4. Chapters 5 and 6 give an overview of the proposed system. Different scenarios for the simulation study are reviewed in chapter 7. Results of tests of the detection system and comparison with other systems are presented in chapters 8 and 9. Here chapter 10 draws conclusions.

The main aim of this research is trying to develop the *WCR* method for use to detect seasonal flu by establishing a simpler method of calculating the critical values (cut-offs). We have developed a method to set the cut-offs based on simulation of previous trends and it is possible to use the simulated data with an ellipse equation to derive cut-offs which could be used.

Our system has extended the *WCR* system of Singh et al. (2010) [142] and we have carried out simulations to examine its performance.

In particular the Singh et al. (2010) approach is to obtain a joint pmf for  $N_{HB}$  and a discretised version of *WCR*, using smoothed historical data, and then to trigger an alarm if an observed value of  $(WCR, N_{HB})$  corresponds to a low value of this pmf, lower than a specified threshold. This setting performed well with pandemic influenza but we look to modify it to work also with seasonal influenza.

In our system, we simulated huge numbers of joint frequencies for  $(WCR, N_{HB})$  in the case of the steady state, then we develop the *WCR* algorithm for different numbers of



spatial areas. We created another data structure using 30, 40 and 50 HBs from the original Scottish data, using different population size structures, then simulated more than 3,000,000 cases of ILI counts, considering the rate as flat during the year, then we got the joint distribution for  $WCR$  and  $N_{HB}$  in the case of no epidemic. Then we created a rejection region for a test of whether or not an observation of  $(WCR, N_{HB})$  comes from the steady state and used more than one significance level for this test (5% and 1%). We used more than one approach to set up our baseline of the steady state.

## 10.1 Conclusions

In this section, some general conclusions are listed, related to our work and the work of others:

- Influenza is one of the most important infectious diseases, which kills many people annually in the world, so it is an important topic to study for researchers.
- Influenza cases do not occur with the same pattern throughout the world but they are different in each region. However throughout the world, the influenza cases are reported in the winter more than in other seasons.
- There are many methods to monitor influenza cases and also there are many methods to use those monitored data to detect the next increase in influenza rates.
- Those detection methods are different in their performance.

Now, we will list some specific results related to our work in this thesis:

- In chapter 2, we observed in the Scottish data that we used during data analysis that ARI and ILI rates have the same patterns but because ILI is more specific to influenza, consultation rates for ILI are lower than for ARI, which includes all respiratory diseases. Both recorded a rise in rates in winter compared to the rest of the year. There was also a shortfall for both ARI and ILI (zeroes in consultation rates) on public holidays, when the GP surgeries were closed. Another point worth mentioning is that the population size varies widely between the regions of Scotland. Patterns of consultations were especially different in the smaller health boards.
- When we talk about approaches for choosing different baselines of simulation in chapter 5, section 5.2.1, we found it is not acceptable to use equal population sizes in different sub-regions when we use real ILI counts to simulate other ILI counts by using a Poisson mean. The reason for this is the big variation between population sizes in different sub-regions.
- In chapter 6, section 6.4, and also in chapter 9, section 9.2, all equations for the

$WCR$  mean and standard deviation in each  $N_{HB}$  did not work well when using a normal approximation for  $WCR$  to calculate cut-offs to use in the  $WCR$  system. This means we must do a new simulation with each change in number of HBs.

- Throughout our tests in chapter 8 and 9 of the performance of  $WCR$ , MEM and CUSUM, generally no one method outperformed the others in all cases that have been tested but there is a disparity in performance between the various methods.
- In cases of a sudden rise (which represents an outbreak rather than a seasonal increase each winter), the  $WCR$  method recorded remarkable superiority in its performance compared to other methods used in chapters 8 and 9. This was actually the reason why the  $WCR$  system of Singh et al. (2010) [142] was initially devised.
- In general, the  $WCR$  and CUSUM methods are similar in performance in different situations and this is clear in chapters 8 and 9.
- In all methods used, there is an inverse relationship between sensitivity and  $MED$  because the more careful we are to get early alarms, the more false alarms there are likely to be.

Our  $WCR$  method should not be used for seasonal influenza but it can still be used for pandemic detection, like the version of Singh et al. (2010) [142].

## 10.2 Further work

To extend the work done and improve its performance, we recommend to simulate millions of weekly cases rather than simulate thousands of weekly cases as this may affect the results.

Secondly, we recommend to try using varying rates in each HBs, with the emphasis not to use equal population sizes in each HBs but using an equivalent approach for the Scottish population.

Some factors such as sex, age, rurality were not considered in our work. Also potential differences between HBs because of these factors were not considered. Our modelling is limited because of these factors. It is a good recommendation to consider these factors if they are available.

We did not find equations for the mean and variance of  $WCR$  that worked especially well. This could be investigated further.

Updated ILI data could be used in the historical data set.

It could be useful to vary more the significance level  $\alpha$  used in the  $WCR$  system.

We did not make a direct comparison of our *WCR* approach to the *WCR* approach of Singh et al. (2010) [142] in our experiments. Both these two methods work well with an outbreak scenario, but it would be useful to do a comparison as we expect to get similar conclusions.

It would be interesting to do similar work based on the data of other countries when the data are available through collaboration with any health institution as we do not know if there any available data in the public domain, or using ARI data, or also using different infectious diseases to find out if similar conclusions would be reached.

As the weather changes in different places, from rainy weather to dusty weather in Saudi Arabia, for example, it could be interesting to try similar work based on data from those different environments.

# Appendix A

## Chapter 5: Overview of the system - Simulation

### A.1 Different spatial locations - Steady State

#### A.1.1 Investigation of Normality of $WCR$ : histograms of $WCR$

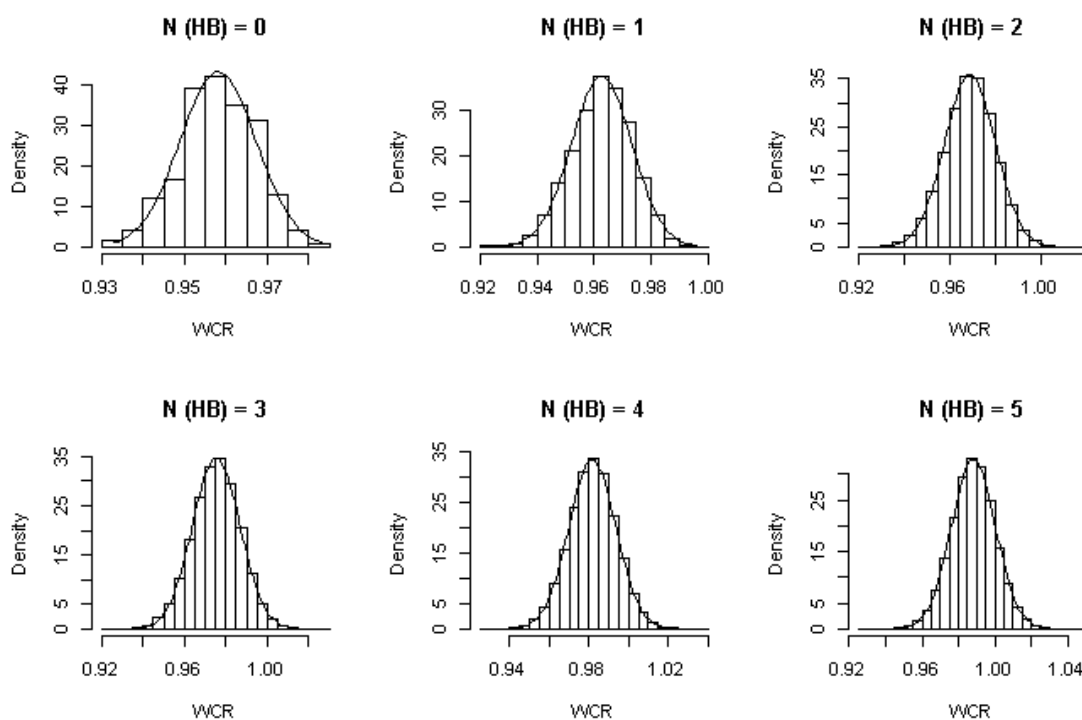


Figure A.1: Histograms of  $WCR$  for each  $N_{HB}$  in the case of HBs=14 with best fitting normal pdf. 3,110,000 simulated cases from the steady state are used here.

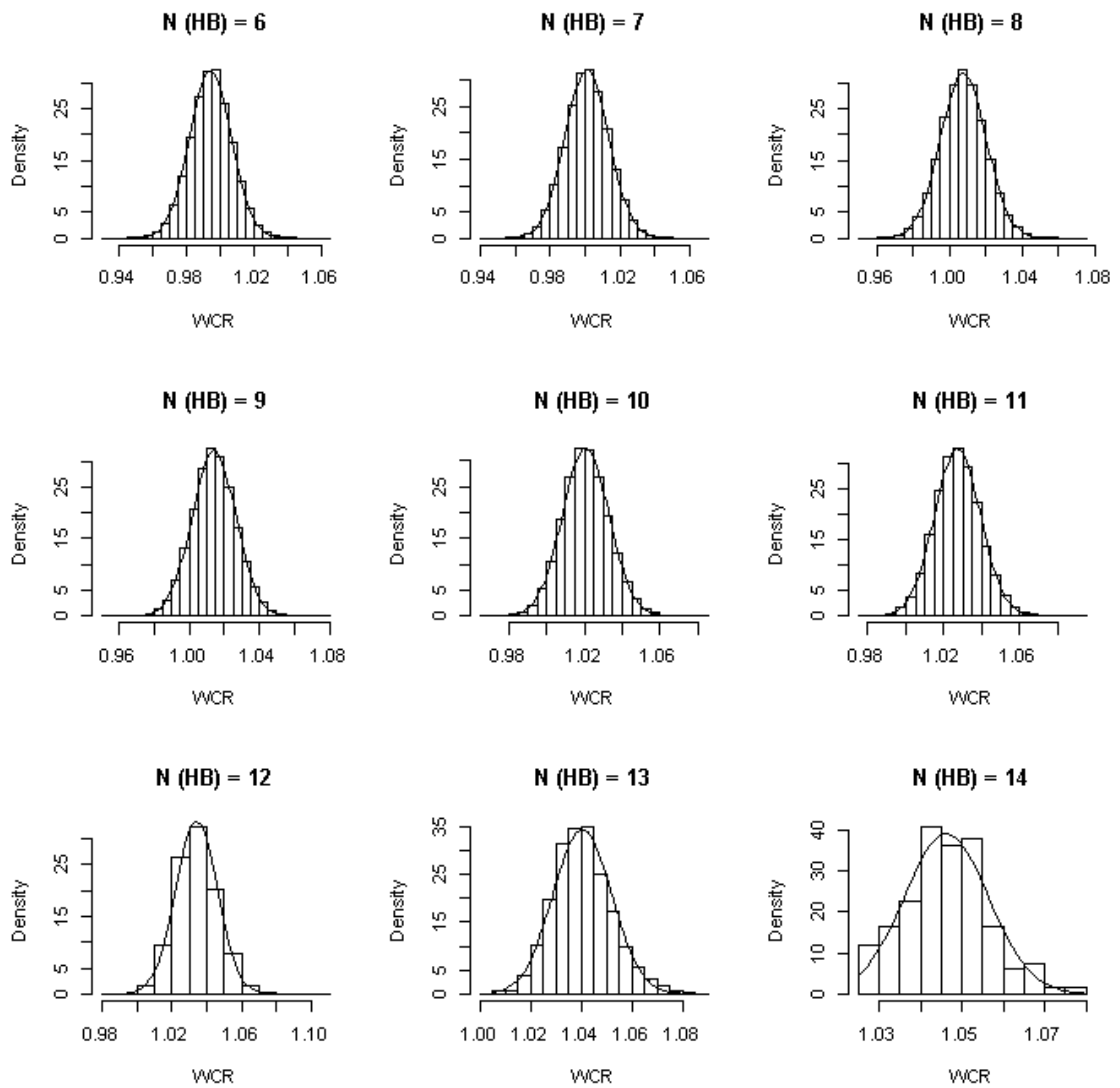


Figure A.2: Histograms of  $WCR$  for each  $N_{HB}$  in the case of HBs=14 with best fitting normal pdf. 3,110,000 simulated cases from the steady state are used here.

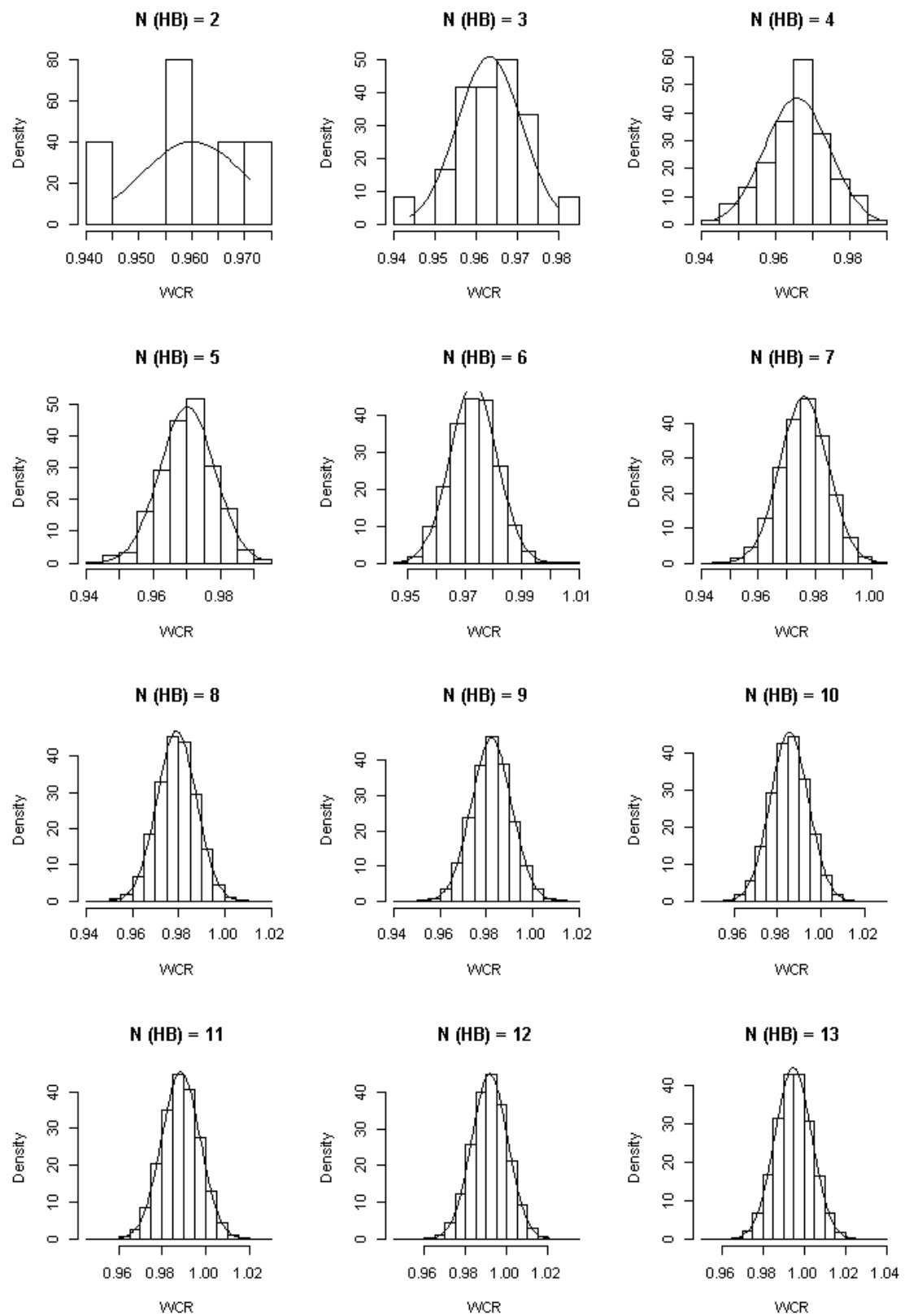


Figure A.3: Histograms of  $WCR$  for each  $N_{HB}$  in the case of HBs=30 with best fitting normal pdf. 3,110,000 simulated cases from the steady state are used here.  $WCR$  does not start from zero of  $N_{HB}$  because the joint distributions of  $WCR$  and  $N_{HB}$  in this cases do not start from zero, as shown in figure 5.2.

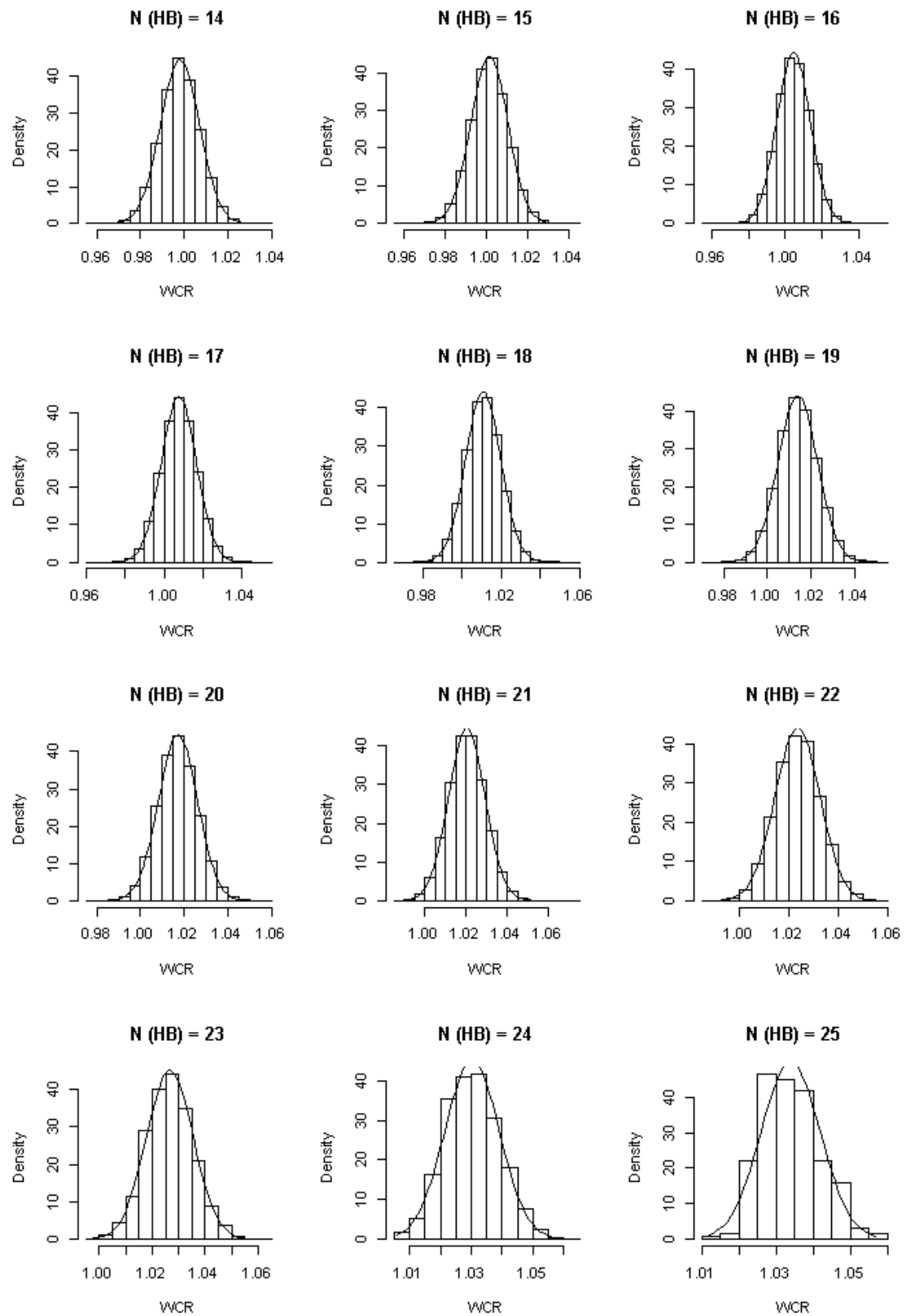


Figure A.4: Histograms of  $WCR$  for each  $N_{HB}$  in the case of HBs=30 with best fitting normal pdf. 3,110,000 simulated cases from the steady state are used here.

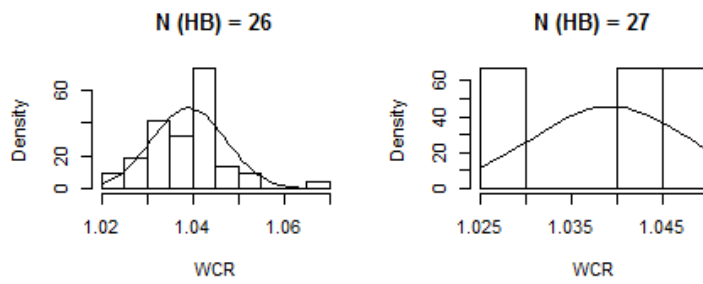


Figure A.5: Histograms of  $WCR$  for each  $N_{HB}$  in the case of  $HBs=30$  with best fitting normal pdf. 3,110,000 simulated cases from the steady state are used here.  $WCR$  does not extend to the end of  $N_{HB}$  because the joint distributions of  $WCR$  and  $N_{HB}$  in this cases do not extend to the end of  $N_{HB}$ , as shown in figure 5.2.



