Spatial Modelling of Alcohol-Related Deaths and Hospitalisations in Scotland

Emma Robinson

A Dissertation Submitted to the
University of Glasgow
for the degree of
Master of Science

Department of Statistics

May 2012

Abstract

Alcohol misuse in Scotland is a major issue which is extremely detrimental to the health of the population and the economy (Donnelley (2008)). This thesis aims to explore the extent of alcohol-related health risks in Scotland at a finer geographical scale than previous research. A major objective of this research is to geographically map alcohol-related health risks in Scotland for males and females separately.

The Scottish data zone geographical areas are used in this study. These areas split Scotland into 6505 small sections each with a population of approximately 500-1000 people where this is reasonable. Details of all alcohol-related deaths and hospitalistions in Scotland during years 2002 to 2006 inclusive recorded at the data zone level are available. Information regarding area deprivation and at-risk population structure at the data zone level has also been obtained. Indirect age and sex standardisation is used to work out how many cases are expected to arise in each data zone.

Firstly, the standardised incidence ratio is explored as an estimate of the relative alcohol-related health risk in each data zone in Scotland. This is calculated separately for the combined male and female data, the male-only data and the female-only data. The results are mapped and discussed for each.

Further sections go on to use spatial Bayesian hierarchical modelling techniques to estimate the relative alcohol-related health risk in each data zone in Scotland. Again these methods are considered separately for the combined male and female data, for the male-only data and for the female-only data.

The basis for the models considered is the Besag, York and Mollié model (Besag et al. (1991)). The models explore both uncorrelated and correlated heterogeneity random effects. The correlated heterogeneity effects are fitted by means of the conditional autoregressive (CAR) prior. Fixed effects for area deprivation are also considered.

A further chapter explores a possible link between the location of single-malt whisky distilleries and alcohol-related health risks. This is done by incorporating the minimum Euclidian distance from the centroid of each data zone to a distillery into the Bayesian models already fitted to the combined male and female data.

The final chapter gives a discussion of the project limitations, difficulties and possibilities for future research.

Acknowledgements

There are several people I would like to thank in relation to this research.

Firstly, I would like to thank both my internal and external supervisors, Professor Mike Titterington and Professor Alastair Leyland. Without their help and guidance this would not have been possible. Their advice and study materials have been invaluable.

I would also like to thank NHS Scotland for funding this research and providing such useful medical data to study.

Some of my analysis would not have been possible without the Scottish postcode centroids, which were kindly supplied to me by Kate Trafford a GIS Analyst for the Rural and Environment Research and Analysis Directorate, part of the Scottish Government. I would like to take this opportunity to thank her for her help.

I am extremely grateful for the OpenBUGS assistance offered by Steve Miller at the MRC Biostatistics Unit in Cambridge. Without his help and investigations, which I believe led to OpenBUGS developments, I would not have been able to monitor vital properties of my models.

Last but definitely not least, I would like to thank my parents Colin and Joyce Robinson and my boyfriend Richard Briggs for encouraging me to start and finish this project. Without Richard's help I could not have done it. I truly appreciate all the support, both practical and emotional, he has provided; especially during the last 10 months while I have been commuting, working, studying for actuarial exams and writing my thesis simultaneously. He made these months bearable - our next holiday is on me!

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Alcohol misuse in Scotland is a major issue which is extremely detrimental to the health of the population and the economy (Donnelley (2008)). Studies which attempt to increase public understanding in this area are crucial in the fight to solve, or at least to reduce, this problem.

The analysis of public health data at a small geographic scale has become possible due to the recent availability of local geographically labelled health and population data. Such research has also been greatly encouraged by improvements in the fields of computing and geographic information systems. The results from studies which use small areas are more interpretable, less susceptible to ecological bias and capable of exposing highly localised effects, such as pockets of extreme deprivation. Conversely, small-scale studies often need more complicated and sophisticated statistical techniques because the data are often sparse due to low populations in each area.

1.1 Alcohol-Related Mortality in Scotland

Alcohol-related mortality is a major public health concern in Scotland with large increases in recent years (McLoone (2003)). There are marked geographical differences in such deaths, and the patterning is known to be related to social deprivation (Leyland et al. (2007)). Some evidence of spatial

clustering has been found for relatively large areas (census tracts - mean population 35,000 (Emslie & Mitchell (2009))). This paper will explore the spatial clustering of alcohol-related mortality in Scotland on a smaller scale.

Scottish alcohol mortality data is available for the years 2002-2006 at the level of data zone, a small area with mean population 780. The small area scale of this analysis should improve our understanding of the spatial concentration of such deaths. The relationship at data zone level between such deaths and the Scottish Index of Multiple Deprivation will also be investigated.

The poor effects of alcohol on Scotland's health have been known for many years. Several historic papers mention such problems, including Glaister (1886), which prophesied that an increase in cholera cases would arise due to "the festivities of the New Year season" and later mentions the effects of "holiday-drinking" on health.

Although Scotland's poor health record has been extensively studied, it cannot yet be explained. Scotland has notably worse health than the rest of Britain and has one of the lowest life expectancies in Western Europe for both men and women (Research Unit in Health, Behaviour & Change (2007)). Scotland has comparatively high mortality rates in most age groups for causes including lung cancer, strokes, accidents, suicide and alcohol-related mortality compared to England and Wales. Many mortality rates are known to be related to deprivation. However, Scotland's higher mortality rates do not seem to be completely explained by its higher rates of socio-economic deprivation. This is known as the 'Scottish Effect' and it is not well understood (Research Unit in Health, Behaviour & Change (2007)).

1.2 Disease Mapping

Geographic monitoring of disease is fundamental to understanding spatial patterns that can help to identify differences in disease prominence among

different regions or communities.

Mapping disease incidence data is now established as a primary tool in the analysis of regional public health data and there has been considerable development in this area in recent years due to an increase in computer capabilities.

Disease maps can be useful in many areas including public policy health care and ecological studies. They allow the analyst to identify areas which have unusually low or high values, highlight areas where cases seem to cluster together or comment on any evident patterns in disease distribution.

The data which can be used ranges from individual cases of disease with associated location to counts of disease cases within ceratin areas. The type of data available greatly affects the path future analysis will take and what statistical tools can be used. Various disease mapping methods are discussed in Chapter 3.

The purpose of disease mapping studies is often to produce smoothed maps of the risk of disease across the study region.

1.3 Objectives/Aims

This thesis aims to model alcohol-related health risks in Scotland spatially on a finer geographical scale than previous studies such as (Emslie & Mitchell (2009)). This will be done using Scottish data zone level of geography. It is hoped that mapping the relative risks of alcohol-related mortality at a finer resolution will increase understanding of the distribution of alcohol-related deaths across Scotland.

Previous research by Emslie & Mitchell (2009) investigated whether the kind of social environment which tends to produce higher or lower rates of alcohol-related mortality is the same for both men and women across Scotland. The results of this study showed that, as was expected, alcohol-related mortality rates for men substantially exceeded those for women, and that

there was significant spatial variation in the rates for both sexes. However, they found little spatial variation between male and female rates; in areas where men had high rates women also tended to have relatively high rates. This thesis hopes to examine the differences between alcohol-related mortality rates between men and women on a smaller area scale. This will be done by creating disease maps of alcohol-related risk for the combined population, males and females separately. It is expected that the risk pattern will be similar in each, but by looking at each group separately it allows any potential differences to be examined and if the chosen models for each have a similar structure, it adds confidence to the model results.

It is also of interest to investigate whether or not single malt whisky distilleries affect the risk of alcohol-related mortality. It is proposed to fit further models to our mortality data to see if the proximity of a distillery to the data zones explains some of the variation in alcohol-related mortality risk.

Bayesian hierarchical models will be used to fit the relative risk models and the program OpenBUGS [1] will be used to implement them.

[1] http://www.openbugs.info/w/

Chapter 2

Data

2.1 Data Source and Descriptions

The types of data that arise in disease mapping exercises can vary from the location of each disease case to counts of disease cases within small areas. It is necessary to use information about the underlying population at risk when trying to interpret any patterns that arise.

This section describes all data that was used in the project and where it was obtained.

2.1.1 Scottish Data Zones Data

Scottish data zones are the geographical areas used in this study. The data zone geography covers the whole of Scotland and splits it into 6505 areas. Due to the large number of these zones, which even split relatively small villages, it is not practical to give each a meaningful name. Instead each data zone is assigned an individual code, for example S01003313. Each data zone was created by combining groups of Census output areas as at 2001; these zones nest completely within Intermediate Geographies, which in turn nest entirely within local authority boundaries, as illustrated by Figure 2.1

obtained from the Scottish Government website [1]. Where possible each data zone has a household population of between 500 and 1000, groups together output areas with similar social attributes and respects physical boundaries such as rivers and lochs. More detailed information about the creation of the Scottish data zones can be found in the report by Flowerdew et al. (2004).

The crucial feature of the data zones is that they are considerably smaller than previous areas for which health statistics are routinely available, such as postcode sector or ward, but are large enough to protect patient confidentiality adequately. A further positive aspect of their small size is that they are more effective at identifying small areas with particular social attributes, such as pockets of extreme deprivation.

Various types of geographic information about the data zones were obtained from the Scottish Neighbourhood Statistics (SNS) website [2] including the physical boundary and centroid of each area. The boundary file allows a data zone map of Scotland to be created using geographic information system (GIS) software such as ArcGIS [3] and is used by WinBUGS/OpenBUGS to create the required adjacency matrix (discussed later in Chapter 3). A look-up table was also obtained from the SNS website which identifies the Intermediate Geography and local authority in which each data zone lies.