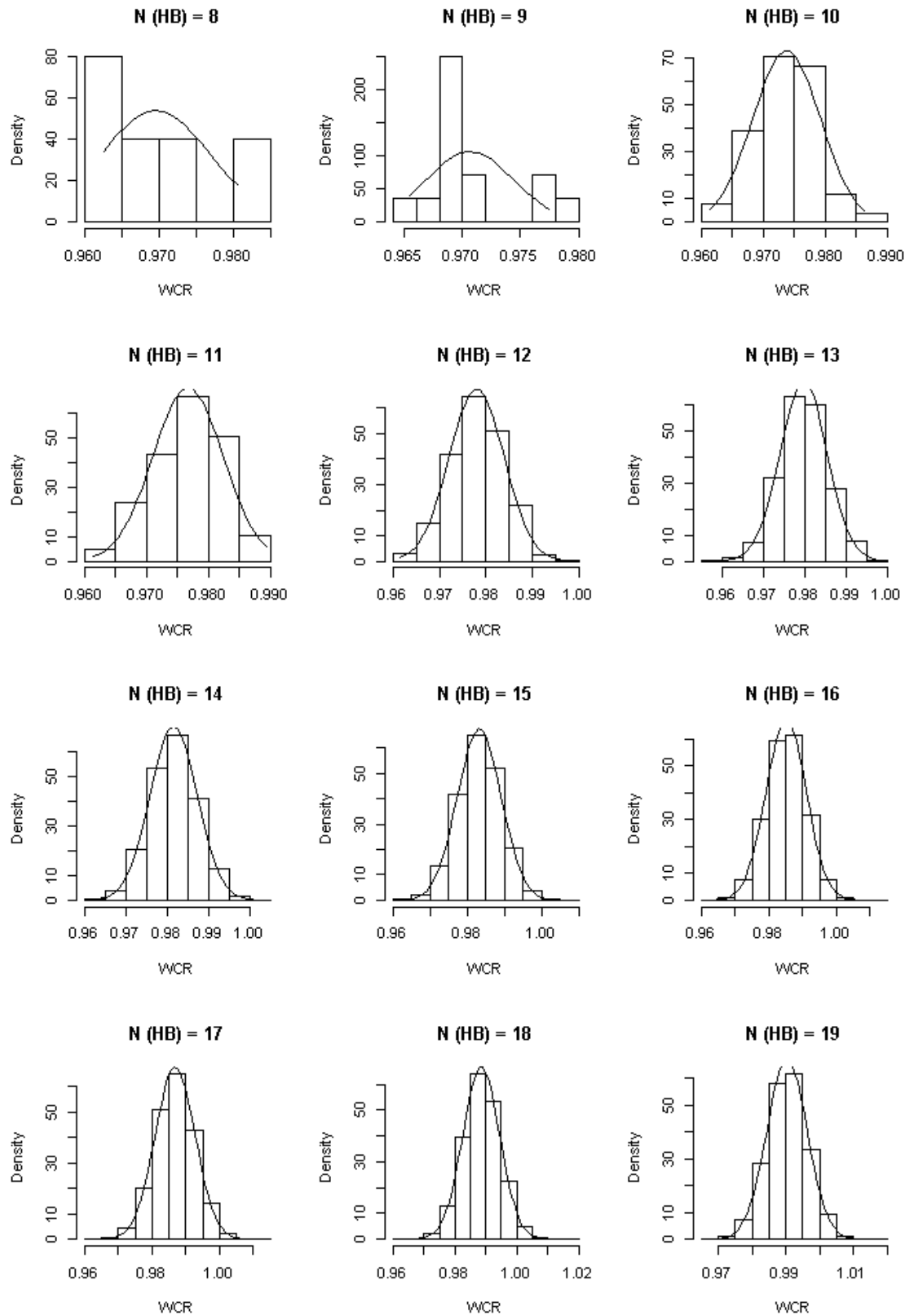


Figure A.6: Histograms of  $WCR$  for each  $N_{HB}$  in the case of  $HBs=50$  with best fitting normal pdf. 3,110,000 simulated cases from the steady state are used here.  $WCR$  does not start from zero of  $N_{HB}$  because the joint distributions of  $WCR$  and  $N_{HB}$  in these cases do not start from zero, as shown in figure 5.3.

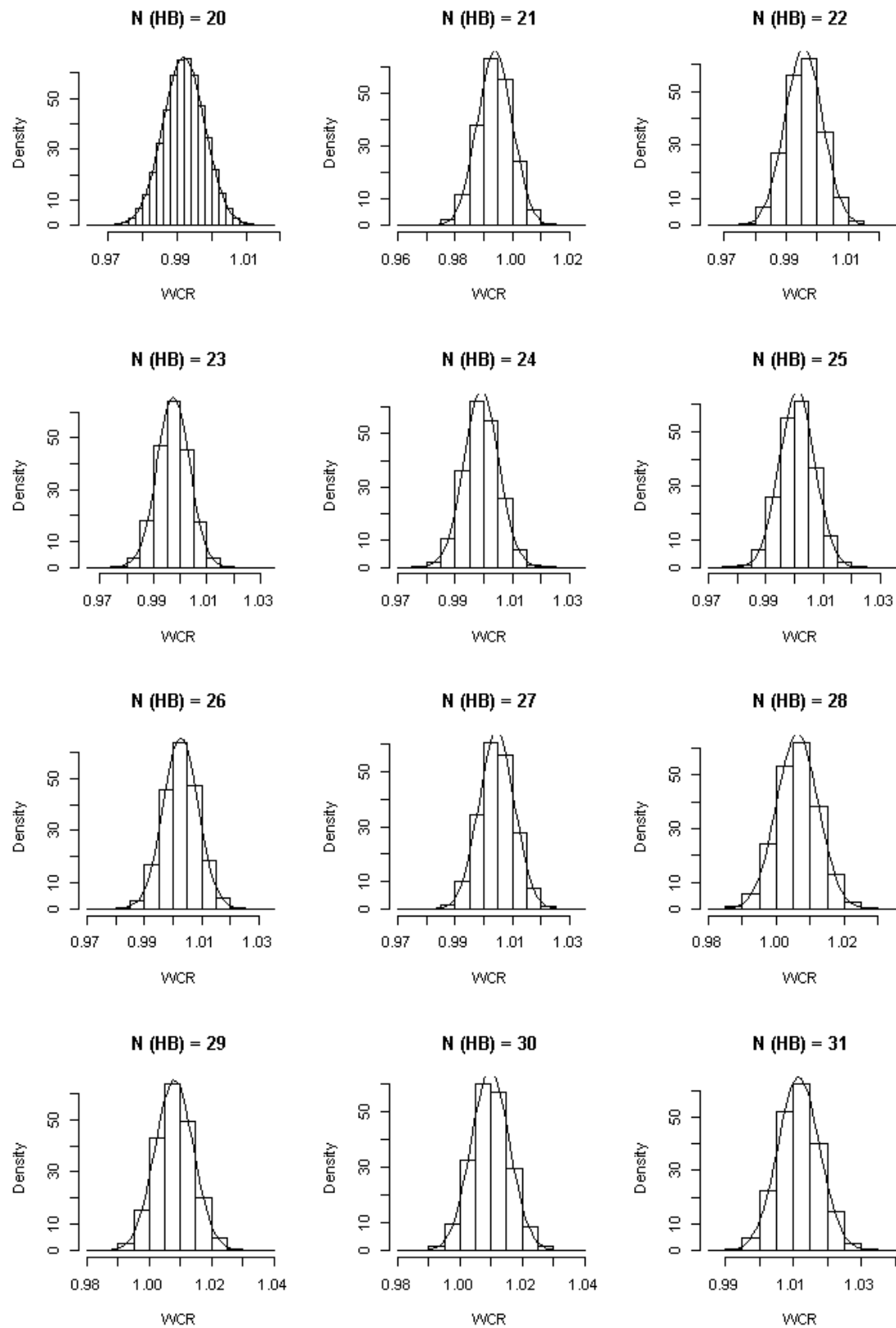


Figure A.7: Histograms of  $WCR$  for each  $N_{HB}$  in the case of HBs=50 with best fitting normal pdf. 3,110,000 simulated cases from the steady state are used here.

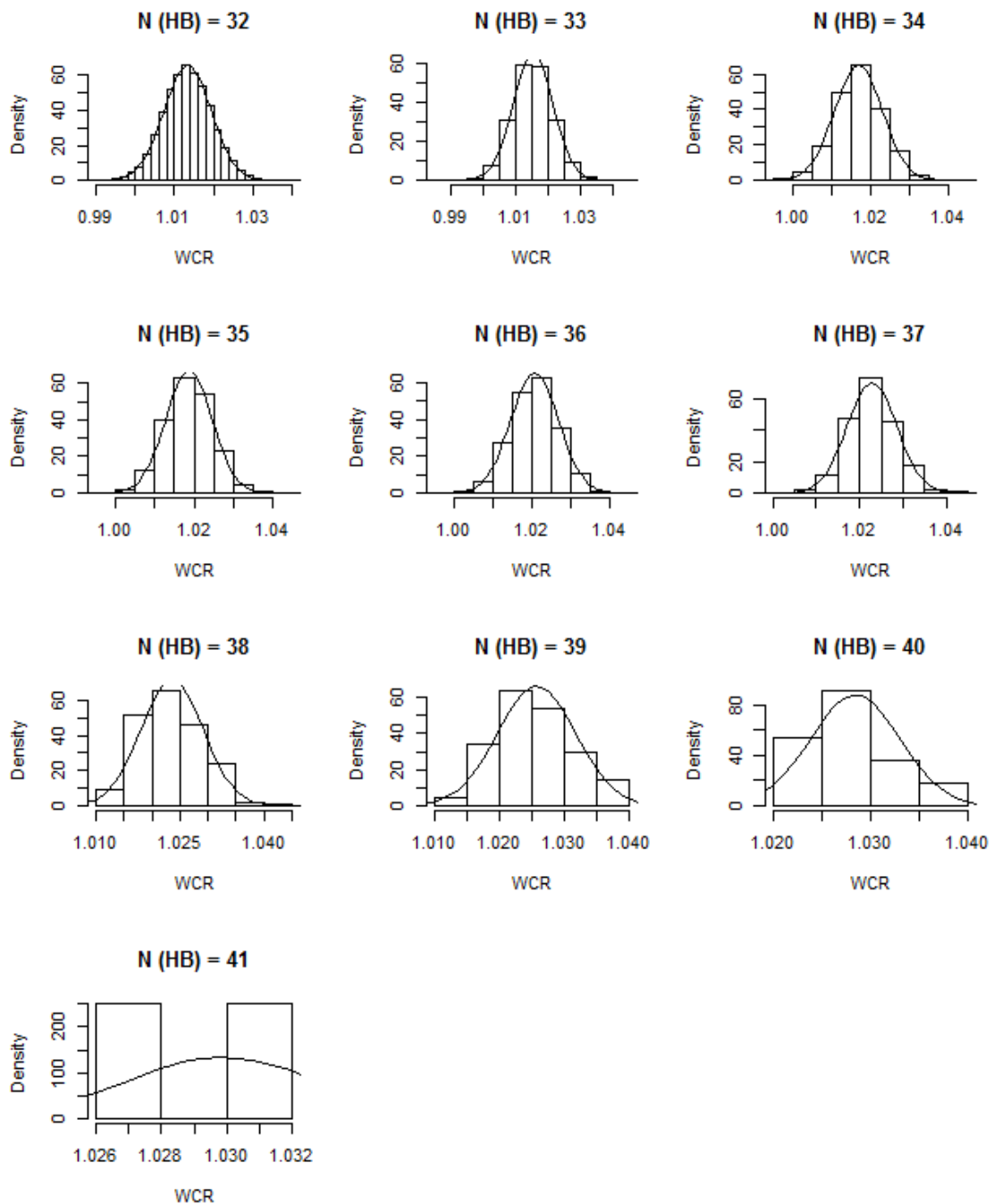


Figure A.8: Histograms of  $WCR$  for each  $N_{HB}$  in the case of  $HBs=50$  with best fitting normal pdf. 3,110,000 simulated cases from the steady state are used here.  $WCR$  does not extend to the end of  $N_{HB}$  because the joint distributions of  $WCR$  and  $N_{HB}$  in these cases do not extend to the end of  $N_{HB}$ , as shown in figure 5.3.

A.1.2 Investigation of Normality of  $WCR$ : normal QQ plots

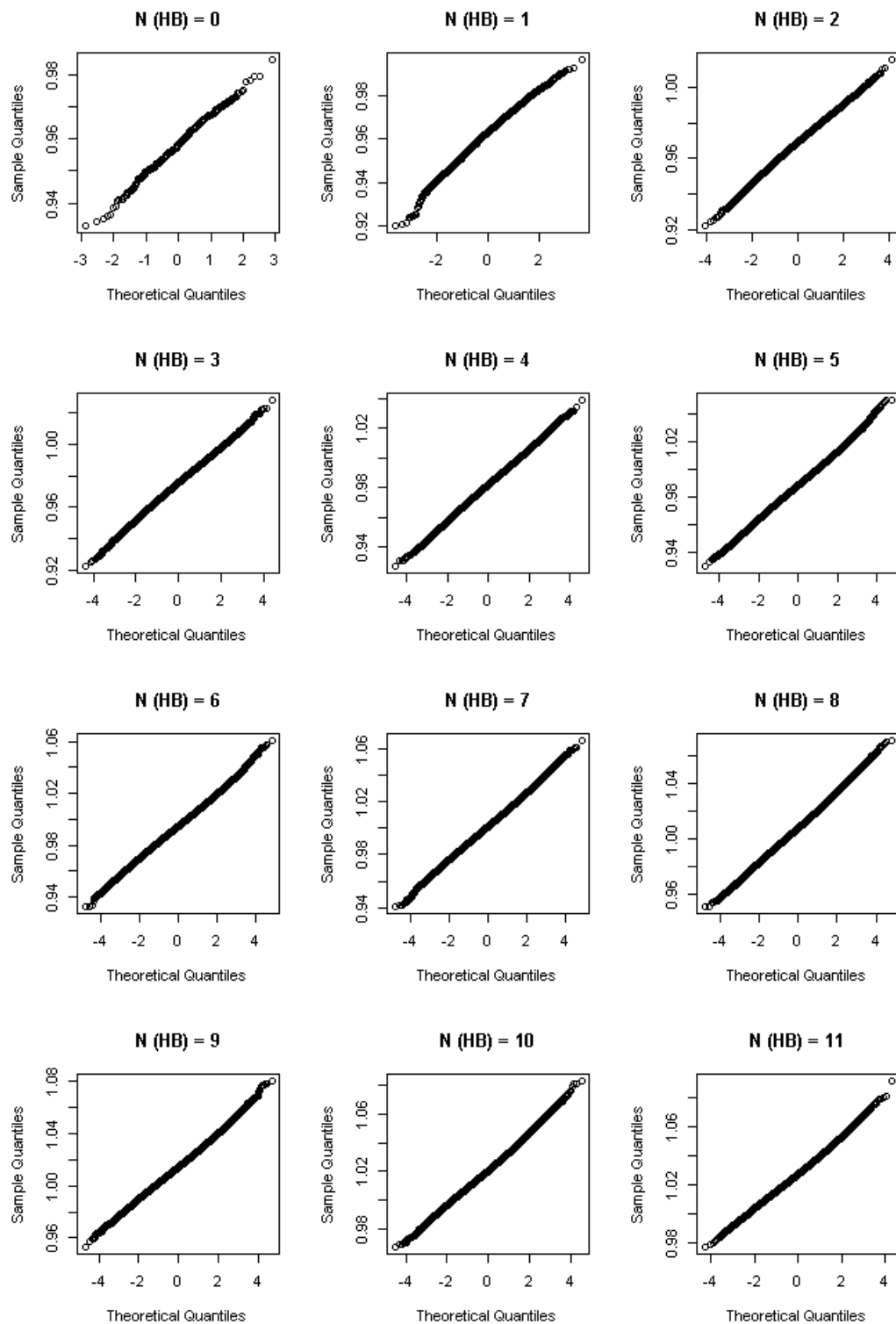


Figure A.9: Normal QQ plot of  $WCR$  for each  $N_{HB}$  in the case of HBs=14.

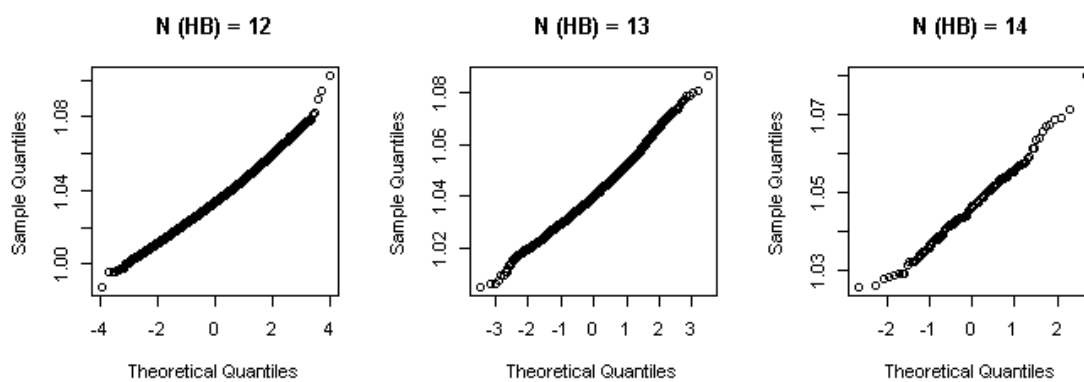


Figure A.10: Normal QQ plot of  $WCR$  for each  $N_{HB}$  in the case of HBs=14.

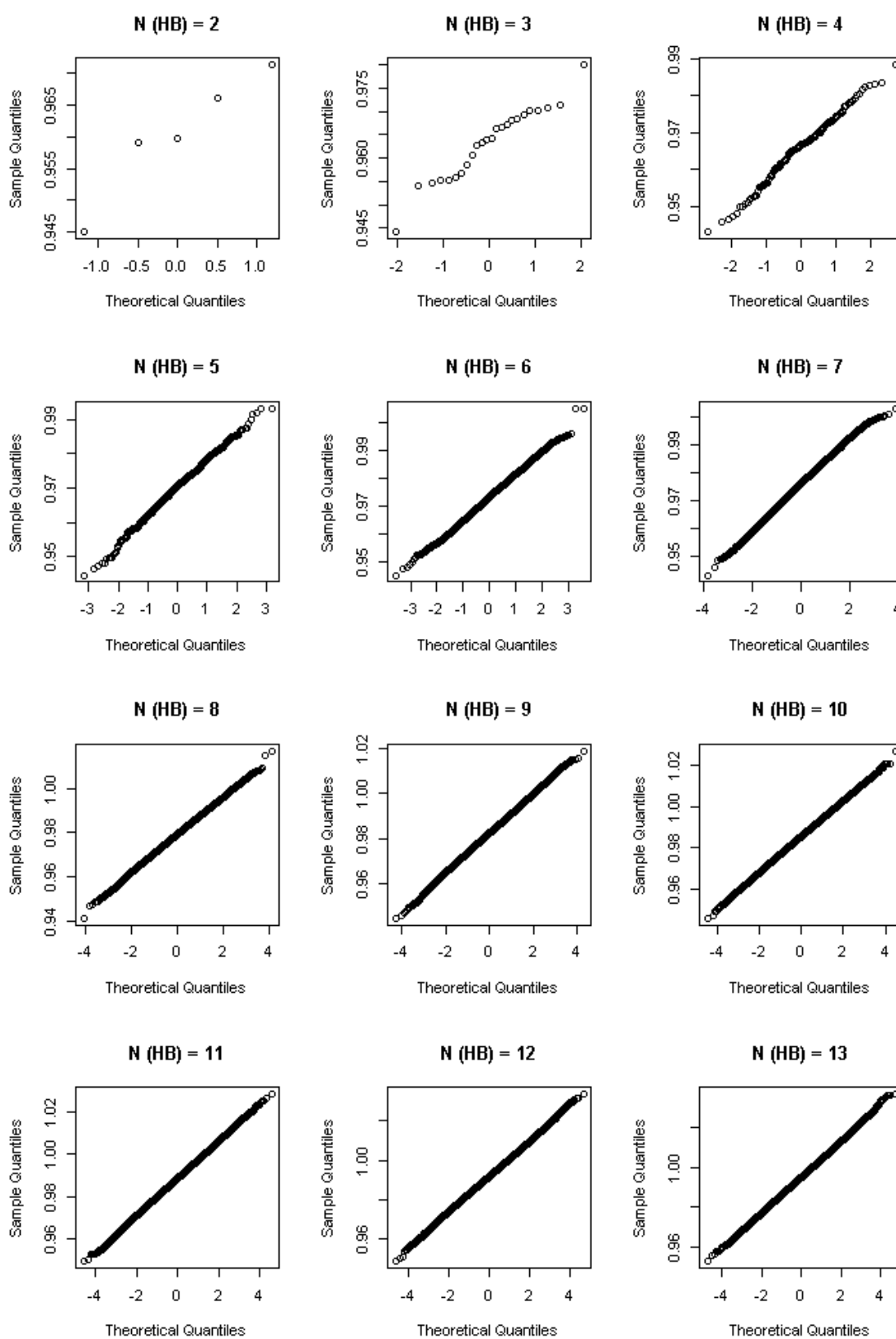


Figure A.11: Normal QQ plot of  $WCR$  for each  $N_{HB}$  in the case of  $HBs=30$ .  $WCR$  does not start from zero of  $N_{HB}$  because the joint distributions of  $WCR$  and  $N_{HB}$  in these cases do not start from zero, as shown in figure 5.2.