2.1.2 Death and Hospitalisation Data

Although this project is concerned with mortality, due to the small population size of each data zone, it has been decided to look at both alcohol-related deaths and hospitalisations due to alcohol. The conditions which are considered to be related to alcohol consumption are set out by the General

[1] http://www.scotland.gov.uk/Publications/2005/02/20697/52626 (accessed on 02/11/09)

[2] http://www.sns.gov.uk/Downloads/DownloadGeography.aspx (accessed on 10/10/09)

[3] http://www.esri.com/software/arcgis/index.html

Figure 2.1: Scottish Geography Relationships (obtained from Scottish Government website as referenced above)

Register Office for Scotland [4] and were agreed by the Office for National Statistics in 2006. Table 2.1 gives the causes of death that are considered to

[4]http://www.gro-scotland.gov.uk/statistics/deaths/alcohol-related-deaths/alcohol-related-deaths-the-coverage-of-the-statistics.html (accessed on 13/12/09)

be related to alcohol consumption during years 2000 to 2007 along with the corresponding code from the International Classification of Diseases Tenth Revision (ICD-10). It should be noted that some deaths which may be thought of as alcohol-related by many are not covered by this definition. These include deaths caused by road accidents, suicide, violence, falls or fires which occur under the influence of alcohol. Medical problems which are considered "partly attributable to alcohol" are also not included in the definition, and these include certain forms of cancer.

The data used was provided by the Information Services Division of NHS Scotland [5] and consists of all alcohol-related deaths and first alcohol-related hospitalisations in Scotland during the years 2002 to 2006. First alcohol-related hospitalisation means that the patient has never been admitted due to alcohol before or that they have not been admitted due to alcohol in the last ten years. The reason for including only patients who have not been admitted in the previous ten years is that it helps to avoid multiple counting, such as recording 10 events when one person is admitted with the same problem 10 times in a year. It should also be noted that deaths were only recorded if the patient had not been admitted to hospital due to alcohol in the last 10 years for similar reasons. This means that it is not possible to count the same individual as both a death and a hospitalisation in this study. This should lead to more accurate estimates of the alcohol-related health risk in each area.

Each data entry gives a code to identify whether it represents a death or a first hospital admission, the date of event and some patient information. The patient information consists of their sex, age and data zone of residence at time of admission. An age group indicator has been created to match the age groups used in the population data discussed below. Since only the data zone of residence is given as opposed to an exact point location or address, subsequent analysis is based upon tract-counts within each data zone.

[5] http://www.isdscotland.org/isd/CCC_FirstPage

ICD-10 Code	Description
F10	Mental and behavioural disorders due to use of alcohol
G31.2	Degeneration of nervous system due to alcohol
G62.1	Alcoholic polyneuropathy
I42.6	Alcoholic cardiomyopathy
K29.2	Alcoholic gastritis
K70	Alcoholic liver disease
K73	Chronic hepatitis, not elsewhere classified
K74.0	Hepatic fibrosis
K74.1	Hepatic sclerosis
K74.2	Hepatic fibrosis with hepatitic sclerosis
K74.6	Other and unspecified cirrhosis of liver
K86.0	Alcohol induced chronic pancreatitis
X45	Accidental poisoning by and exposure to alcohol
X65	Intentional self-poisoning by and exposure to alcohol
Y15	Poisoning by and exposure to alcohol, undetermined intent

Table 2.1: Alcohol-Related Conditions During Years 2000 to 2007

2.1.3 Population Data

Scottish population data has been obtained from the General Register Office for Scotland website [6] separately for the years 2002 to 2006. For each year the population is broken down by data zone, age group and sex. The age groups used are zero to four years, five to nine years, 10 to 15 years, 16 to 19 years, 20 to 24 years, 25 to 29 years, then continuing in 5-year bands until 90 plus years. Note that the age bands 10 to 15 years and 16 to 19 years do not cover five years as most bands do. This is so that it is possible to split the data into children (less than 16 years), working age (16 to 59/64 years) and pensionable age (60/65 years or more) if it proves

[6] http://www.gro-scotland.gov.uk/statistics/publications-and-data/population-estimates/special-area/sape/index.html (accessed on 30/11/09)

desirable. Further information about how these population estimates are calculated can be found on the General Register Office for Scotland website referenced earlier in this section.

2.1.4 Possible Risk Factors

Deprivation

Previous studies such as Leyland et al. (2007) have shown that alcohol-related mortality rates tend to be higher in more deprived areas and it is of interest to investigate this relationship. The measure of deprivation used in this study is the Scottish Index of Multiple Deprivation (SIMD) 2004. This index aims to locate small areas of concentrated multiple deprivation across Scotland as fairly as possible. It is based on the data zone geography and combines 31 separate deprivation indicators including current income, employment, health, education, housing and geographical access. The index is based on methodology developed by Oxford University and also implements changes as recommended in the report by Bailey et al. (2003). Further information on the SIMD can be found on the Scottish Government website [7].

Using the SIMD estimates for 2004 seems reasonable as this is in the middle of the study period, 2002 to 2006, and most of the data used to calculate the estimates actually represents 2002.

The SIMD 2004 values were used to create a categorical deprivation variable ranging from one to ten, 1 representing the most deprived 10% of data zones and 10 representing the least deprived 10% of data zones.

[7] http://www.scotland.gov.uk/Publications/2005/01/20458/49127 (accessed on 02/08/10)

Whisky Distilleries

This study focuses on single malt whisky distilleries in Scotland. There are many sources online giving conflicting lists of Scottish distilleries and it has been decided to use those listed in Jackson (1999). This book was published in 2001, one year before the study period begins. In this book Jackson lists "every Scottish malt distillery that has ever witnessed its product in a bottle". Some of these distilleries have long been closed but are included in the text because the whisky can still be found. It has been decided to omit distilleries that closed over ten years before the study period, i.e. all distilleries which closed before 1992. The postcode of each distillery is given in Jackson (1999) and the centroid of each of these postcodes has been provided by the Scottish Government (although without permission to pass on or publish). The centroid of the postcode in which each distillery falls has been used as their approximate location.

The Euclidian distance between each data zone centroid and each distillery location was calculated and for each data zone the minimum distance to a distillery was recorded in meters.

2.2 Data Summaries

2.2.1 Death and Hospitalisation Data

There were 67742 alcohol-related events of interest in Scotland during 2002 to 2006, of which 65212 (96.3%) are first hospital admissions and 2530 (3.7%) are deaths. Given that significantly more hospitalisations have been observed than deaths, this study is de facto looking at alcohol-related hospitalisations. Any spatial patterns present amongst the deaths will be 'overshadowed' by patterns present in the hospitalisation data. However, it is expected that such patterns should be similar.

As expected there are significantly more male occurrences, with 69.4%

of first-alcohol-related admissions and 70.7% of alcohol-related deaths attributed to males. For males and females, however, a very similar proportion of cases corresponded to deaths with 3.8% for males and 3.6% for females.

There are 1409 data entries (2.1%) for which the data zone of residence has not been recorded. This may be due to unknown area of residence at time of admission or administrative errors. These events have been included when estimating the overall Scottish rates for males and females and for each age group, but obviously cannot be used when counting occurrences in each data zone. Since there is only a small percentage missing this should not affect the risk estimates much and there is no reason to believe that there is any systematic reason for the missing information.

Of the 6505 data zones across Scotland only 63 (fewer than 0.1%) experienced no alcohol-related deaths or hospitalisations during the study period. All of these zones have a deprivation score of 6 or more, i.e. are part of the least deprived half of data zones, and over half (32) had a deprivation score of 10.

Further, the highest number of alcohol related deaths and hospitalisations in a single zone over the period is 87. This occurred in data zone S01003313 an area of Parkhead West and Barrowfield in Glasgow's EastEnd which ranks in the top 10% of most deprived areas in Scotland. This supports the findings of previous studies such as Leyland et al. (2007) which suggest that high deprivation levels are linked to high alcohol-related mortality and that alcohol-related mortality is particularly high in the Glasgow area.

2.2.2 Age Groups

Table 2.2 gives the number of alcohol-related events in each age group during the period 2002 to 2006, as well as giving the percentage of events that each age group accounts for. This table shows that 0.2% of the events considered in this study correspond to children less than ten years of age. Although this is a small percentage it consists of 114 hospitalisations which

is more than one might expect for such young ages. It has been decided to include all age ranges in the later model-fitting since all age groups experience at least 38 events during the five years in question. The general pattern of the data appears to be an increase in alcohol-related deaths and hospitalisations in successive age groups, peaking at 45 to 49 years, followed by a general decline through to the highest age group of 90-plus years.

Age Group (years)	Frequency	Percent	Cumulative Percent
0 to 4	76	0.1	0.1
5 to 9	38	0.1	0.2
10 to 15	2640	3.9	4.1
16 to 19	4609	6.8	10.9
20 to 24	4806	7.1	18.0
25 to 29	3706	5.5	23.4
30 to 34	4351	6.4	29.9
35 to 39	5350	7.9	37.8
40 to 44	5984	8.8	46.6
45 to 49	6176	9.1	55.7
50 to 54	5763	8.5	64.2
55 to 59	5876	8.7	72.9
60 to 64	5595	8.3	81.1
65 to 69	4673	6.9	88.0
70 to 74	3625	5.4	93.4
75 to 79	2460	3.6	97.0
80 to 84	1366	2.0	99.0
85 to 89	483	0.7	99.8
90 +	165	0.2	100
Total	67742	100	

Table 2.2: Age Group Percentages

Table 2.3 breaks down the number of alcohol-related events in each age

group into male and female occurrences. This table shows that, as was expected, there are many more alcohol-related deaths and hospitalisations among men than among women. Males have a higher number of alcohol-related events in every age group apart 10 to 15 years. This may be because females tend to hit puberty earlier and may start adolescent drinking at an earlier age than males. The difference between the male and female counts increases in general until 60 to 64 years, with the exception of 20 to 24 years where there is a bulge, and then it begins to decrease in successive age groups.