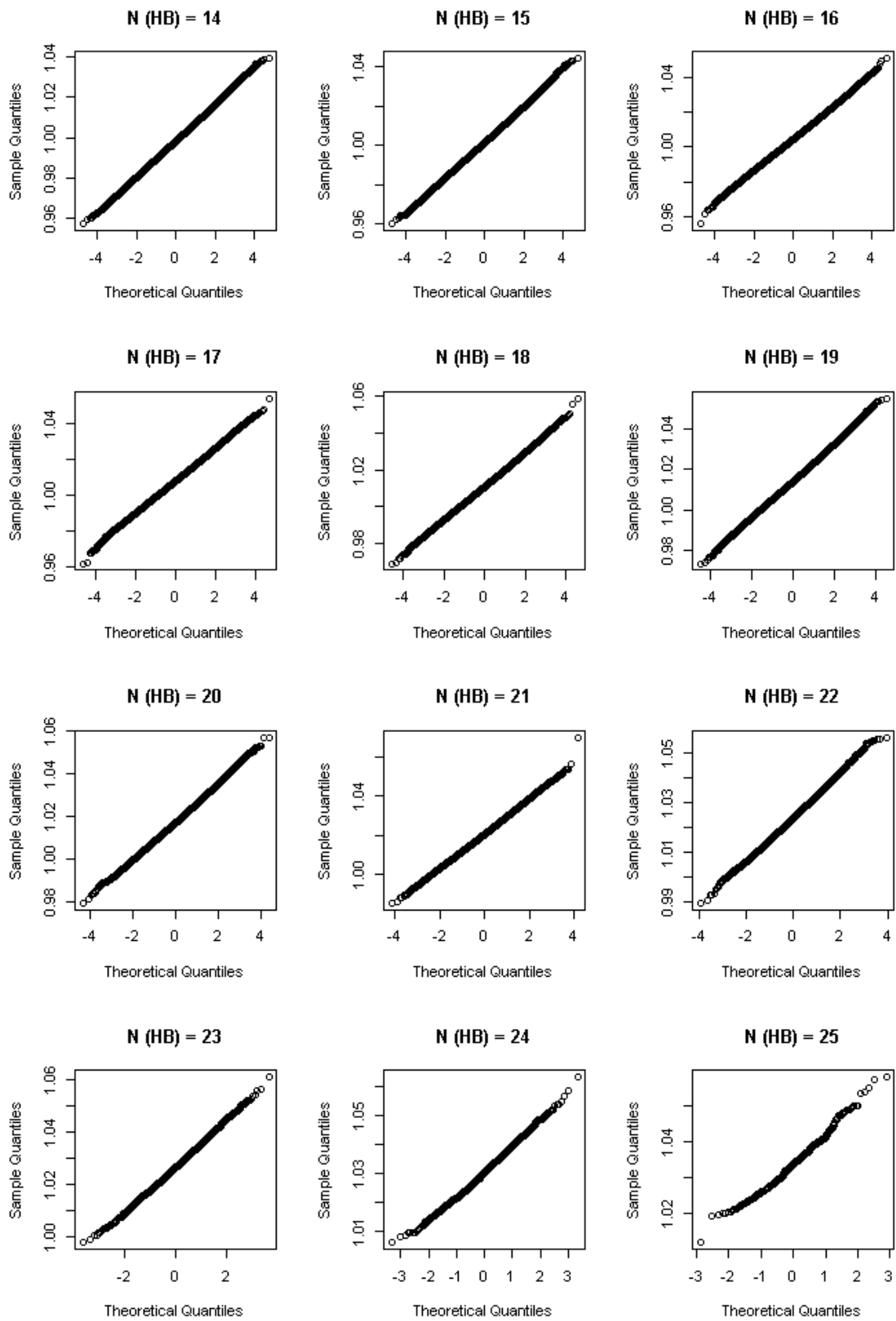


Figure A.12: Normal QQ plot of  $WCR$  for each  $N_{HB}$  in the case of  $HBs=30$ .

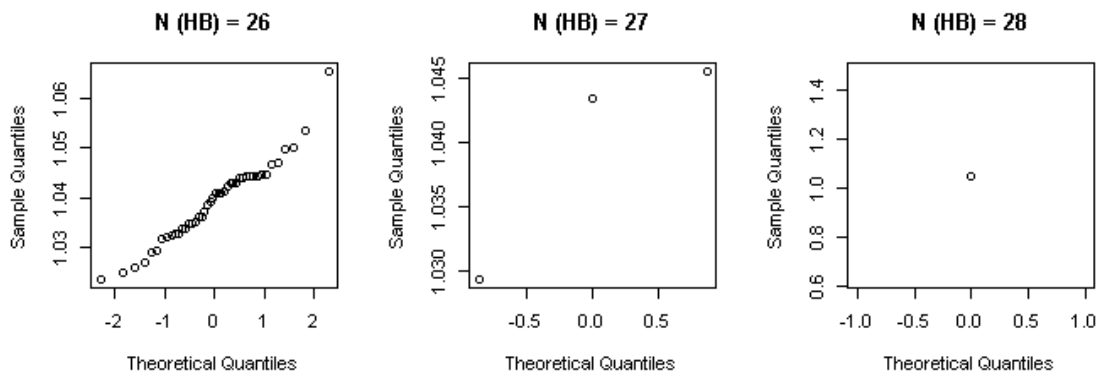


Figure A.13: Normal QQ plot of  $WCR$  for each  $N_{HB}$  in the case of  $HBs=30$ .  $WCR$  does not extend to the end of  $N_{HB}$  because the joint distributions of  $WCR$  and  $N_{HB}$  in these cases do not extend to the end of  $N_{HB}$ , as shown in figure 5.2.



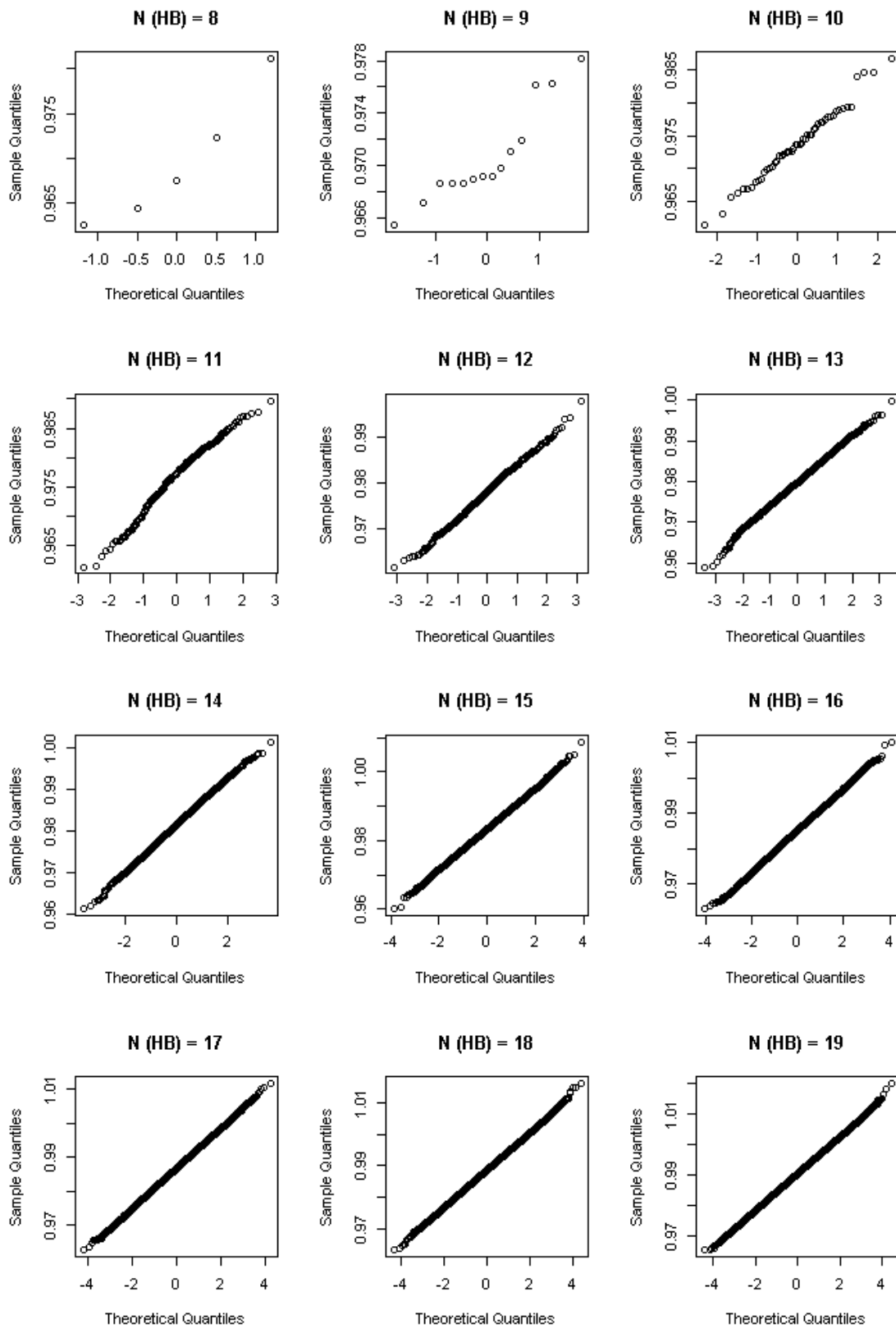
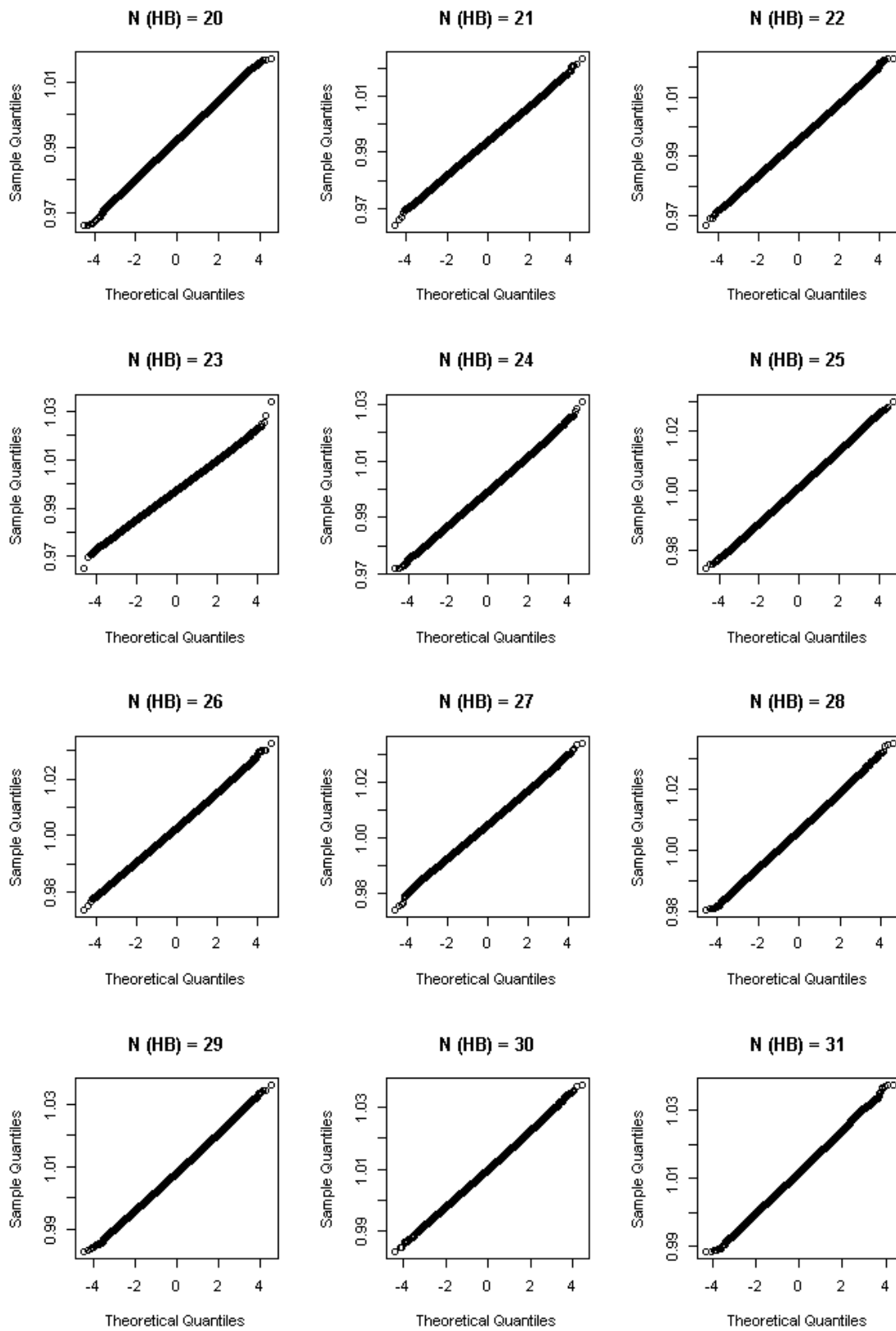


Figure A.14: Normal QQ plot of  $WCR$  for each  $N_{HB}$  in the case of  $HBs=50$ .  $WCR$  does not start from zero of  $N_{HB}$  because the joint distributions of  $WCR$  and  $N_{HB}$  in these cases do not start from zero, as shown in figure 5.3.

Figure A.15: Normal QQ plot of  $WCR$  for each  $N_{HB}$  in the case of  $HBs=50$ .

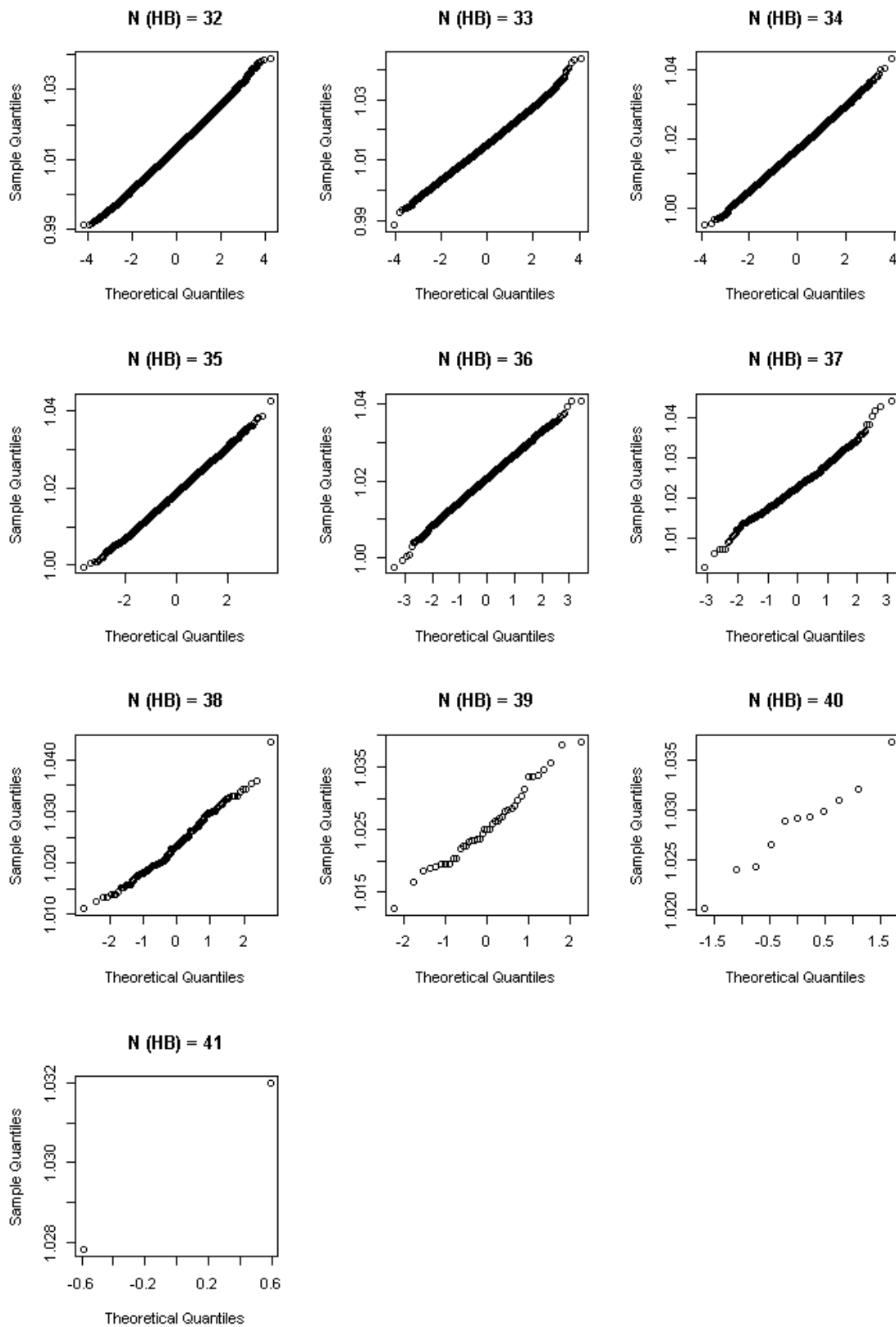


Figure A.16: Normal QQ plot of  $WCR$  for each  $N_{HB}$  in the case of  $HBs=50$ .  $WCR$  does not extend to the end of  $N_{HB}$  because the joint distributions of  $WCR$  and  $N_{HB}$  in these cases do not extend to the end of  $N_{HB}$ , as shown in figure 5.3.