Although these figures show the patterns that one would expect, it should be noted that this is a crude analysis which only looks at count data and takes no account of the size or distribution of the population at risk. For example, it takes no account of the fact that there tends to be a higher proportion of women in the older age groups.

2.2.3 Possible Risk Factors

Deprivation

In order to summarise the level and patterning of deprivation scores across Scotland various maps have been produced. A full map of Scotland showing the area deprivation score in each data zone is given in Figure 2.2 along with magnified areas of this map for Aberdeen (Figure 2.3), Ayrshire (2.4), the Dundee area (2.5), Edinburgh (2.6), Glasgow (2.7), the Inverness area (2.8) and Stirling (2.9).

From the full map of Scottish deprivation scores in Figure 2.2 it appears that areas tend to be more deprived towards the north and west of the country. An important and obvious observation is that deprivation levels appear to be extremely high in the Glasgow City area, even when compared to another large city such as Edinburgh. This appears to be especially true in the East of the City. On the whole, the most deprived areas with a score of 1 tend to be very small and densely populated. There is also some evidence

Age Group (years)	Male	Female	Total
0 to 4	45	31	76
5 to 9	31	7	38
10 to 15	1293	1347	2640
16 to 19	2990	1619	4609
20 to 24	3411	1395	4806
25 to 29	2623	1083	3706
30 to 34	3025	1326	4351
35 to 39	3581	1769	5350
40 to 44	3999	1985	5984
45 to 49	4244	1932	6176
50 to 54	4151	1612	5763
55 to 59	4254	1622	5876
60 to 64	4166	1429	5595
65 to 69	3490	1183	4673
70 to 74	2666	959	3625
75 to 79	1816	644	2460
80 to 84	923	443	1366
85 to 89	281	202	483
90 +	87	78	165
Total	47076	20666	67742

Table 2.3: Age and Sex Frequency Table

of cluster of high and low levels of deprivation.

Over 41% of the data zones in Scotland with the worst deprivation score of 1 fall within Glasgow City and these zones represent roughly 39% of the data zones in Glasgow City. On the other hand the local authority areas of Moray, Shetland Islands, Orkney Islands and Eilean Siar have no data zones with the most severe level of deprivation.

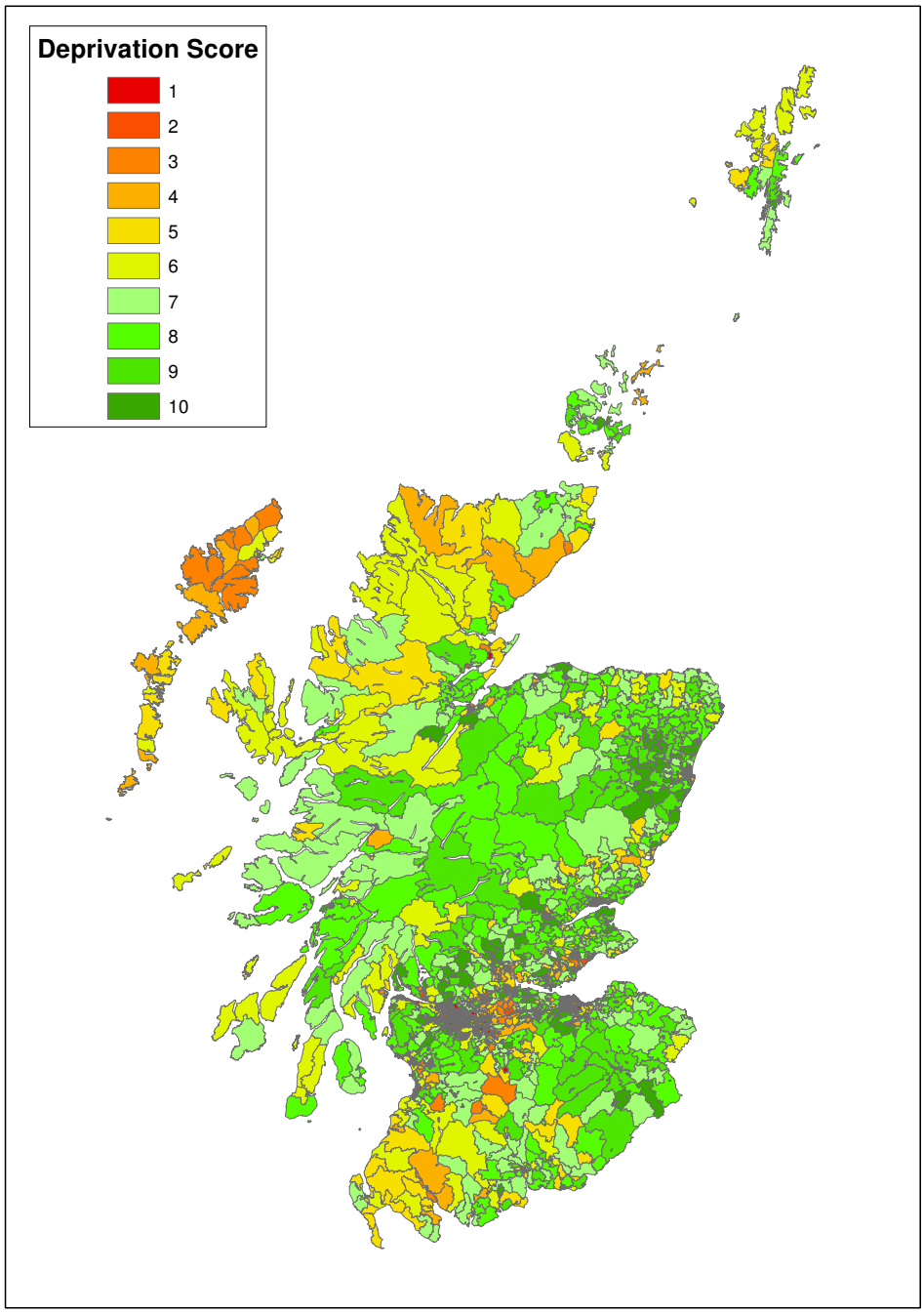


Figure 2.2: Map of Scottish Data Zone Deprivation Scores

Single Malt Whisky Distilleries

In total there are 98 single malt whisky distilleries in Scotland that meet our criteria. The minimum (approximate) distance from a data zone to a