**A.2** Finding functions to relate  $WCR$  and  $N_{HB}$  (1)

```

> m14 #(Modelling for mean with HBs=14)
Call:
glm(formula = m ~ ni, weights = w2)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-8.279e-04 -4.966e-04  7.009e-05  4.152e-04  9.153e-04

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 9.570e-01  7.612e-04 1257.30 < 2e-16 ***
ni          6.232e-03  7.901e-05   78.88 2.8e-10 ***
---

(Dispersion parameter for gaussian family taken to be 4.368385e-07)

Null deviance: 2.7204e-03  on 7  degrees of freedom
Residual deviance: 2.6210e-06  on 6  degrees of freedom
AIC: -96.369

Number of Fisher Scoring iterations: 2

#####
> sd14 #(Modelling for standard deviation with HBs=14)
Call:
glm(formula = sd ~ ni, weights = w2)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-8.540e-04 -1.705e-04  8.546e-05  2.328e-04  2.839e-04

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 1.448e-02  5.103e-04  28.368 1.27e-07 ***
ni         -2.517e-04  5.297e-05  -4.753 0.00315 **
---

(Dispersion parameter for gaussian family taken to be 1.963305e-07)

Null deviance: 5.6126e-06  on 7  degrees of freedom
Residual deviance: 1.1780e-06  on 6  degrees of freedom

```

AIC: -102.77

Number of Fisher Scoring iterations: 2

#####

> sd14BC #(Modelling for standard deviation

#(using Box-Cox) with HBs=14)

Call:

glm(formula = sd~14 ~ ni, weights = w2)

Deviance Residuals:

	Min	1Q	Median	3Q	Max
	-4.269e-28	-8.887e-29	-6.842e-29	1.020e-28	2.874e-28

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	4.599e-27	2.830e-28	16.25	3.46e-06 ***
ni	-3.145e-28	2.938e-29	-10.70	3.92e-05 ***
---				

(Dispersion parameter for gaussian family taken to be 6.040332e-56)

Null deviance: 7.2836e-54 on 7 degrees of freedom

Residual deviance: 3.6242e-55 on 6 degrees of freedom

AIC: -996.39

Number of Fisher Scoring iterations: 1

#####

> m30 #(Modelling for mean with HBs=30)

Call:

glm(formula = m ~ ni, weights = w2)

Deviance Residuals:

	Min	1Q	Median	3Q	Max
	-5.881e-04	-2.587e-04	-6.522e-05	4.098e-05	1.637e-03

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	9.545e-01	7.763e-04	1229.50	<2e-16 ***
ni	3.058e-03	4.042e-05	75.64	<2e-16 ***
---				

(Dispersion parameter for gaussian family taken to be 4.002341e-07)

Null deviance: 2.2949e-03 on 13 degrees of freedom  
 Residual deviance: 4.8028e-06 on 12 degrees of freedom  
 AIC: -166.53

Number of Fisher Scoring iterations: 2

#####

> sd30 #(Modelling for standard deviation with HBs=30)

Call:

glm(formula = sd ~ ni, weights = w2)

Deviance Residuals:

	Min	1Q	Median	3Q	Max
	-3.910e-04	-1.489e-04	2.893e-05	9.169e-05	6.479e-04

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	9.814e-03	3.185e-04	30.815	8.57e-13 ***
ni	-4.718e-05	1.658e-05	-2.845	0.0148 *

---

(Dispersion parameter for gaussian family taken to be 6.735385e-08)

Null deviance: 1.3535e-06 on 13 degrees of freedom  
 Residual deviance: 8.0825e-07 on 12 degrees of freedom  
 AIC: -191.48

Number of Fisher Scoring iterations: 2

#####

> sd30BC #(Modelling for standard deviation

#(using Box-Cox) with HBs=30)

Call:

glm(formula = sd^(-1) ~ ni, weights = w2)

Deviance Residuals:

	Min	1Q	Median	3Q	Max
	-7.7218	-1.1902	-0.3907	1.7699	5.7653

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	100.1242	4.1012	24.414	1.34e-11 ***
ni	0.6368	0.2135	2.982	0.0114 *

---

(Dispersion parameter for gaussian family taken to be 11.16999)

Null deviance: 233.36 on 13 degrees of freedom  
 Residual deviance: 134.04 on 12 degrees of freedom  
 AIC: 73.489

Number of Fisher Scoring iterations: 2

#####

> m50 #(Modelling for mean with HBs=50)

Call:

glm(formula = m ~ ni, weights = w2)

Deviance Residuals:

Min	1Q	Median	3Q	Max
-3.620e-04	-2.050e-04	-8.854e-05	9.188e-05	1.294e-03

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	0.9573079	0.0006249	1532.05	<2e-16 ***
ni	0.0017151	0.0000205	83.66	<2e-16 ***

---

(Dispersion parameter for gaussian family taken to be 2.089265e-07)

Null deviance: 1.4655e-03 on 17 degrees of freedom  
 Residual deviance: 3.3428e-06 on 16 degrees of freedom  
 AIC: -225.86

Number of Fisher Scoring iterations: 2

#####

> sd50 #(Modelling for standard deviation with HBs=50)

Call:

glm(formula = sd ~ ni, weights = w2)

Deviance Residuals:

Min	1Q	Median	3Q	Max
-----	----	--------	----	-----

```
-1.082e-03  -1.751e-05  2.079e-05  7.897e-05  7.037e-04
```

Coefficients:

```
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  6.934e-03  4.641e-04  14.940 8.11e-11 ***
ni           -2.854e-05  1.523e-05  -1.874  0.0793 .
---
```

(Dispersion parameter for gaussian family taken to be 1.15281e-07)

```
Null deviance: 2.2495e-06  on 17  degrees of freedom
Residual deviance: 1.8445e-06  on 16  degrees of freedom
AIC: -236.57
```

Number of Fisher Scoring iterations: 2

```
#####
```

```
> sd50BC #(Modelling for standard deviation
#(using Box-Cox) with HBs=50)
```

Call:

```
glm(formula = sd^3 ~ ni, weights = w2)
```

Deviance Residuals:

```
          Min           1Q           Median           3Q          Max
-7.539e-08  -5.254e-09   2.146e-09   6.165e-09   8.822e-08
```

Coefficients:

```
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.978e-07  4.204e-08   7.083 2.59e-06 ***
ni           -2.406e-09  1.379e-09  -1.745  0.1
---
```

(Dispersion parameter for gaussian family taken to be 9.456423e-16)

```
Null deviance: 1.8009e-14  on 17  degrees of freedom
Residual deviance: 1.5130e-14  on 16  degrees of freedom
AIC: -571.7
```

Number of Fisher Scoring iterations: 1

### A.3 Finding functions to relate $WCR$ and $N_{HB}$ (2)

```
>z1 #(Modelling for mean with different HBs=30, 40 and 50
```



```

#comparing to HB= 14)
Call:
lm(formula = m ~ n.i.midway * factor(ng), data = Data)

Residuals:
      Min       1Q   Median       3Q      Max
-0.0016042 -0.0001003  0.0001032  0.0002810  0.0016564

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    9.998e-01  3.773e-04 2649.778  <2e-16 ***
n.i.midway     6.472e-03  9.019e-05  71.763  <2e-16 ***
factor(ng)30   -3.932e-05  4.798e-04  -0.082    0.935
factor(ng)40    3.900e-04  4.693e-04   0.831    0.410
factor(ng)50   -1.118e-04  4.607e-04  -0.243    0.809
n.i.midway:factor(ng)30 -3.268e-03  9.817e-05 -33.292  <2e-16 ***
n.i.midway:factor(ng)40 -4.254e-03  9.560e-05 -44.501  <2e-16 ***
n.i.midway:factor(ng)50 -4.653e-03  9.402e-05 -49.484  <2e-16 ***
---

Residual standard error: 0.0005845 on 48 degrees of freedom
Multiple R-squared:  0.9979,    Adjusted R-squared:  0.9976
F-statistic: 3259 on 7 and 48 DF,  p-value: < 2.2e-16

> z2 #(Modelling for mean between different HBs=14, 30, 40 and 50)

Call:
lm(formula = m ~ n.i.midway:factor(ng), data = Data)

Residuals:
      Min       1Q   Median       3Q      Max
-1.558e-03 -1.111e-04  6.911e-05  2.959e-04  1.680e-03

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    9.998e-01  1.466e-04 6817.9  <2e-16 ***
n.i.midway:factor(ng)14 6.459e-03  5.701e-05  113.3  <2e-16 ***
n.i.midway:factor(ng)30 3.193e-03  2.596e-05  123.0  <2e-16 ***
n.i.midway:factor(ng)40 2.250e-03  2.171e-05  103.6  <2e-16 ***
n.i.midway:factor(ng)50 1.805e-03  1.859e-05   97.1  <2e-16 ***
---

```

Residual standard error: 0.0005785 on 51 degrees of freedom  
 Multiple R-squared: 0.9978, Adjusted R-squared: 0.9976  
 F-statistic: 5821 on 4 and 51 DF, p-value: < 2.2e-16

```
>z3 #(Modelling for mean between different HBs=14, 30, 40 and 50
# using weighted regression)
```

Call:

```
lm(formula = m ~ n.i.midway:factor(ng), data = Data,
weights = weight)
```

Weighted Residuals:

Min	1Q	Median	3Q	Max
-1.366e-03	-2.052e-04	-8.959e-05	1.400e-06	9.056e-04

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )	
(Intercept)	9.994e-01	7.794e-05	12823.51	<2e-16	***
n.i.midway:factor(ng)14	6.745e-03	8.002e-05	84.29	<2e-16	***
n.i.midway:factor(ng)30	3.342e-03	4.717e-05	70.85	<2e-16	***
n.i.midway:factor(ng)40	2.433e-03	3.873e-05	62.81	<2e-16	***
n.i.midway:factor(ng)50	1.905e-03	3.328e-05	57.25	<2e-16	***

---

Residual standard error: 0.0004223 on 51 degrees of freedom  
 Multiple R-squared: 0.9957, Adjusted R-squared: 0.9954  
 F-statistic: 2986 on 4 and 51 DF, p-value: < 2.2e-16

```
>z4 #(Modelling for log(slope) for mean and log(ng))
```

Call:

```
lm(formula = log(slope) ~ log(ng), data = Results)
```

Residuals:

n.i.midway:factor(ng)14	n.i.midway:factor(ng)30
-0.014480	0.035331
n.i.midway:factor(ng)40	n.i.midway:factor(ng)50
0.001722	-0.022573

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )	
(Intercept)	-2.38015	0.11245	-21.17	0.00222	**

```
log(ng)      -0.98682      0.03265  -30.23  0.00109 **
---
```

```
Residual standard error: 0.03139 on 2 degrees of freedom
Multiple R-squared:  0.9978,    Adjusted R-squared:  0.9967
F-statistic: 913.5 on 1 and 2 DF,  p-value: 0.001093
```

```
>z5 #(Modelling for variance with different HBs=30, 40 and 50
#comparing to HB= 14
# using weighted regression)
Call:
glm(formula = var ~ n.i.midway * factor(ng), family = Gamma,
     data = Data, weights = weight)
```

```
Deviance Residuals:
      Min       1Q   Median       3Q      Max
-0.018070 -0.006605 -0.003130  0.001493  0.017452
```

```
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)      6386.82     21.34 299.279 < 2e-16 ***
n.i.midway        145.10     13.02  11.142 6.54e-15 ***
factor(ng)30     5931.42     42.23 140.458 < 2e-16 ***
factor(ng)40    12544.25     57.36 218.682 < 2e-16 ***
factor(ng)50    20493.32     75.59 271.112 < 2e-16 ***
n.i.midway:factor(ng)30  -62.39     19.46  -3.206  0.0024 **
n.i.midway:factor(ng)40  -57.32     22.29  -2.572  0.0133 *
n.i.midway:factor(ng)50  -66.64     25.41  -2.622  0.0117 *
---
```

```
(Dispersion parameter for Gamma family taken to be 6.194123e-05)
```

```
Null deviance: 17.5004015 on 55 degrees of freedom
Residual deviance: 0.0030914 on 48 degrees of freedom
AIC: -1706.1
```

```
Number of Fisher Scoring iterations: 3
```

```
>z6 #(Modelling for variance between different HBs=14, 30, 40 and 50
# using weighted regression)
Call:
```

```
glm(formula = var ~ n.i.midway + factor(ng) - 1, family = Gamma,
data = Data, weights = weight)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-0.025708	-0.005840	-0.001104	0.002024	0.017690

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
n.i.midway	106.483	8.732	12.2	<2e-16 ***
factor(ng)14	6431.593	19.561	328.8	<2e-16 ***
factor(ng)30	12274.279	31.531	389.3	<2e-16 ***
factor(ng)40	18890.132	43.465	434.6	<2e-16 ***
factor(ng)50	26810.344	56.933	470.9	<2e-16 ***

---

(Dispersion parameter for Gamma family taken to be 7.486974e-05)

Null deviance: NaN on 56 degrees of freedom  
Residual deviance: 0.0039733 on 51 degrees of freedom  
AIC: -1695.7

Number of Fisher Scoring iterations: 3

```
>z7 #(Modelling for log(variance intercept) for variance and ng)
```

Call:

```
lm(formula = log(Var.Int) ~ ng, data = Results)
```

Residuals:

n.i.midway:factor(ng)14	n.i.midway:factor(ng)30
-0.007374	-0.001454
n.i.midway:factor(ng)40	n.i.midway:factor(ng)50
0.029453	-0.020626

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	8.2160336	0.0352087	233.35	1.84e-05 ***
ng	0.0400227	0.0009769	40.97	0.000595 ***

---

Residual standard error: 0.02597 on 2 degrees of freedom

Multiple R-squared: 0.9988, Adjusted R-squared: 0.9982

F-statistic: 1679 on 1 and 2 DF, p-value: 0.0005952

# Appendix B

## Chapter 6: Detection system

### B.1 Making the system more general

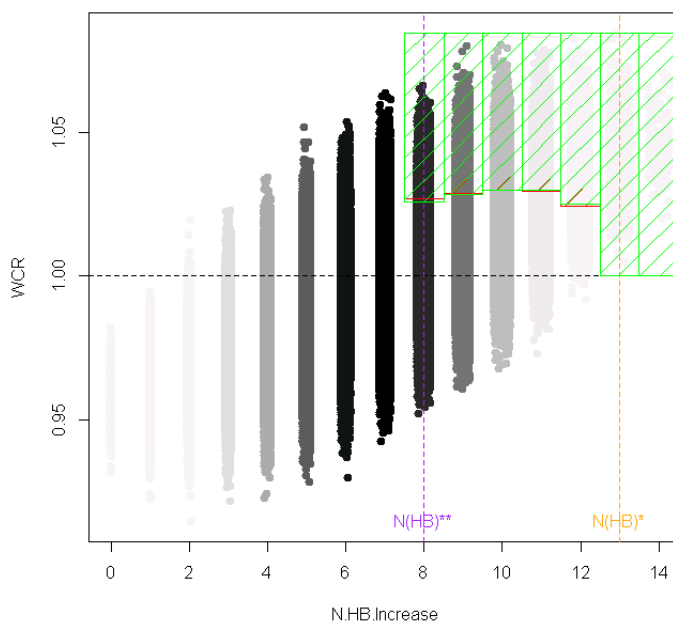


Figure B.1:  $WCR$  against  $N_{HB}$  for steady rates using  $r = 1$  with 14 HBs and a 5 million population.  $N_{HB}^*$  is presented in the dotted vertical orange line while the dotted vertical purple line represents  $N_{HB}^{**}$ . The red shaped area columns came from calculating the quantile (cut-offs) using  $\mu$  and  $\sigma$  from many simulated cases with a flat rate. They represent around 5% of the most extreme observations under the steady state. The green shaped area columns came from calculating  $qnorm$  (cut-offs) using  $\mu$  and  $\sigma$  from equations 5.4.7 and 5.4.10. They represent around 5% of the most extreme observations under the steady state. Darkness of the black colour reflects the number of simulated cases.

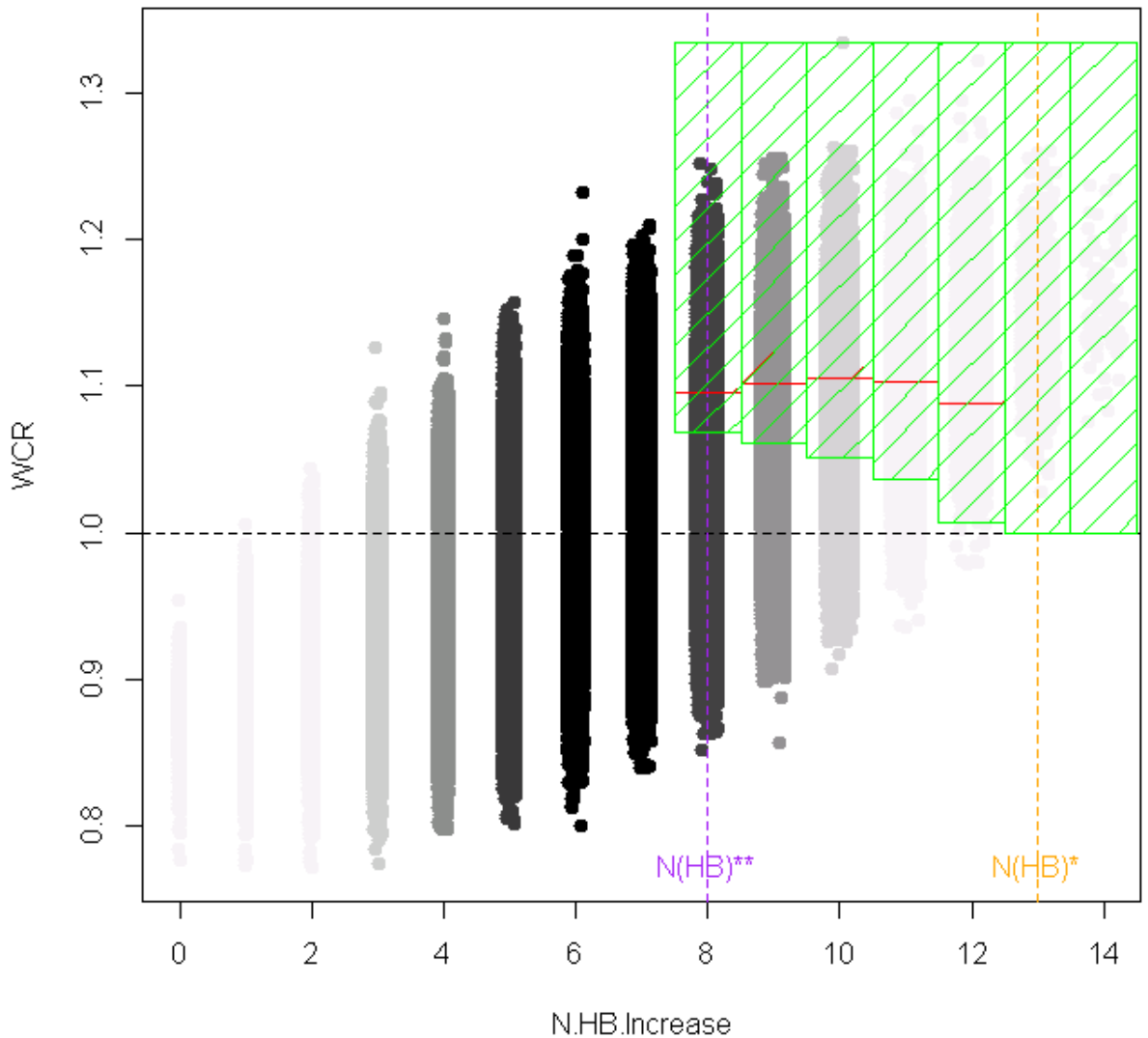


Figure B.2:  $WCR$  against  $N_{HB}$  for steady rates using  $r = 10$  with 14 HBs and a 5 million population.  $N_{HB}^*$  is presented in the dotted vertical orange line while the dotted vertical purple line represents  $N_{HB}^{**}$ . The red shaped area columns came from calculating the quantile (cut-offs) using  $\mu$  and  $\sigma$  from many simulated cases with a flat rate. They represent around 5% of the most extreme observations under the steady state. The green shaped area columns came from calculating  $qnorm$  (cut-offs) using  $\mu$  and  $\sigma$  from equations 5.4.7 and 5.4.10. They represent around 10% of the most extreme observations under the steady state. Darkness of the black colour reflects the number of simulated cases.

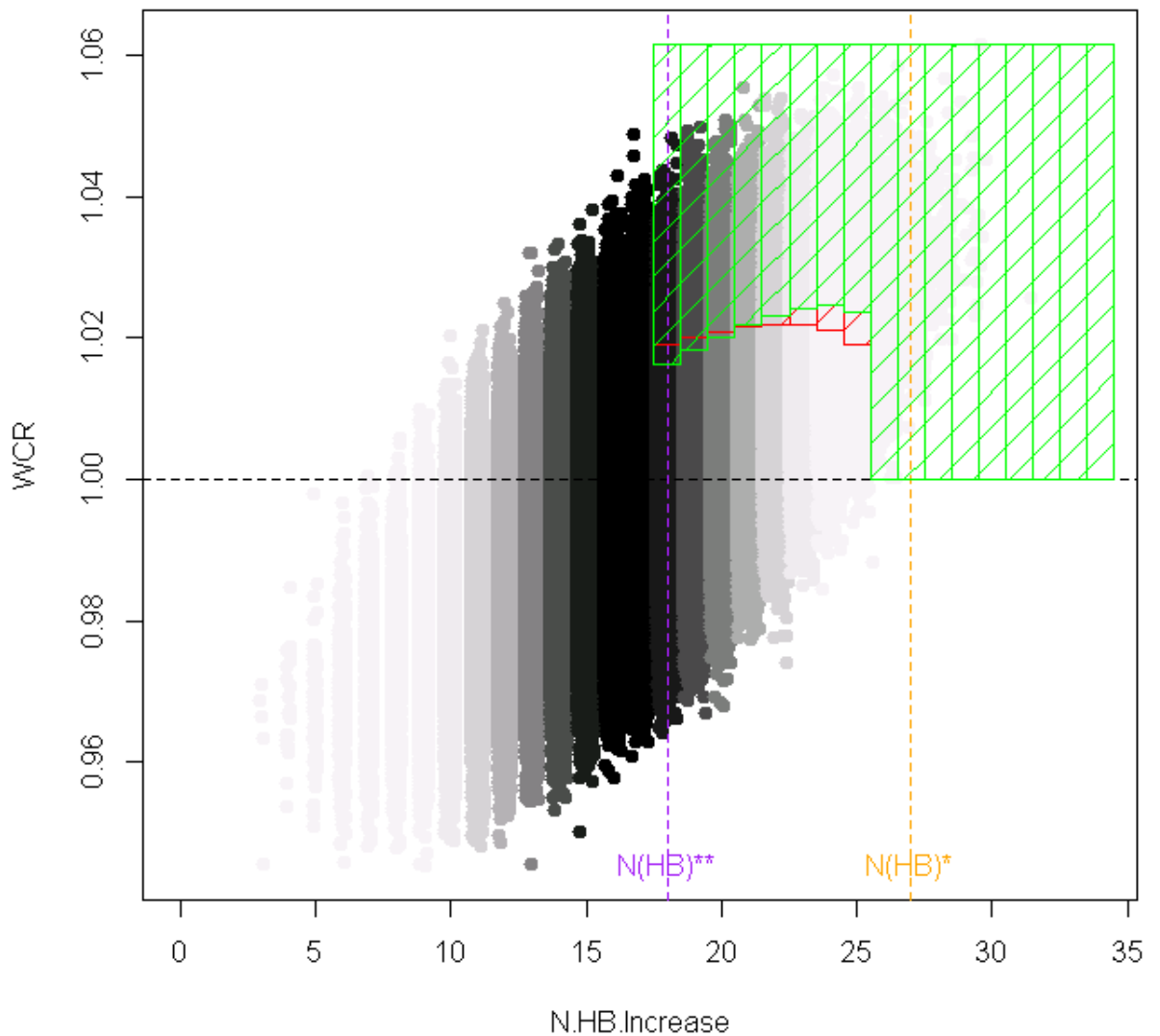


Figure B.3:  $WCR$  against  $N_{HB}$  for steady rates using  $r = 1$  with 34 HBs and a 5 million population.  $N_{HB}^*$  is presented in the dotted vertical orange line while the dotted vertical purple line represents  $N_{HB}^{**}$ . The red shaped area columns came from calculating the quantile (cut-offs) using  $\mu$  and  $\sigma$  from many simulated cases with a flat rate. They represent around 5% of the most extreme observations under the steady state. The green shaped area columns came from calculating  $qnorm$  (cut-offs) using  $\mu$  and  $\sigma$  from equations 5.4.7 and 5.4.10. They represent around 4% of the most extreme observations under the steady state. Darkness of the black colour reflects the number of simulated cases.



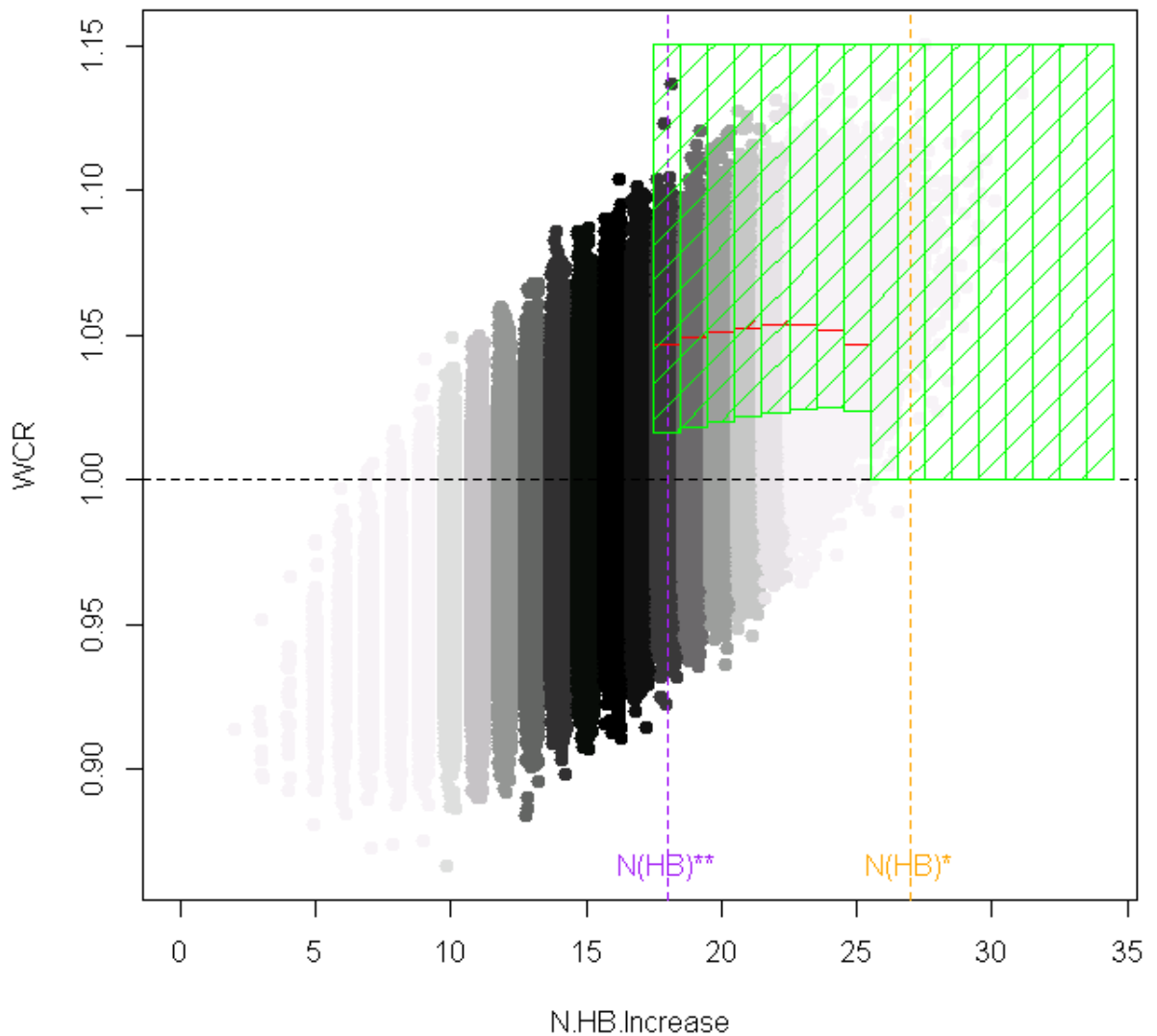


Figure B.4:  $WCR$  against  $N_{HB}$  for steady rates using  $r = 5$  with 34 HBs and a 5 million population.  $N_{HB}^*$  is presented in the dotted vertical orange line while the dotted vertical purple line represents  $N_{HB}^{**}$ . The red shaped area columns came from calculating the quantile (cut-offs) using  $\mu$  and  $\sigma$  from many simulated cases with a flat rate. They represent around 5% of the most extreme observations under the steady state. The green shaped area columns came from calculating  $qnorm$  (cut-offs) using  $\mu$  and  $\sigma$  from equations 5.4.7 and 5.4.10. They represent around 16% of the most extreme observations under the steady state. Darkness of the black colour reflects the number of simulated cases.

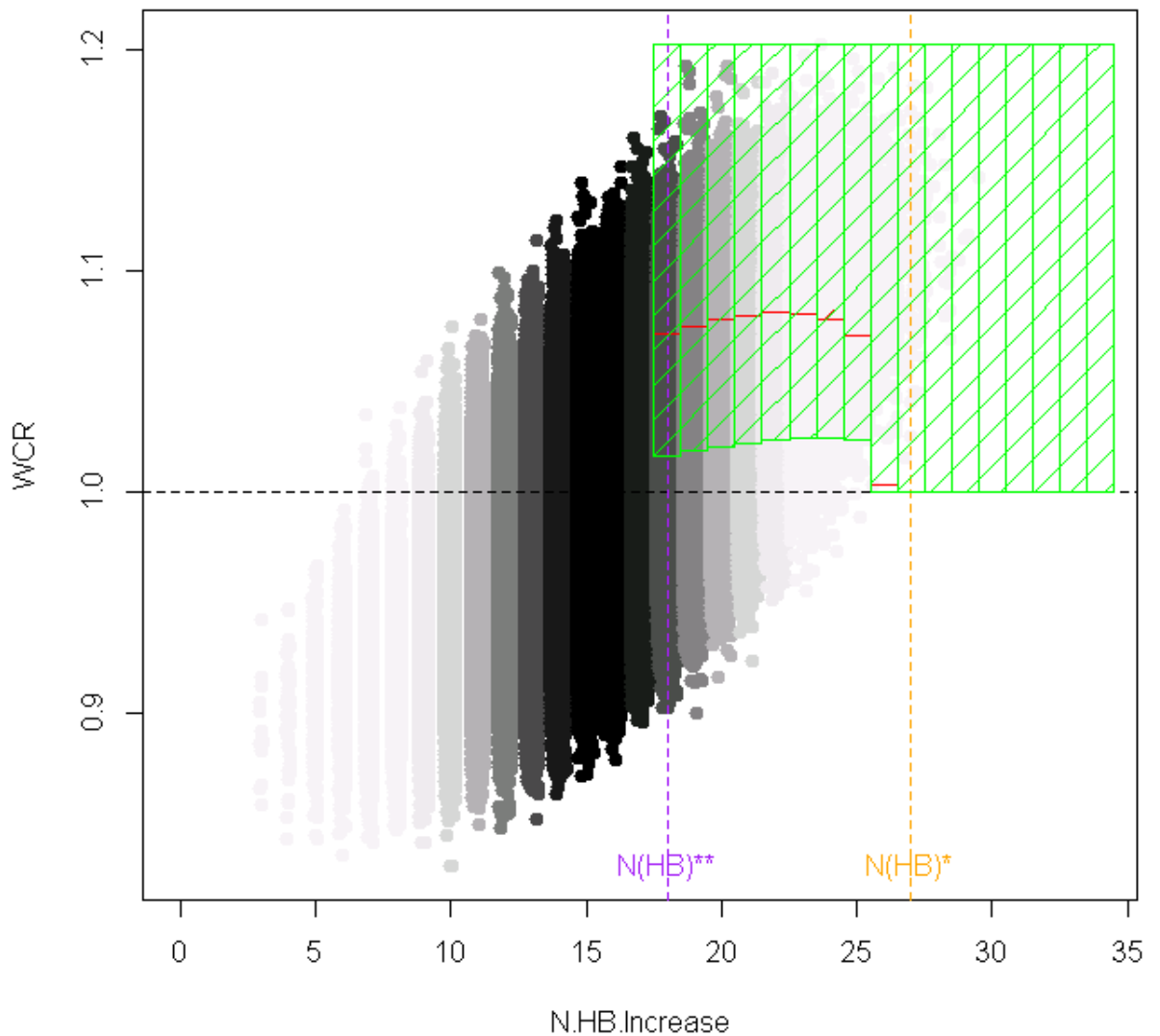


Figure B.5:  $WCR$  against  $N_{HB}$  for steady rates using  $r = 10$  with 34 HBs and a 5 million population.  $N_{HB}^*$  is presented in the dotted vertical orange line while the dotted vertical purple line represents  $N_{HB}^{**}$ . The red shaped area columns came from calculating the quantile (cut-offs) using  $\mu$  and  $\sigma$  from many simulated cases with a flat rate. They represent around 5% of the most extreme observations under the steady state. The green shaped area columns came from calculating  $qnorm$  (cut-offs) using  $\mu$  and  $\sigma$  from equations 5.4.7 and 5.4.10. They represent around 17% of the most extreme observations under the steady state. Darkness of the black colour reflects the number of simulated cases.

# Appendix C

## Chapter 7: Simulation study

### C.1 Illustration

$W$	$W L$	$W B$	$W$	$W L$	$W B$	$W$	$W L$	$W B$	$W$	$W L$	$W B$
1	31	04 Aug	14	44	03 Nov	27	5	02 Feb	40	18	04 May
2	32	11 Aug	15	45	10 Nov	28	6	09 Feb	41	19	11 May
3	33	18 Aug	16	46	17 Nov	29	7	16 Feb	42	20	18 May
4	34	25 Aug	17	47	24 Nov	30	8	23 Feb	43	21	25 May
5	35	01 Sep	18	48	01 Dec	31	9	02 March	44	22	01 June
6	36	08 Sep	19	49	08 Dec	32	10	09 March	45	23	08 June
7	37	15 Sep	20	50	15 Dec	33	11	16 March	46	24	15 June
8	38	22 Sep	21	51	22 Dec	34	12	23 March	47	25	22 June
9	39	29 Sep	22	52	29 Dec	35	13	30 March	48	26	29 June
10	40	06 Oct	23	1	05 Jan	36	14	06 April	49	27	06 July
11	41	13 Oct	24	2	12 Jan	37	15	13 April	50	28	13 July
12	42	20 Oct	25	3	19 Jan	38	16	20 April	51	29	20 July
13	43	27 Oct	26	4	26 Jan	39	17	27 April	52	30	27 July

Table C.1: Week locations and numbers during the year, where  $W$  = week number (1-52), and  $W L$ : week location in the year (31-52) and (1-30), and  $W B$ : week begin date.

# Appendix D

## Chapter 9: Testing *WCR* with other systems (2)

### D.1 Testing *WCR* with other systems (2)

```
>z7 #(Modelling for mean between different HBs=14, 30, 40 and 50
# using weighted regression)
Call:
lm(formula = m ~ n.i.midway + rate + ng, data = Data.sct,
    weights = weight)
```

Weighted Residuals:

Min	1Q	Median	3Q	Max
-0.054526	-0.010021	-0.000423	0.011122	0.083715

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	1.0429080	0.0045276	230.342	< 2e-16 ***
n.i.midway	0.0157269	0.0011714	13.426	< 2e-16 ***
rate	-0.0089893	0.0004590	-19.585	< 2e-16 ***
ng	0.0008943	0.0001192	7.503	8.34e-13 ***

---

Residual standard error: 0.02373 on 280 degrees of freedom  
Multiple R-squared: 0.7007, Adjusted R-squared: 0.6975  
F-statistic: 218.5 on 3 and 280 DF, p-value: < 2.2e-16