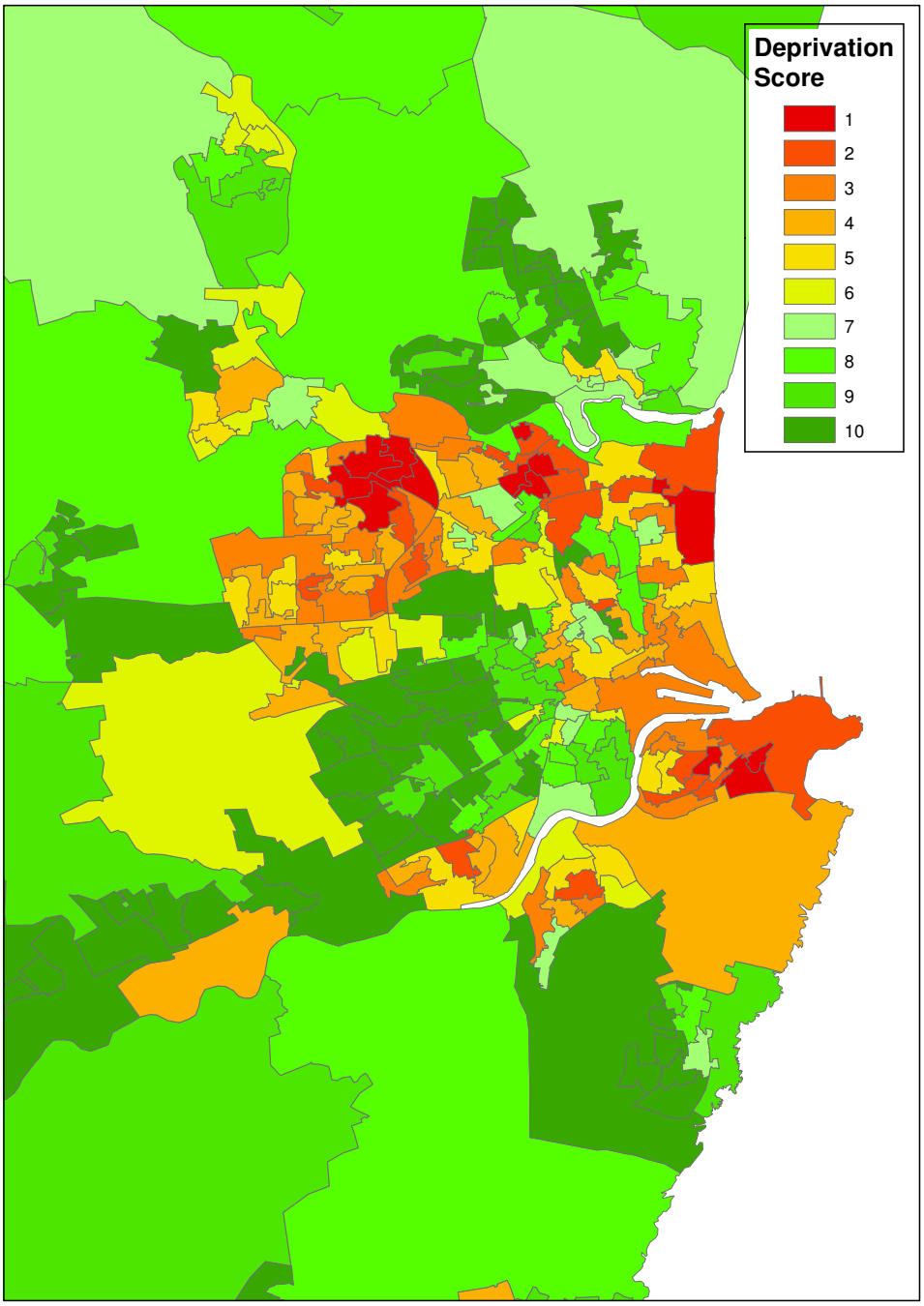


Figure 2.3: Map of Aberdeen Data Zone Deprivation Scores

single malt whisky distillery ranges from 0.0127 to 231.47 km and has a mean of 22.53 km.

A Scottish map of estimated minimum Euclidean distance between each data zone centroid and a single malt whisky distillery in meters is shown

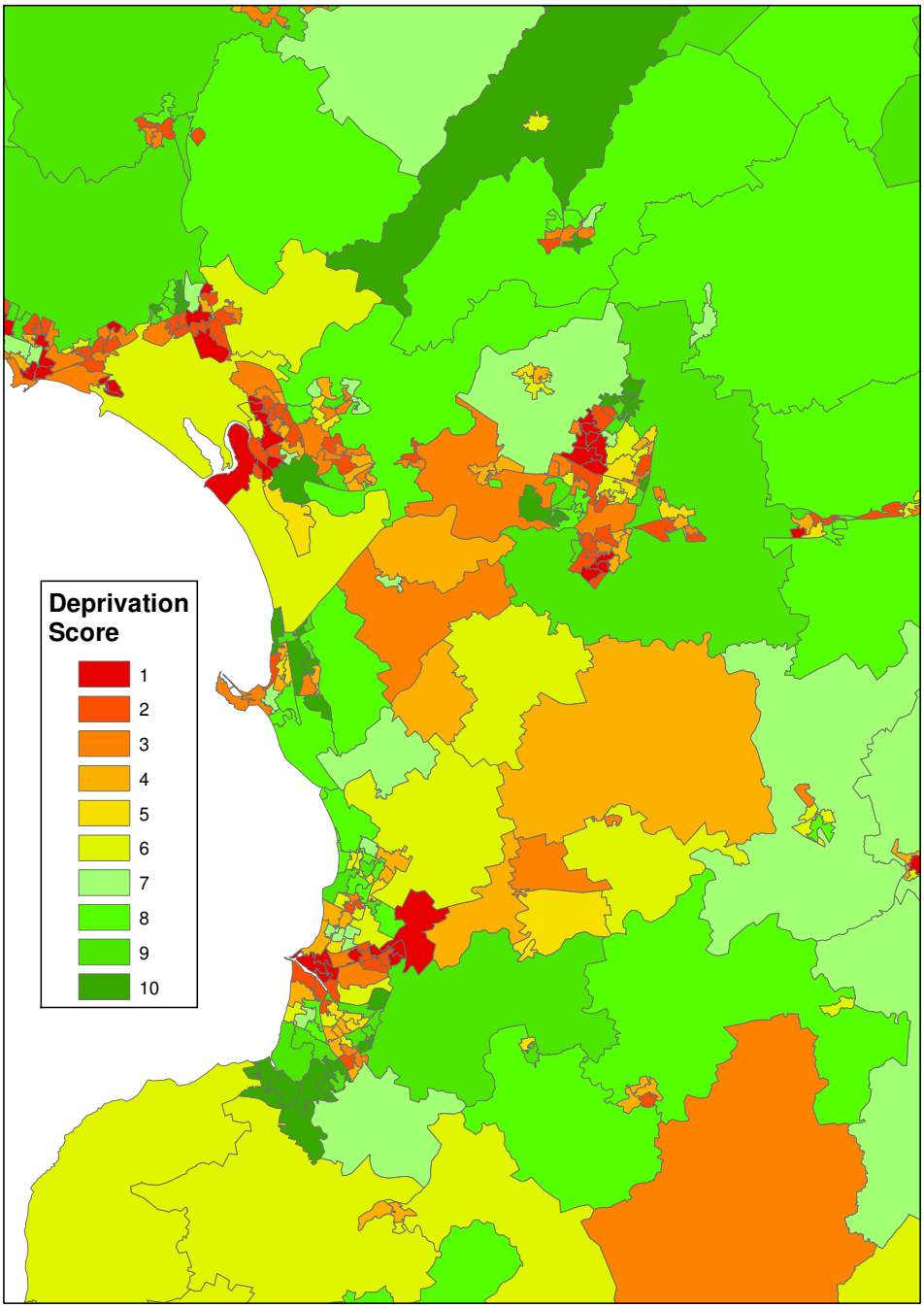


Figure 2.4: Map of Ayrshire Data Zone Deprivation Scores

in Figure 2.11; magnified areas of this map are shown for Aberdeen (Figure 2.11), Ayrshire (Figure 2.12), the Dundee area (Figure 2.13), Edinburgh (Figure 2.14), Glasgow (Figure 2.15), the Inverness area (Figure 2.16) and Stirling (Figure 2.17).