```
>z8 #(Modelling for variance between different HBs=14, 30, 40 and 50
# using weighted regression)
Call:
glm(formula = var1 ~ n.i.midway + rate + ng, family = Gamma,
    data = Data.sct, weights = weight)
```

Deviance Residuals:

	Min	1Q	Median	3Q	Max
	-0.262284	-0.013496	-0.000588	0.012021	0.274348

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )	
(Intercept)	-11.2329	4.0867	-2.749	0.006373	**
n.i.midway	-9.3168	0.9155	-10.176	< 2e-16	***
rate	265.4865	1.6027	165.650	< 2e-16	***
ng	-0.3855	0.1022	-3.772	0.000198	***
---					

(Dispersion parameter for Gamma family taken to be 0.003306198)

Null deviance: 194.24104 on 283 degrees of freedom  
Residual deviance: 0.92075 on 280 degrees of freedom  
AIC: -3747.3

Number of Fisher Scoring iterations: 4

## D.2 Testing of mean and variance equation

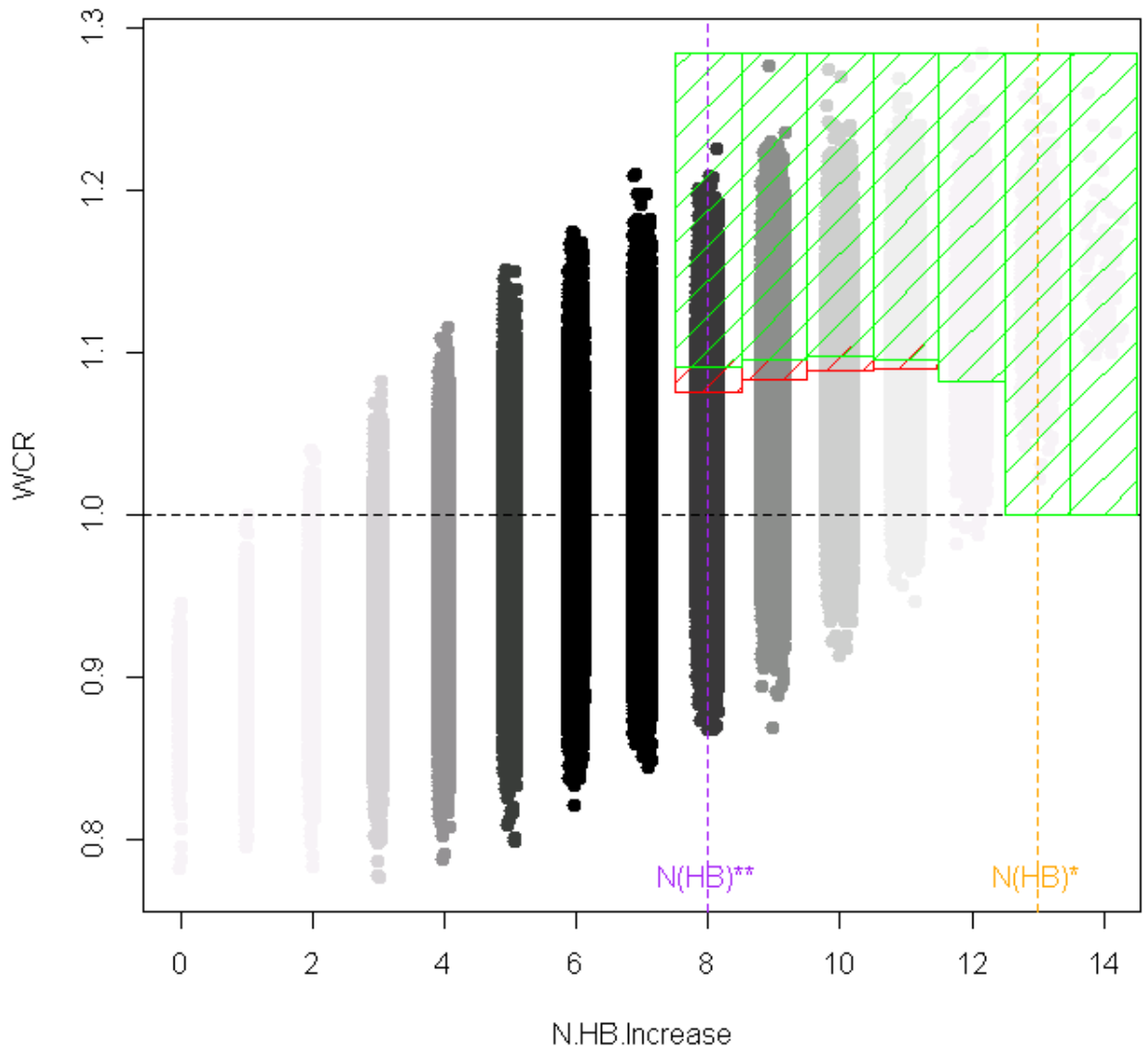


Figure D.1:  $WCR$  against  $N_{HB}$  for steady rates using fixed rate = 3 per 100,000 population with 14 HBs.  $N_{HB}^*$  is presented in the dotted vertical orange line while the dotted vertical purple line represents  $N_{HB}^{**}$ . The red shaped columns came from calculating the quantiles (cut-offs) using  $\mu$  and  $\sigma$  from many simulated cases (3,110,000) in the flat rate situation. They represent around 5% most extreme observations under the steady state. The green columns came from using `qnorm` to calculate cut-offs using  $\mu$  and  $\sigma$  from equations 9.3.1 and 9.3.2. They represent around 3% most extreme observations under the steady state. Darkness of the black colour reflects the number of simulated cases.

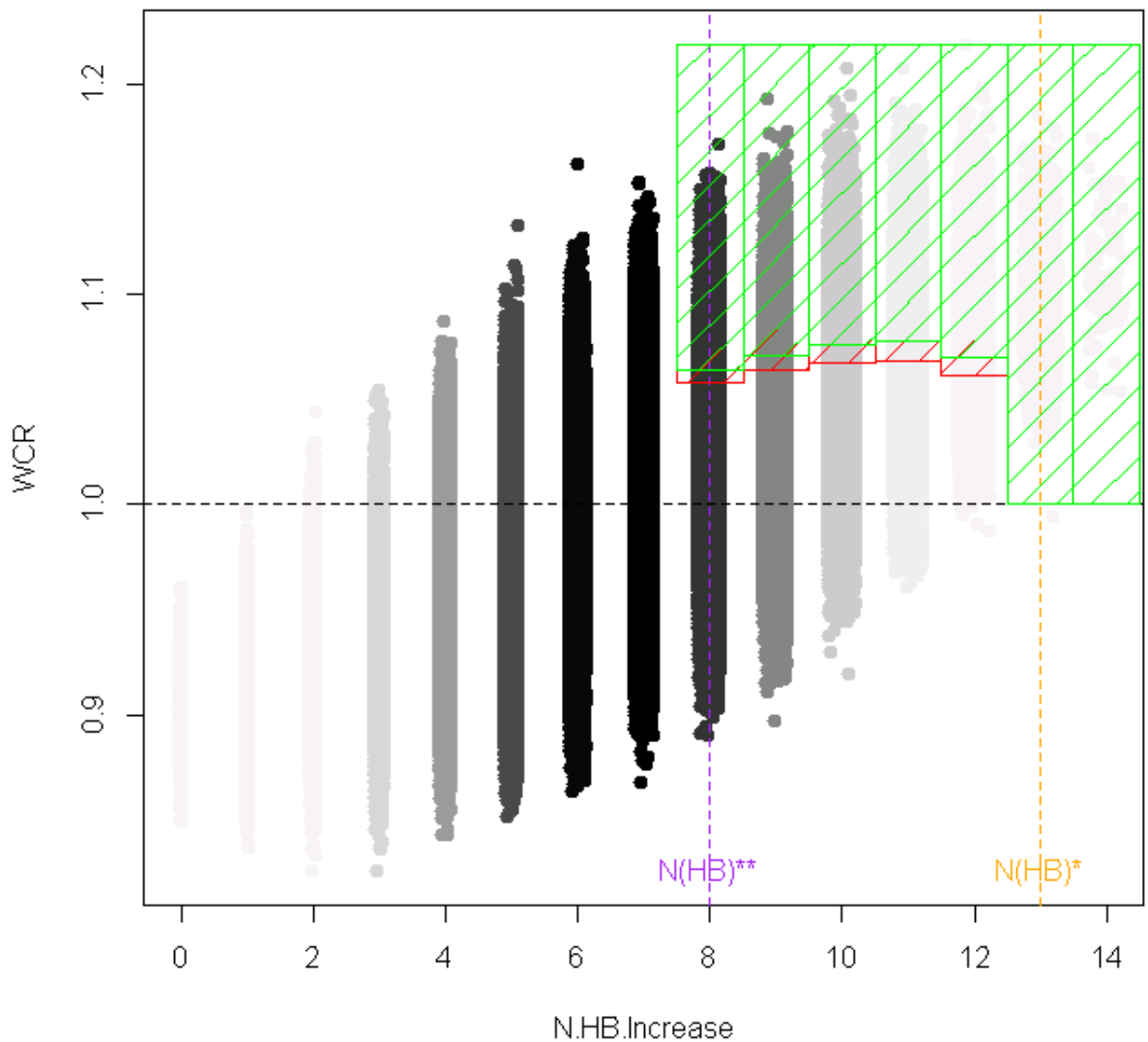


Figure D.2:  $WCR$  against  $N_{HB}$  for steady rates using fixed rate = 5 per 100,000 population with 14 HBs.  $N_{HB}^*$  is presented in the dotted vertical orange line while the dotted vertical purple line represents  $N_{HB}^{**}$ . The red shaped columns came from calculating the quantiles (cut-offs) using  $\mu$  and  $\sigma$  from many simulated cases (3,110,000) in the flat rate situation. They represent around 5% most extreme observations under the steady state. The green columns came from using `qnorm` to calculate cut-offs using  $\mu$  and  $\sigma$  from equations 9.3.1 and 9.3.2. They represent around 4% most extreme observations under the steady state. Darkness of the black colour reflects the number of simulated cases.

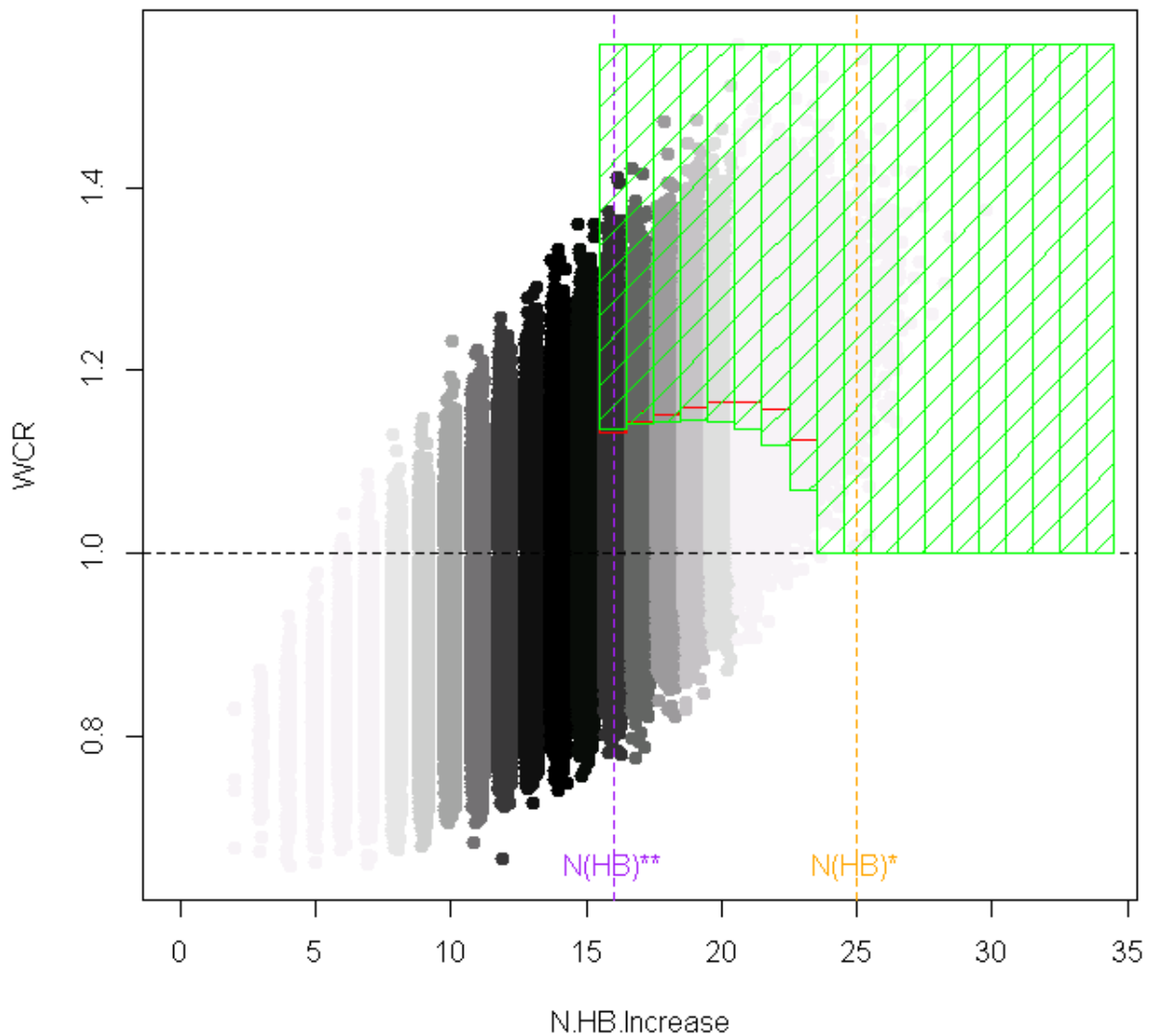


Figure D.3:  $WCR$  against  $N_{HB}$  for steady rates using fixed rate = 1 per 100,000 population with 34 HBs.  $N_{HB}^*$  is presented in the dotted vertical orange line while the dotted vertical purple line represents  $N_{HB}^{**}$ . The red shaped columns came from calculating the quantiles (cut-offs) using  $\mu$  and  $\sigma$  from many simulated cases (3,110,000) in the flat rate situation. They represent around 5% most extreme observations under the steady state. The green columns came from using `qnorm` to calculate cut-offs using  $\mu$  and  $\sigma$  from equations 9.3.1 and 9.3.2. They represent around 6% most extreme observations under the steady state. Darkness of the black colour reflects the number of simulated cases.



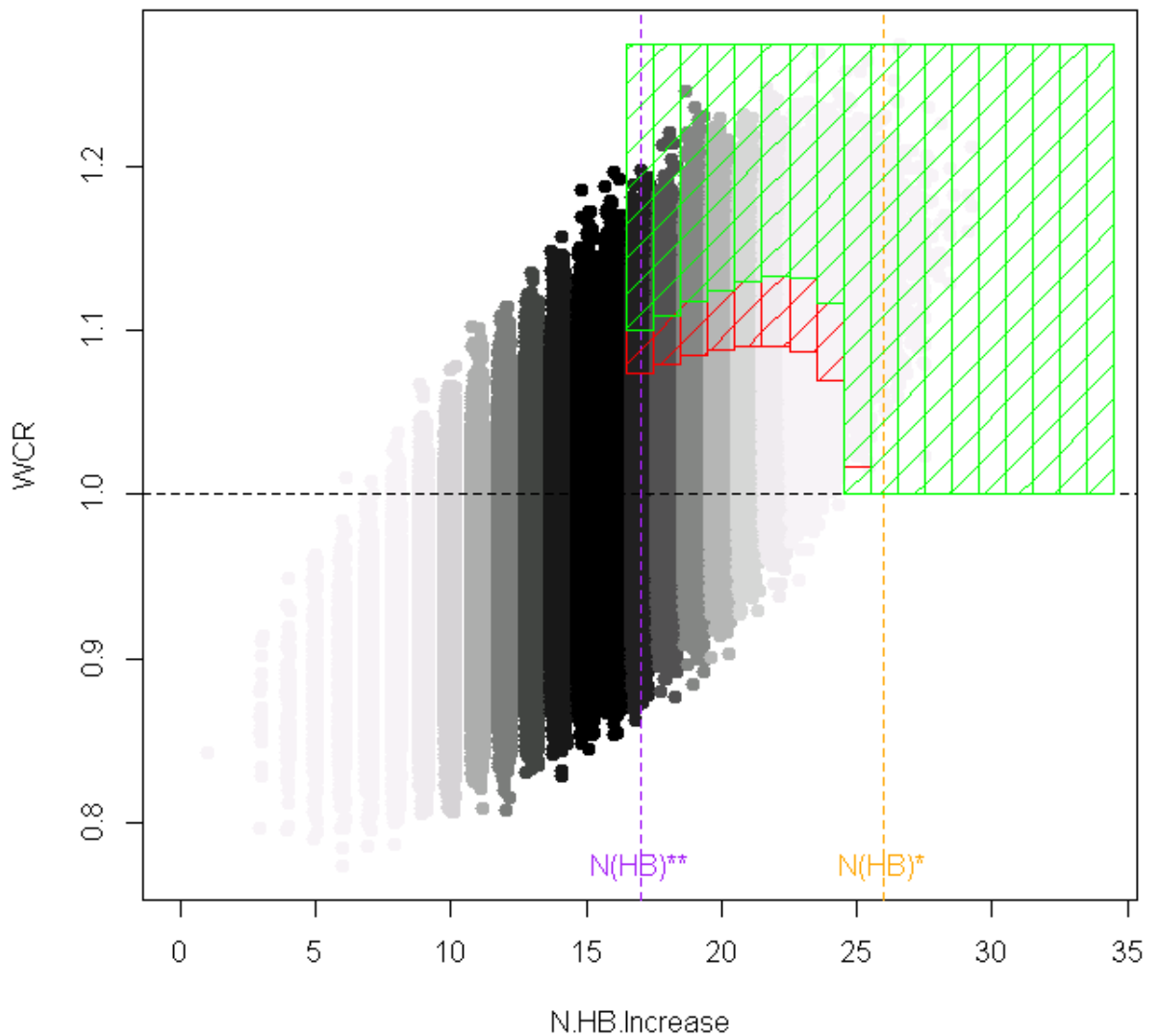


Figure D.4:  $WCR$  against  $N_{HB}$  for steady rates using fixed rate = 3 per 100,000 population with 34 HBs.  $N_{HB}^*$  is presented in the dotted vertical orange line while the dotted vertical purple line represents  $N_{HB}^{**}$ . The red shaped columns came from calculating the quantiles (cut-offs) using  $\mu$  and  $\sigma$  from many simulated cases (3,110,000) in the flat rate situation. They represent around 5% most extreme observations under the steady state. The green columns came from using `qnorm` to calculate cut-offs using  $\mu$  and  $\sigma$  from equations 9.3.1 and 9.3.2. They represent around 7% most extreme observations under the steady state. Darkness of the black colour reflects the number of simulated cases.

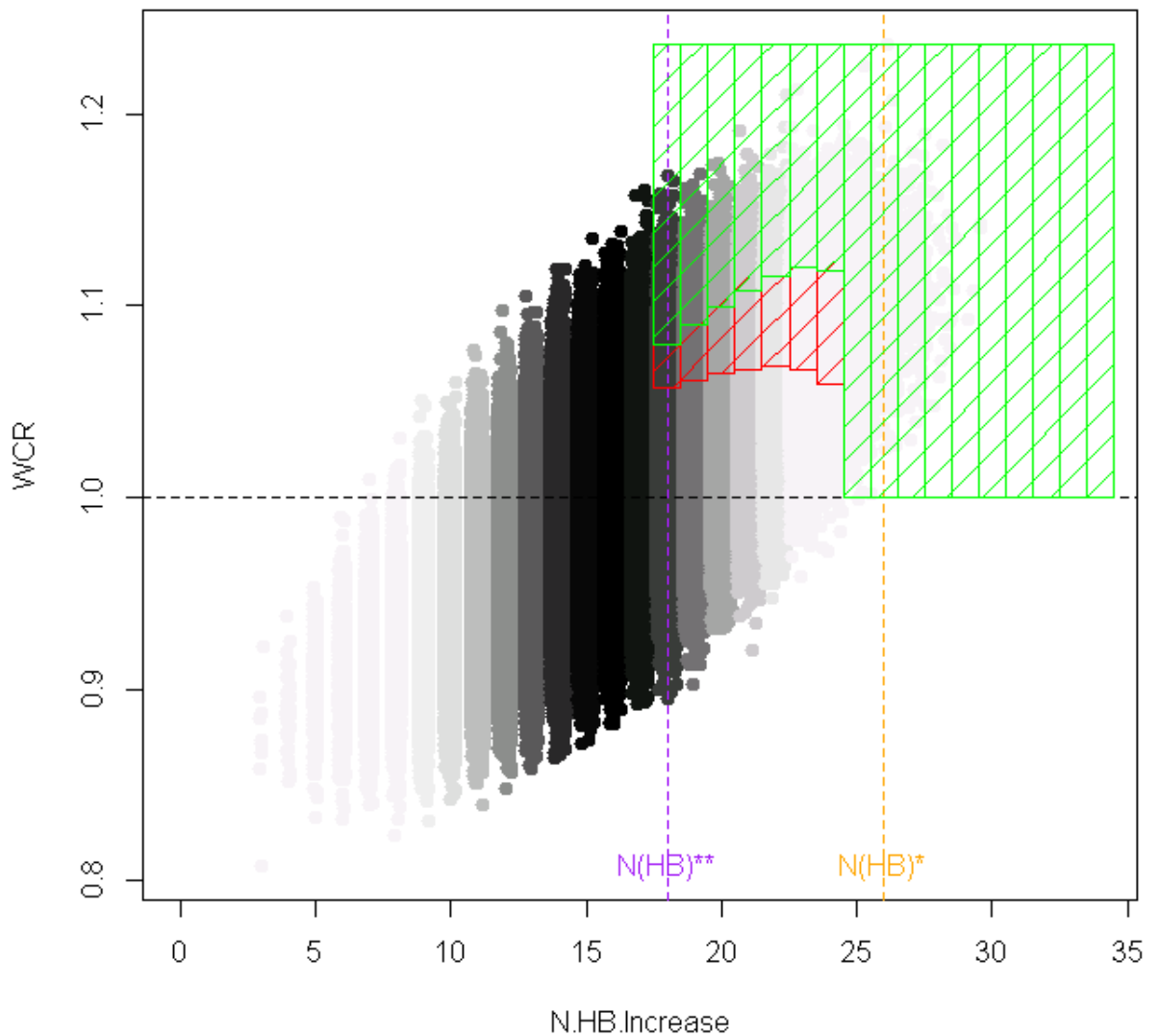


Figure D.5:  $WCR$  against  $N_{HB}$  for steady rates using fixed rate = 5 per 100,000 population with 34 HBs.  $N_{HB}^*$  is presented in the dotted vertical orange line while the dotted vertical purple line represents  $N_{HB}^{**}$ . The red shaped columns came from calculating the quantiles (cut-offs) using  $\mu$  and  $\sigma$  from many simulated cases (3,110,000) in the flat rate situation. They represent around 5% most extreme observations under the steady state. The green columns came from using `qnorm` to calculate cut-offs using  $\mu$  and  $\sigma$  from equations 9.3.1 and 9.3.2. They represent around 0.7% most extreme observations under the steady state. Darkness of the black colour reflects the number of simulated cases.

# Bibliography

- [1] W D E Adamson, S Maddi, C Robertson, S McDonagh, P J Molyneaux, K E Templeton, and W F Carman. 2009 pandemic influenza A (H1N1) virus in Scotland: geographically variable immunity in Spring 2010, following the winter outbreak. *European Surveillance*, 15(24):1–4, 2010.
- [2] O Y E Albarracin, A P Alencar, and L L Ho. CUSUM chart to monitor auto-correlated counts using Negative Binomial GARMA model. *Statistical Methods in Medical Research*, 27(9):2859–2871, 2017.
- [3] D J Alexander. A review of avian influenza in different bird species. *Veterinary Microbiology*, 74(1):3–13, 2000.
- [4] M Alheim, S M Johansson, D Hauzenberger, P Grufman, and J Holgersson. A flow cytometric crossmatch test for simultaneous detection of antibodies against donor lymphocytes and endothelial precursor cells. *Tissue Antigens*, 75(3):269–277, 2010.
- [5] M Almuqrin, C Robertson, and A Gray. Comparing methods for early detection systems for seasonal and pandemic influenza. In *Royal Statistical Society Conference*, Glasgow, 4 Sep 2017, University of Strathclyde.
- [6] R Amorós, D Conesa, M A Martínez-Beneito, and A López-Qulez. Statistical methods for detecting the onset of influenza outbreaks: a review. *REVSTAT-Statistical Journal*, 13(1):41–62, 2015.
- [7] ANZIC Influenza Investigators. Critical care services and 2009 H1N1 influenza in Australia and New Zealand. *New England Journal of Medicine*, 2009(361):1925–1934, 2009.
- [8] F Y Aoki, M D Macleod, P Paggiaro, O Carewicz, A E I Sawy, C Wat, M Griffiths, E Waalberg, P Ward, and the IMPACT Study Group. Early administration of oral oseltamivir increases the benefits of influenza treatment. *Journal of Antimicrobial Chemotherapy*, 51(1):123–129, 2003.

- [9] E Aramaki, S Maskawa, and M Morita. Twitter catches the flu: detecting influenza epidemics using Twitter. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pages 1568–1576. Association for Computational Linguistics, 27-31 July 2011, John McIntyre Conference Centre, Edinburgh, UK.
- [10] E P Armstrong and J Abarca. Pharmacoeconomic model to evaluate new influenza treatments. *Formulary*, 35(2):169–181, 2000.
- [11] N Balakrishnan. *Continuous Multivariate Distributions*. American Cancer Society, 2014. <https://onlinelibrary.wiley.com/doi/full/10.1002/9781118445112.stat01249>. Last accessed 15 January 2019.
- [12] M Bangert, H Gil, J Oliva, C Delgado, T Vega, S De Mateo, A Larrauri, and the Epidemiology Working Group of the Spanish Influenza Sentinel Surveillance System (SISSS). Pilot study to harmonize the reported influenza intensity levels within the Spanish Influenza Sentinel Surveillance System (SISSS) using the Moving Epidemic Method (MEM). *Epidemiology and Infection*, 145(4):715–722, 2017.
- [13] S Barik. New treatments for influenza. *BMC Medicine*, 10(1):104, 2012.
- [14] W H Barker and J P Mullooly. Pneumonia and influenza deaths during epidemics: implications for prevention. *Archives of Internal Medicine*, 142(1):85–89, 1982.
- [15] R W Barnard, K Pearce, and K C Richards. An inequality involving the generalized hypergeometric function and the arc length of an ellipse. *SIAM Journal on Mathematical Analysis*, 31(3):693–699, 2000.
- [16] R A Becker, J M Chambers, and A R Wilks. *The New S Language*. Pacific Grove, CA: Wadsworth & Brooks/Cole, 1988.
- [17] G Bédubourg and Y Le Strat. Evaluation and comparison of statistical methods for early temporal detection of outbreaks: A simulation-based study. *PLoS One*, 12(7):e0181227, 2017.
- [18] R Beyene and J A Boockvar. Disease-specific induced pluripotent stem cells. *Neurosurgery*, 63(6):N12–N12, 2008.
- [19] J F Bishop, M P Murnane, and R Owen. Australia’s winter with the 2009 pandemic influenza A (H1N1) virus. *New England Journal of Medicine*, 361(27):2591–2594, 2009.
- [20] G Boivin, I Hardy, G Tellier, and J Maziade. Predicting influenza infections during epidemics with use of a clinical case definition. *Clinical Infectious Diseases*, 31(5):1166–1169, 2000.

- [21] A W Bowman and A Azzalini. *Applied Smoothing Techniques for Data Analysis: the Kernel approach with S-Plus Illustrations*. Oxford University Press, Oxford, 1997.
- [22] G E P Box, G M Jenkins, and G C Reinsel. *Time Series Analysis: Forecasting and Control*, volume 734, Fifth edition. Wiley Series in Probability and Statistics, Canada, 2011.
- [23] C B Bridges, W W Thompson, M I Meltzer, G R Reeve, W J Talamonti, N J Cox, H A Lilac, H Hall, A Klimov, and K Fukuda. Effectiveness and cost-benefit of influenza vaccination of healthy working adults: a randomized controlled trial. *Journal of the American Medical Association*, 284(13):1655–1663, 2000.
- [24] I H Brown. The epidemiology and evolution of influenza viruses in pigs. *Veterinary Microbiology*, 74(1):29–46, 2000.
- [25] Centers for Disease Control and Prevention (CDC). Outbreak of swine-origin influenza A (H1N1) virus infection-Mexico, March-April 2009. *Morbidity and Mortality Weekly Report*, 58(17):467, 2009.
- [26] J D M Chambers and T D J Hastie. (eds.) *Statistical Models in S*, volume 251. Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, CA, 1992.
- [27] T Chan, Y Teng, and J Hwang. Detection of influenza-like illness aberrations by directly monitoring Pearson residuals of fitted negative binomial regression models. *BMC Public Health*, 15(1):168, 2015.
- [28] H Chen and C Huang. The use of a CUSUM residual chart to monitor respiratory syndromic data. *IIE Transactions*, 46(8):790–797, 2014.
- [29] K Choi and S B Thacker. An evaluation of influenza mortality surveillance, 1962–1979 i. Time series forecasts of expected pneumonia and influenza deaths. *American Journal of Epidemiology*, 113(3):215–226, 1981.
- [30] K Choi and S B Thacker. Mortality during influenza epidemics in the United States, 1967-1978. *American Journal of Public Health*, 72(11):1280–1283, 1982.
- [31] G Chowell, M A Miller, and C Viboud. Seasonal influenza in the United States, France, and Australia: transmission and prospects for control. *Epidemiology and Infection*, 136(06):852–864, 2008.
- [32] S D Collins. Excess mortality from causes other than influenza and pneumonia during influenza epidemics. *Public Health Reports (1896-1970)*, 47(46):2159–2179, 1932.

- [33] D Conesa, M A Martínez-Beneito, R Amorós, and A López-Quílez. Bayesian hierarchical Poisson models with a hidden Markov structure for the detection of influenza epidemic outbreaks. *Statistical Methods in Medical Research*, 24(2):206–223, 2015.
- [34] D Cooper, N Verlander, A Elliot, C Joseph, and G Smith. Can syndromic thresholds provide early warning of national influenza outbreaks? *Journal of Public Health*, 31(1):17–25, 2007.
- [35] R J Cooper, J R Hoffman, J G Bartlett, R E Besser, R Gonzales, J M Hickner, and M A Sande. Principles of appropriate antibiotic use for acute pharyngitis in adults: background. *Annals of Internal Medicine*, 134(6):509–517, 2001.
- [36] B J Cowling, K-H Chan, V J Fang, C K Y Cheng, R O P Fung, W Wai, J Sin, W H Seto, R Yung, D W S Chu, B C Chiu, P W Lee, M C Chiu, H C Lee, T M Uyeki, P M Houck, J S Peiris, and G M Leung. Facemasks and hand hygiene to prevent influenza transmission in households: a cluster randomized trial. *Annals of Internal Medicine*, 151(7):437–446, 2009.
- [37] B J Cowling, I O L Wong, L-M Ho, S Riley, and G M Leung. Methods for monitoring influenza surveillance data. *International Journal of Epidemiology*, 35(5):1314–1321, 2006.
- [38] B Cox, F Wuillaume, H Van Oyen, and S Maes. Monitoring of all-cause mortality in Belgium (Be-MOMO): a new and automated system for the early detection and quantification of the mortality impact of public health events. *International Journal of Public Health*, 55(4):251–259, 2010.
- [39] F S Dahlgren, D K Shay, H S Izurieta, R A Forshee, M Wernecke, Y Chillarige, Y Lu, J A Kelman, and C Reed. Evaluating oseltamivir prescriptions in Centers for Medicare and Medicaid Services medical claims records as an indicator of seasonal influenza in the United States. *Influenza and other Respiratory Viruses*, 12(4):465–474, 2018.
- [40] L Dailey, R E Watkins, and A J Plant. Timeliness of data sources used for influenza surveillance. *Journal of the American Medical Informatics Association*, 14(5):626–631, 2007.
- [41] M D De Jong, T T Thanh, T Hu Khanh, V M Hien, G J D Smith, N V Chau, B V Cam, P T Qui, D Q Ha, Y Guan, J M Peiris, T T Hien, and J Farrar. Oseltamivir resistance during treatment of influenza A (H5N1) infection. *New England Journal of Medicine*, 353(25):2667–2672, 2005.
- [42] P N Dinh, H T Long, N T K Tien, N T Hien, L T Q Mai, L H Phong, L Van Tuan, H Van Tan, N B Nguyen, P Van Tu, N T Phuong, World Health Organization/-

- Global Outbreak Alert, and Response Network Avian Influenza Investigation Team in Vietnam. Risk factors for human infection with avian influenza A H5N1, Vietnam, 2004. *Emerging Infectious Diseases*, 12(12):1841, 2006.
- [43] F Dominici, A McDermott, S L Zeger, and J M Samet. On the use of generalized additive models in time-series studies of air pollution and health. *American Journal of Epidemiology*, 156(3):193–203, 2002.
- [44] J Dushoff, J B Plotkin, C Viboud, D J D Earn, and L Simonsen. Mortality due to influenza in the United States—an annualized regression approach using multiple-cause mortality data. *American Journal of Epidemiology*, 163(2):181–187, 2006.
- [45] T C Eickhoff, I L Sherman, and R E Serfling. Observations on excess mortality associated with epidemic influenza. *Journal of the American Medical Association*, 176(9):776–782, 1961.
- [46] R C Elandt-Johnson and N L Johnson. *Survival Models and Data Analysis*, volume 74 of Wiley Classics Library. John Wiley & Sons, Canada, 1999.
- [47] C P Farrington, N J Andrews, A D Beale, and M A Catchpole. A statistical algorithm for the early detection of outbreaks of infectious disease. *Journal of the Royal Statistical Society, Series A (Statistics in Society)*, 159:547–563, 1996.
- [48] M Frisé, E Andersson, and L Schiöler. Robust outbreak surveillance of epidemics in Sweden. *Statistics in Medicine*, 28(3):476–493, 2009.
- [49] C Fu, R Chen, and Y Lo. Earlier prediction of influenza epidemic by hospital-based data in Taiwan. In *International Conference on Education, Management and Systems Engineering, Beijing, China, May 21–22, 2017, Eds EMSE 2017 International Committee*, number 2, pages 403–407, 2017, <http://www.dpi-proceedings.com/index.php/dtssehs/article/view/12808>. Last accessed 15 January 2019.
- [50] H Gao, K K Wong, Y Zheteyeva, J Shi, A Uzicanin, and J J Rainey. Comparing observed with predicted weekly influenza-like illness rates during the winter holiday break, United States, 2004–2013. *PLoS One*, 10(12):1–11, 2015.
- [51] T Garske, J Legrand, C A Donnelly, H Ward, S Cauchemez, C Fraser, N M Ferguson, and A C Ghani. Assessing the severity of the novel influenza A/H1N1 pandemic. *British Medical Journal*, 339:b2840, 2009.
- [52] General Authority for Statistics in Saudi Arabia, 2016. <http://www.cdsi.gov.sa/en/node>. Last accessed 02 July 2018.

- [53] S Gerbier-Colomban, V e Potinet-Pagliaroli, and M Metzger. Can epidemic detection systems at the hospital level complement regional surveillance networks: Case study with the influenza epidemic? *BMC Infectious Diseases*, 14(1):381, 2014.
- [54] T C Germann, K Kadau, I M Longini, and C A Macken. Mitigation strategies for pandemic influenza in the United States. *Proceedings of the National Academy of Sciences*, 103(15):5935–5940, 2006.
- [55] J Ginsberg, M H Mohebbi, R S Patel, L Brammer, M S Smolinski, and L Brilliant. Detecting influenza epidemics using search engine query data. *Nature*, 457(7232):1012–1014, 2009.
- [56] N L Goddard, J Kyncl, and J M Watson. Appropriateness of thresholds currently used to describe influenza activity in England. *Communicable Disease and Public Health/PHLS*, 6(3):238–245, 2003.
- [57] W B Goggins, J Woo, S Ho, E Y Y Chan, and P H Chau. Weather, season, and daily stroke admissions in Hong Kong. *International Journal of Biometeorology*, 56(5):865–872, 2012.
- [58] D K Grayson. Donner party deaths: A demographic assessment. *Journal of Anthropological Research*, 46(3):223–242, 1990.
- [59] H K Green, A Charlett, J Moran-Gilad, D Fleming, H Durnall, D Rh Thomas, S Cottrell, B Smyth, C Kearns, A J Reynolds, G E Smith, A J Elliot, J Ellis, M Zambon, J M Watson, J McMenamini, and R G Pebody. Harmonizing influenza primary-care surveillance in the United Kingdom: piloting two methods to assess the timing and intensity of the seasonal epidemic across several general practice-based surveillance schemes. *Epidemiology and Infection*, 143(1):1–12, 2015.
- [60] B A Griffin, A K Jain, J Davies-Cole, C Glymph, G Lum, S C Washington, and M A Stoto. Early detection of influenza outbreaks using the DC Department of Health’s syndromic surveillance system. *BMC Public Health*, 9(1):483, 2009.
- [61] O A Grigg, V T Farewell, and D J Spiegelhalter. Use of risk-adjusted cusum and rsprtcharts for monitoring in medical contexts. *Statistical Methods in Medical Research*, 12(2):147–170, 2003.
- [62] T J Hastie and R J Tibshirani. *Generalized Additive Models, Monographs on Statistics and Applied Probability*, volume 43. Chapman & Hall, London, 1990.
- [63] A C Hayward, R Harling, S Wetten, A M Johnson, S Munro, J Smedley, S Murad, and J M Watson. Effectiveness of an influenza vaccine programme for care home



- staff to prevent death, morbidity, and health service use among residents: cluster randomised controlled trial. *British Medical Journal*, 333(7581):1241, 2006.
- [64] Health Protection Scotland (HPS), 23 June 2009. <http://www.hps.scot.nhs.uk/resourcedocument.aspx?id=425>. Last accessed 02 July 2018.
- [65] Health Protection Scotland. Reporting rates of influenza (flu) consultations from General Practitioners in Scotland, 10 October 2016. <http://www.hps.scot.nhs.uk/resourcedocument.aspx?id=5605>. Last accessed 02 July 2018.
- [66] Healthcare Improvement Scotland, 2015. <https://www.sudiscotland.org.uk/regionalmaps/sudi-paediatricians/>. Last accessed 18 January 2019.
- [67] J D Healy. A note on multivariate CUSUM procedures. *Technometrics*, 29(4):409–412, 1987.
- [68] L Held, M Hofmann, M Höhle, and V Schmid. A two-component model for counts of infectious diseases. *Biostatistics*, 7(3):422–437, 2006.
- [69] R Heller, D Small, and P Rosenbaum. R package ‘crossmatch’. 2012. <http://cran.stat.nus.edu.sg/web/packages/crossmatch/crossmatch.pdf>. Last accessed 02 July 2018.
- [70] L Hobbs, K Wong, and I Johnson. Data-driven approach of CUSUM algorithm in temporal aberrant event detection using interactive web applications. *Canadian Journal of Public Health*, 107(1):9–15, 2016.
- [71] M Höhle and A Mazick. Aberration detection in R illustrated by Danish mortality monitoring. *Biosurveillance: Methods and Case Studies*, 3:215–237, 2010.
- [72] D W Hosmer and S Lemeshow. Confidence interval estimation of interaction. *Epidemiology*, 3(5):452–456, 1992.
- [73] J Housworth and A D Langmuir. Excess mortality from epidemic influenza, 1957–1966. *American Journal of Epidemiology*, 100(1):40–48, 1974.
- [74] C M Hurvich, J S Simonoff, and C-L Tsai. Smoothing parameter selection in non-parametric regression using an improved Akaike information criterion. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 60(2):271–293, 1998.
- [75] R Jones. A previously uncharacterized infectious-like event leading to spatial spread of deaths across England and Wales: Characteristics of the most recent event and a time series for past events. *British Journal of Medicine and Medical Research*, 5(11):1361–1380, 2015.

- [76] M J Kane, N Price, M Scotch, and P Rabinowitz. Comparison of ARIMA and Random Forest time series models for prediction of avian influenza H5N1 outbreaks. *BMC Bioinformatics*, 15(1):276, 2014.
- [77] K M Kunisaki and E N Janoff. Influenza in immunosuppressed populations: a review of infection frequency, morbidity, mortality, and vaccine responses. *The Lancet Infectious Diseases*, 9(8):493–504, 2009.
- [78] K Laake and J M Sverre. Winter excess mortality: a comparison between Norway and England plus Wales. *Age and Ageing*, 25(5):343–348, 1996.
- [79] D Lamson, N Renwick, V Kapoor, Z Liu, G Palacios, J Ju, A Dean, K St. George, T Briese, and W Ian L. Masstag polymerase-chain-reaction detection of respiratory pathogens, including a new rhinovirus genotype, that caused influenza-like illness in New York State during 2004–2005. *The Journal of Infectious Diseases*, 194(10):1398–1402, 2006.
- [80] E H Y Lau, C K Y Cheng, D K M Ip, and B J Cowling. Situational awareness of influenza activity based on multiple streams of surveillance data using multivariate dynamic linear model. *PLoS One*, 7(5):1–6, 2012.
- [81] Y Le Strat and F Carrat. Monitoring epidemiologic surveillance data using hidden Markov models. *Statistics in Medicine*, 18(24):3463–3478, 1999.
- [82] D S Leland and C C Ginocchio. Role of cell culture for virus detection in the age of technology. *Clinical Microbiology Reviews*, 20(1):49–78, 2007.
- [83] J Lessler, N G Reich, and D A T Cummings. Outbreak of 2009 pandemic influenza A (H1N1) at a New York city school. *New England Journal of Medicine*, 361(27):2628–2636, 2009.
- [84] J C Lin and K L Nichol. Excess mortality due to pneumonia or influenza during influenza seasons among persons with acquired immunodeficiency syndrome. *Archives of Internal Medicine*, 161(3):441–446, 2001.
- [85] I M Longini, M E Halloran, A Nizam, and Y Yang. Containing pandemic influenza with antiviral agents. *American Journal of Epidemiology*, 159(7):623–633, 2004.
- [86] I M Longini, A Nizam, S Xu, K Ungchusak, W Hanshaworakul, D A Cummings, and M E Halloran. Containing pandemic influenza at the source. *Science*, 309(5737):1083–1087, 2005.
- [87] B Lupiani and S M Reddy. The history of avian influenza. *Comparative Immunology, Microbiology and Infectious Diseases*, 32(4):311–323, 2009.

- [88] S Ly, T Arashiro, V Ieng, R Tsuyuoka, A Parry, P Horwood, S Heng, S Hamid, K Vandemaële, S Chin, B Sar, and Y Arima. Establishing seasonal and alert influenza thresholds in Cambodia using the WHO method: implications for effective utilization of influenza surveillance in the tropics and subtropics. *Western Pacific Surveillance and Response Journal: WPSAR*, 8(1):22, 2017.
- [89] J Makaroon, C Pittayawonganon, D Gross, and M McMorrow. An evaluation of Influenza-like Illness (ILI) epidemic alert thresholds in two provinces of Thailand, 2007-2010. *OSIR Journal*, 6(1):13–18, 2016.
- [90] M A Martínez-Beneito, D Conesa, A López-Quílez, and A López-Maside. Bayesian Markov switching models for the early detection of influenza epidemics. *Statistics in Medicine*, 27(22):4455–4468, 2008.
- [91] P McCullagh and J A Nelder. *Generalized Linear Models*, volume 37. CRC Press. London, 1989.
- [92] N-A M Molinari, I R Ortega-Sanchez, M L Messonnier, W W Thompson, P M Wortley, E Weintraub, and C B Bridges. The annual impact of seasonal influenza in the US: measuring disease burden and costs. *Vaccine*, 25(27):5086–5096, 2007.
- [93] G C Monamele, M-A Vernet, R FJ Nsaibirni, J J R Bigna, S Kenmoe, M R Njankouo, and R Njouom. Associations between meteorological parameters and influenza activity in a subtropical country: Case of five sentinel sites in Yaoundé-Cameroon. *PLoS One*, 12(10):1–12, 2017.
- [94] D C Montgomery. *Introduction to Statistical Quality Control*. John Wiley & Sons, New York, 2007.
- [95] A S Monto, D M Fleming, D Henry, R De Groot, M Makela, T Klein, M Elliott, O N Keene, and C Y Man. Efficacy and safety of the neuraminidase inhibitor zanamivir in the treatment of influenza A and B virus infections. *Journal of Infectious Diseases*, 180(2):254–261, 1999.
- [96] A S Monto, S Gravenstein, M Elliott, M Colopy, and J Schweinle. Clinical signs and symptoms predicting influenza infection. *Archives of Internal Medicine*, 160(21):3243–3247, 2000.
- [97] A S Monto and K M Sullivan. Acute respiratory illness in the community. Frequency of illness and the agents involved. *Epidemiology and Infection*, 110(01):145–160, 1993.
- [98] K F Mulchrone and K R Choudhury. Fitting an ellipse to an arbitrary shape: implications for strain analysis. *Journal of Structural Geology*, 26(1):143–153, 2004.

- [99] J Mullooly, M Bennett, M Hornbrook, and W Barker. Cost-effectiveness of influenza vaccination programs in an HMO: the experience of Kaiser Permanente, northwest region. In *Excerpta Medica International Congress Series*, volume 1019, page 53. Elsevier, 1993.
- [100] J L K Murray, D F P Marques, R L Cameron, A Potts, J Bishop, B Von Wissmann, N William, A J Reynolds, C Robertson, and J McMenamin. Moving epidemic method (MEM) applied to virology data as a novel real time tool to predict peak in seasonal influenza healthcare utilisation. The Scottish experience of the 2017/18 season to date. *Eurosurveillance*, 23(11), 2018.
- [101] National Health Service (NHS), 08 September 2014. <https://www.nhs.uk/using-the-nhs/nhs-services/visiting-or-moving-to-england/how-to-access-nhs-services-in-england/>. Last accessed 01 December 2018.
- [102] National Health Service (NHS), 16 January, 2016. <http://www.gov.scot/Topics/Health/NHS-Workforce/NHS-Boards>. Last accessed 02 July 2018.
- [103] National Records of Scotland, 26 April 2016. <http://www.nrscotland.gov.uk/news/2015/scotlands-population-at-its-highest-ever>. Last accessed 02 July 2018.
- [104] J A Nelder and R J Baker. Generalized linear models. *Encyclopedia of Statistical Sciences*, 135:370–384, Wiley Online Library New York, USA, 1972.
- [105] K M Neuzil, C Hohlbein, and Y Zhu. Illness among schoolchildren during influenza season: effect on school absenteeism, parental absenteeism from work, and secondary illness in families. *Archives of Pediatrics and Adolescent Medicine*, 156(10):986–991, 2002.
- [106] K M Neuzil, P F Wright, E F Mitchel, and M R Griffin. The burden of influenza illness in children with asthma and other chronic medical conditions. *The Journal of Pediatrics*, 137(6):856–864, 2000.
- [107] A K N’gattia, D Coulibaly, N T Nzussouo, H A Kadjo, D Chérif, Y Traoré, B K Kouakou, P D Kouassi, K D Ekra, N S Dagnan, T Williams, and I Tiembré. Effects of climatological parameters in modeling and forecasting seasonal influenza transmission in Abidjan, Cote d’Ivoire. *BMC Public Health*, 16(1):972, 2016.
- [108] K L Nichol, L Baken, and A Nelson. Relation between influenza vaccination and outpatient visits, hospitalization, and mortality in elderly persons with chronic lung disease. *Annals of Internal Medicine*, 130(5):397–403, 1999.

- [109] K L Nichol, A Lind, K L Margolis, M Murdoch, R McFadden, M Hauge, S Magnan, and M Drake. The effectiveness of vaccination against influenza in healthy, working adults. *New England Journal of Medicine*, 333(14):889–893, 1995.
- [110] K L Nichol, K L Margolis, J Wuorenma, and T Von Sternberg. The efficacy and cost effectiveness of vaccination against influenza among elderly persons living in the community. *New England Journal of Medicine*, 331(12):778–784, 1994.
- [111] K G Nicholson. Impact of influenza and respiratory syncytial virus on mortality in England and Wales from January 1975 to December 1990. *Epidemiology and Infection*, 116(01):51–63, 1996.
- [112] A Noufaily, P Farrington, P Garthwaite, D G Enki, N Andrews, and A Charlett. Detection of infectious disease outbreaks from laboratory data with reporting delays. *Journal of the American Statistical Association*, 111(514):488–499, 2016.
- [113] A Ozonoff, S Sukpraput, and P Sebastiani. Modeling seasonality of influenza with Hidden Markov Models. *Proceedings of the American Statistical Association*, 2006. [https://www.researchgate.net/profile/Paola\\_Sebastiani/publication/267206133\\_Modeling\\_seasonality\\_of\\_influenza\\_with\\_Hidden\\_Markov\\_Models/links/55475f610cf234bdb21dda49/Modeling-seasonality-of-influenza-with-Hidden-Markov-Models.pdf](https://www.researchgate.net/profile/Paola_Sebastiani/publication/267206133_Modeling_seasonality_of_influenza_with_Hidden_Markov_Models/links/55475f610cf234bdb21dda49/Modeling-seasonality-of-influenza-with-Hidden-Markov-Models.pdf). Last accessed 15 January 2019.
- [114] X Pang, X Wang, Q Wang, T Gao, and X Li. Infectious disease surveillance and early warning system during Beijing Olympic Games. In *Early Warning for Infectious Disease Outbreak, Theory and Practice (Ed. W. Yang)*, pages 163–180. Academic Press, 2017.
- [115] R Páscoa, A P Rodrigues, S Silva, B Nunes, and C Martins. Comparison between influenza coded primary care consultations and national influenza incidence obtained by the General Practitioners Sentinel Network in Portugal from 2012 to 2017. *PLoS One*, 13(2):1–10, 2018.
- [116] S Paul, O Mgbere, R Arafat, B Yang, and E Santos. Modeling and forecasting Influenza-like Illness (ILI) in Houston, Texas using three surveillance data capture mechanisms. *Online Journal of Public Health Informatics*, 9(2):e187, 2017.
- [117] K I Pedersen, P E Mogensen, and B H Fleury. A stochastic model of the temporal and azimuthal dispersion seen at the base station in outdoor propagation environments. *IEEE Transactions on Vehicular Technology*, 49(2):437–447, 2000.
- [118] J Peiris, Y Guan, D Markwell, P Ghose, R G Webster, and K F Shortridge. Co-circulation of avian H9N2 and contemporary “human” H3N2 influenza A viruses in

- pigs in southeastern China: potential for genetic reassortment? *Journal of Virology*, 75(20):9679–9686, 2001.
- [119] M Peiris, K Y Yuen, C W Leung, K H Chan, PLS Ip, R W M Lai, W K Orr, and K F Shortridge. Human infection with influenza H9N2. *The Lancet*, 354(9182):916–917, 1999.
- [120] C Pelat, I Bonmarin, M Ruello, A Fouillet, C Caserio-Schönemann, D Levy-Bruhl, Y Le Strat, and The Regional Influenza Study Group. Improving regional influenza surveillance through a combination of automated outbreak detection methods: the 2015/16 season in France. *European Surveillance*, 22:32, 2017. <https://books.google.co.uk/books?hl=en&lr=&id=rNt5CgAAQBAJ&oi=fnd&pg=PR7&dq=G+EP+Box,+G+M+Jenkins,+and+G+C+Reinsel.+Time+Series+Analysis:+Forecasting+and+Control,+volume+734.+John+Wiley+%26+Sons,+2011&ots=DJ53s0pZRF&sig=dzkCqqtv448LtfP6v8fRG465a3U#v=onepage&q&f=false>. Last accessed 15 January 2018.
- [121] R Perez-Padilla, D De La Rosa-Zamboni, S Ponce De Leon, M Hernandez, F Quiñones-Falconi, E Bautista, A Ramirez-Venegas, J Rojas-Serrano, C E Ormsby, A Corrales, A Higuera, E Mondragon, J A Cordova-Villalobos, and the INER Working Group on Influenza. Pneumonia and respiratory failure from swine-origin influenza A (H1N1) in Mexico. *New England Journal of Medicine*, 361(7):680–689, 2009.
- [122] F Pervaiz, M Pervaiz, N A Rehman, and U Saif. Flubreaks: early epidemic detection from Google flu trends. *Journal of Medical Internet Research*, 14(5):e725, 2012.
- [123] Population Reference Bureau, 2016. <http://www.prb.org/Publications/Articles/2010/northkorea-population.aspx>. Last accessed 02 July 2018.
- [124] Population Reference Bureau, 2016. <http://www.prb.org/DataFinder/Topic/Rankings.aspx?ind=15>. Last accessed 02 July 2018.
- [125] Population Reference Bureau, 2016. <http://www.prb.org/Publications/Articles/2012/south-korea-population.aspx>. Last accessed 02 July 2018.
- [126] J Potter, D J Stott, M A Roberts, A G Elder, B O’Donnell, P V Knight, and W F Carman. Influenza vaccination of health care workers in long-term-care hospitals reduces the mortality of elderly patients. *Journal of Infectious Diseases*, 175(1):1–6, 1997.
- [127] A M Presanis, D De Angelis, Swine Flu Investigation Team, New York City Department of Health, A Hagy, C Reed, S Riley, B Cooper, P Biedrzycki, L Finelli,

- and M Lipsitch. The severity of pandemic H1N1 influenza in the United States, April–July 2009. *PLoS Currents*, 1:1–12, 2009.
- [128] G F Pyle. *The Diffusion of Influenza: Patterns and Paradigms*. Rowman & Littlefield, USA, 1986.
- [129] R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2008. ISBN 3-900051-07-0.
- [130] L R Rabiner. A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2):257–286, 1989.
- [131] Y Rao and B McCabe. Real-time surveillance for abnormal events: the case of influenza outbreaks. *Statistics in Medicine*, 35(13):2206–2220, 2016.
- [132] T M Rath, M Carreras, and P Sebastiani. Automated detection of influenza epidemics with hidden Markov models. In *Advances in Intelligent Data Analysis V, Proceedings of the 5th Intelligent Symposium on Intelligent Data Analysis, IDA 2003, Berlin, Germany, August 28-30, 2003, Eds M R Berthold, H-T Lenz, E Bradley, R Kruse, and C Borgelt, Lecture Notes in Computer Science, vol 2810, Springer, Berlin, Heidelberg*.
- [133] T A Reichert, L Simonsen, A Sharma, S A Pardo, D S Fedson, and M A Miller. Influenza and the winter increase in mortality in the United States, 1959–1999. *American Journal of Epidemiology*, 160(5):492–502, 2004.
- [134] J Rocklöv and B Forsberg. The effect of temperature on mortality in Stockholm 1998–2003: A study of lag structures and heatwave effects. *Scandinavian Journal of Social Medicine*, 36(5):516–523, 2008.
- [135] P A Rogerson and I Yamada. Approaches to syndromic surveillance when data consist of small regional counts. *Morbidity and Mortality Weekly Report*, 53:79–85, 2004.
- [136] R Sakia. The Box-Cox transformation technique: a review. *Journal of the Royal Statistical Society. Series D (The Statistician)*, 41(2):169–178, 1992.
- [137] D Schanzer, T Tam, J Langley, and B Winchester. Influenza-attributable deaths, Canada 1990–1999. *Epidemiology and Infection*, 135(7):1109–1116, 2007.
- [138] S Schrell, A Ziemann, L Garcia-Castrillo Riesgo, N Rosenkötter, J Llorca, D Popa, and T Krafft. Local implementation of a syndromic influenza surveillance system using emergency department data in Santander. *Public Health*, 35(3):397–403, 2013.

- [139] R E Serfling. Methods for current statistical analysis of excess pneumonia-influenza deaths. *Public Health Reports*, 78(6):494, 1963.
- [140] A Signorini, A M Segre, and P M Polgreen. The use of Twitter to track levels of disease activity and public concern in the US during the influenza A H1N1 pandemic. *PLoS One*, 6(5):1–10, 2011.
- [141] L Simonsen, W C Blackwelder, T A Reichert, and M A Miller. Estimating deaths due to influenza and respiratory syncytial virus. *Journal of the American Medical Association*, 289(19):2499–2502, 2003.
- [142] B Singh, N Savill, N Ferguson, C Robertson, and M Woolhouse. Rapid detection of pandemic influenza in the presence of seasonal influenza. *BMC Public Health*, 10(1):726, 2010.
- [143] R Snacken, J Lion, V Van Casteren, R Cornelis, F Yane, M Mombaerts, W Aelvoet, and A Stroobant. Five years of sentinel surveillance of acute respiratory infections (1985–1990): the benefits of an influenza early warning system. *European Journal of Epidemiology*, 8(4):485–490, 1992.
- [144] E Spackman, D A Senne, T J Myers, L L Bulaga, L P Garber, M L Perdue, K Lohman, L T Daum, and D L Suarez. Development of a real-time reverse transcriptase PCR assay for type A influenza virus and the avian H5 and H7 hemagglutinin subtypes. *Journal of Clinical Microbiology*, 40(9):3256–3260, 2002.
- [145] A Spreco, O Eriksson, Ö Dahlström, B J Cowling, and T Timpka. Evaluation of nowcasting for detecting and predicting local influenza epidemics, Sweden, 2009–2014. *Emerging Infectious Diseases*, 24(10):1868, 2018.
- [146] S H Steiner, K Grant, M Coory, and H A Kelly. Detecting the start of an influenza outbreak using exponentially weighted moving average charts. *BMC Medical Informatics and Decision Making*, 10(1):37, 2010.
- [147] A Stuart and K Ord. *Kendall's Advanced Theory of Statistics*, Arnold, volume 1. London, 1994.
- [148] K Subbarao, A Klimov, J Katz, H Regnery, W Lim, H Hall, M Perdue, D Swayne, C Bender, J Huang, M Hemphill, T Rowe, M Shaw, X Xu, K Fukuda, and N Cox. Characterization of an avian influenza A (H5N1) virus isolated from a child with a fatal respiratory illness. *Science*, 279(5349):393–396, 1998.
- [149] B L Taylor, H E Montgomery, A Rhodes, and C L Sprung. Protection of patients and staff during a pandemic. *Intensive Care Medicine*, 36(1):45–54, 2010.



- [150] Y Teng, D Bi, X Guo, D Feng, and Y Tong. Epidemic potential for human infection with influenza A (H7N9) virus in China through web search behaviors: A data-driven study. *bioRxiv*, page 168, 2017. doi 10.1101/168112. <https://www.biorxiv.org/content/early/2017/08/26/168112>. Last accessed 06 December 2018.
- [151] The Writing Committee of the World Health Organization (WHO) Consultation on Human Influenza A/H5. Avian influenza A (H5N1) infection in humans. *New England Journal of Medicine*, 353(13):1374–1385, 2005.
- [152] W W Thompson, L Comanor, and D K Shay. Epidemiology of seasonal influenza: use of surveillance data and statistical models to estimate the burden of disease. *Journal of Infectious Diseases*, 194(Supplement 2):S82–S91, 2006.
- [153] W W Thompson, M R Moore, E Weintraub, P-Y Cheng, X Jin, C B Bridges, J S Bresee, and D K Shay. Estimating influenza-associated deaths in the United States. *American Journal of Public Health*, 99(S2):S225–S230, 2009.
- [154] W W Thompson, D K Shay, E Weintraub, L Brammer, N Cox, L J Anderson, and K Fukuda. Mortality associated with influenza and respiratory syncytial virus in the United States. *Journal of the American Medical Association*, 289(2):179–186, 2003.
- [155] S A Tweed, D M Skowronski, S T David, A Larder, M Petric, W Lees, Y Li, J Katz, M Krajdén, R Tellier, C Halpert, M Hirst, C Astell, D Lawrence, and A Mak. Human illness from avian influenza H7N3, British Columbia. *Emerging Infectious Diseases*, 10(12):2196, 2004.
- [156] S Unkel, C Farrington, P H Garthwaite, C Robertson, and N Andrews. Statistical methods for the prospective detection of infectious disease outbreaks: a review. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 175(1):49–82, 2012.
- [157] M Urashima, T Segawa, M Okazaki, M Kurihara, Y Wada, and H Ida. Randomized trial of vitamin D supplementation to prevent seasonal influenza A in schoolchildren. *The American Journal of Clinical Nutrition*, 91(5):1255–1260, 2010.
- [158] C Van Den Dool, E Hak, J Wallinga, A M Van Loon, J W J Lammers, and M J M Bonten. Symptoms of influenza virus infection in hospitalized patients. *Infection Control*, 29(04):314–319, 2008.
- [159] R A Van Deusen, V S Hinshaw, D A Senne, and D Pellacani. Micro neuraminidase-inhibition assay for classification of influenza A virus neuraminidases. *Avian Diseases*, 27(3):745–750, 1983.

- [160] T Vega and J E Lozano, November 2018. <https://cran.r-project.org/web/packages/mem/mem.pdf>. Last accessed 18 January 2019.
- [161] T Vega, J E Lozano, T Meerhoff, R Snacken, J Beauté, P Jorgensen, R Ortiz De Lejarazu, L Domegan, J Mossong, J Nielsen, R Born, A Larrauri, and C Brown. Influenza surveillance in Europe: comparing intensity levels calculated using the moving epidemic method. *Influenza and other Respiratory Viruses*, 9(5):234–246, 2015.
- [162] T Vega, J E Lozano, T Meerhoff, R Snacken, J Mott, R Ortiz De Lejarazu, B Nunes, and R Born. Influenza surveillance in Europe: establishing epidemic thresholds by the moving epidemic method. *Influenza and other Respiratory Viruses*, 7(4):546–558, 2013.
- [163] C Viboud, P-Y Boëlle, K Pakdaman, F Carrat, A-J Valleron, and A Flahault. Influenza epidemics in the United States, France, and Australia, 1972-1997. *Emerging Infectious Diseases*, 10(1):9–32, 2004.
- [164] M M Wagner, F-C Tsui, J U Espino, V M Dato, D F Sittig, R A Caruana, L F McGinnis, D W Deerfield, M J Druzdel, and D B Fridsma. The emerging science of very early detection of disease outbreaks. *Journal of Public Health Management and Practice*, 7(6):51–59, 2001.
- [165] C Wang, Y Li, W Feng, K Liu, S Zhang, F Hu, S Jiao, X Lao, H Ni, and G Xu. Supplementary Materials: Epidemiological features and forecast model analysis for the morbidity of influenza in Ningbo, China, 2006–2014. *International Journal for Environmental Research and Public Health*, 14:1, 2016.
- [166] C Wang, Y Li, W Feng, K Liu, S Zhang, F Hu, S Jiao, X Lao, H Ni, and G Xu. Epidemiological features and forecast model analysis for the morbidity of influenza in Ningbo, China, 2006–2014. *International Journal for Environmental Research and Public Health*, 14(6):559, 2017.
- [167] R D Wang, Y G Jiang, G M Zhao, X Q Guo, and E Michael. Outbreak Gold Standard selection to provide optimized threshold for infectious diseases early-alert based on China Infectious Disease Automated-alert and Response System. *Journal of Huazhong University of Science and Technology [Medical Sciences]*, 37(6):833–841, 2017.
- [168] X Wang, S Wu, C R MacIntyre, H Zhang, W Shi, X Peng, W Duan, P Yang, Y Zhang, and Q Wang. Using an adjusted Serfling regression model to improve the early warning at the arrival of peak timing of influenza in Beijing. *PLoS One*, 10(3):1–14, 2015.

- [169] R E Watkins, S Eagleson, B Veenendaal, G Wright, and A J Plant. Applying CUSUM-based methods for the detection of outbreaks of Ross River virus disease in Western Australia. *BMC Medical Informatics and Decision Making*, 8(1):37, 2008.
- [170] R E Watkins, S Eagleson, B Veenendaal, G Wright, and A J Plant. Disease surveillance using a hidden Markov model. *BMC Medical Informatics and Decision Making*, 9(1):39, 2009.
- [171] R J Webby and R G Webster. Are we ready for pandemic influenza? *Science*, 302(5650):1519–1522, 2003.
- [172] R A Weinstein, C B Bridges, M J Kuehnert, and C B Hall. Transmission of influenza: implications for control in health care settings. *Clinical Infectious Diseases*, 37(8):1094–1101, 2003.
- [173] Y G Weldeselassie, H J Whitaker, and C P Farrington. Use of the self-controlled case-series method in vaccine safety studies: review and recommendations for best practice. *Epidemiology and Infection*, 139(12):1805–1817, 2011.
- [174] M West and J Harrison. *Bayesian Forecasting and Dynamic Models*. Springer Science & Business Media. New York, 2006.
- [175] E Whimbey, R E Champlin, R B Couch, J A Englund, J M Goodrich, I Raad, D Przepiorka, V A Lewis, N Mirza, H Yousuf, J J Tarrand, and G P Bodey. Community respiratory virus infections among hospitalized adult bone marrow transplant recipients. *Clinical Infectious Diseases*, 22(5):778–782, 1996.
- [176] H J Whitaker, M N Hocine, and C P Farrington. The methodology of self-controlled case series studies. *Statistical Methods in Medical Research*, 18(1):7–26, 2009.
- [177] R Wilcox. Kolmogorov–Smirnov Test. *Encyclopedia of Biostatistics*, 4(1):2151–2169, 2005.
- [178] S Wood and F Scheipl. gamm4: Generalized additive mixed models using mgcv and lme4. *R package version 0.2-3*, 2014.
- [179] S N Wood. mgcv: GAMs and generalized ridge regression for R. *R News*, 1(2):20–25, 2001.
- [180] World Health Organization, Regional Office for Europe, U.S. Centers, C S Brown, P Jorgensen, J A Mott, D Pereyaslov, L Martiryosan, T Meerhoff, and D Gross. WHO Regional Office for Europe guidance for sentinel influenza surveillance in humans. *Copenhagen, Denmark: WHO Regional Office for Europe*, 2011. Available at:[http://www.euro.who.int/\\_\\_data/assets/pdf\\_file/0020/90443/E92738.pdf](http://www.euro.who.int/__data/assets/pdf_file/0020/90443/E92738.pdf). Last accessed 14 August 2018.

- [181] World Health Organization (WHO). *International travel and health: situation as on 1 January 2010*. World Health Organization, Switzerland, 2010.
- [182] World Health Organization (WHO), 20 October 2017. [http://www.who.int/influenza/surveillance\\_monitoring/updates/latest\\_update\\_GIP\\_surveillance/en/](http://www.who.int/influenza/surveillance_monitoring/updates/latest_update_GIP_surveillance/en/). Last accessed 02 July 2018.
- [183] World Health Organization (WHO). Vaccines against influenza. WHO position paper—November 2012. *Weekly Epidemiological Record*, 87:461–76, 2012.
- [184] World Health Organization (WHO) surveillance case definitions for ILI and ARI, 08 January 2014. [http://www.who.int/influenza/surveillance\\_monitoring/ili\\_sari\\_surveillance\\_case\\_definition/en/](http://www.who.int/influenza/surveillance_monitoring/ili_sari_surveillance_case_definition/en/). Last accessed 02 July 2018.
- [185] Worldometers, 09 June 2018. <http://www.worldometers.info/world-population/netherlands-population/>. Last accessed 02 July 2018.
- [186] Worldometers, 09 June 2018. <http://www.worldometers.info/world-population/romania-population/>. Last accessed 02 July 2018.
- [187] Worldometers, 09 June 2018. <http://www.worldometers.info/world-population/italy-population/>. Last accessed 02 July 2018.
- [188] Worldometers, 09 June 2018. <http://www.worldometers.info/world-population/spain-population/>. Last accessed 02 July 2018.
- [189] W Xu, T Chen, X Dong, M Kong, X Lv, and L Li. Outbreak detection and evaluation of a school-based influenza-like-illness syndromic surveillance in Tianjin, China. *PLoS One*, 12(9):1–10, 2017.
- [190] J Yang, W Lu, and A Waibel. Skin-color modeling and adaptation. In *R Chin and T C Pong (eds) Computer Vision — ACCV’98 Asian Conference on Computer Vision 1998, Lecture Notes in Computer Science, vol 1352*, pages 687–694, Springer, Berlin, Heidelberg.
- [191] H Yasuda and K Suzuki. Measures against transmission of pandemic H1N1 influenza in Japan in 2009: simulation model. *Eurosurveillance*, 14(44):672, 2009.
- [192] C M Yuan, S Love, and M Wilson. Syndromic surveillance at hospital emergency departments—southeastern Virginia. *Morbidity and Mortality Weekly Report*, 53:56–58, 2004.
- [193] M Yuan and Y Lin. Model selection and estimation in regression with grouped variables. *Journal of the Royal Statistical Society, Series B (Statistical Methodology)*, 68(1):49–67, 2006.

- 
- [194] A Zuur, E N Ieno, N Walker, A A Saveliev, and G M Smith. *Mixed Effects Models and Extensions in Ecology with R*. Springer Science & Business Media. New York, 2009.