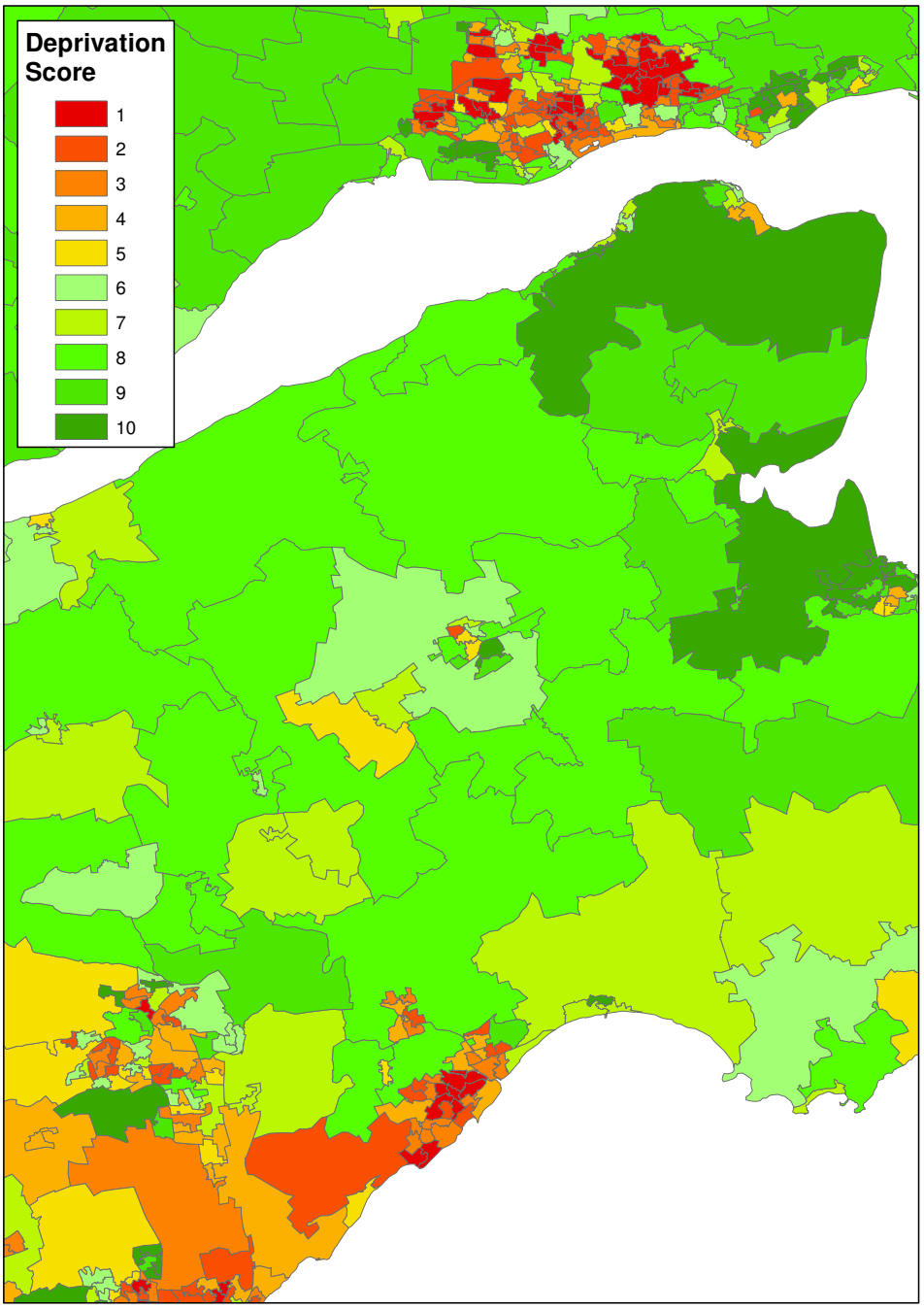


Figure 2.5: Map of Dundee & Fife Data Zone Deprivation Scores

These maps show no clear pattern or strong similarities to the deprivation maps discussed above. However, Glasgow City appears to be very close to a whisky distillery, and given previous research findings of very high alcoholism rates in the city, this may prove important.

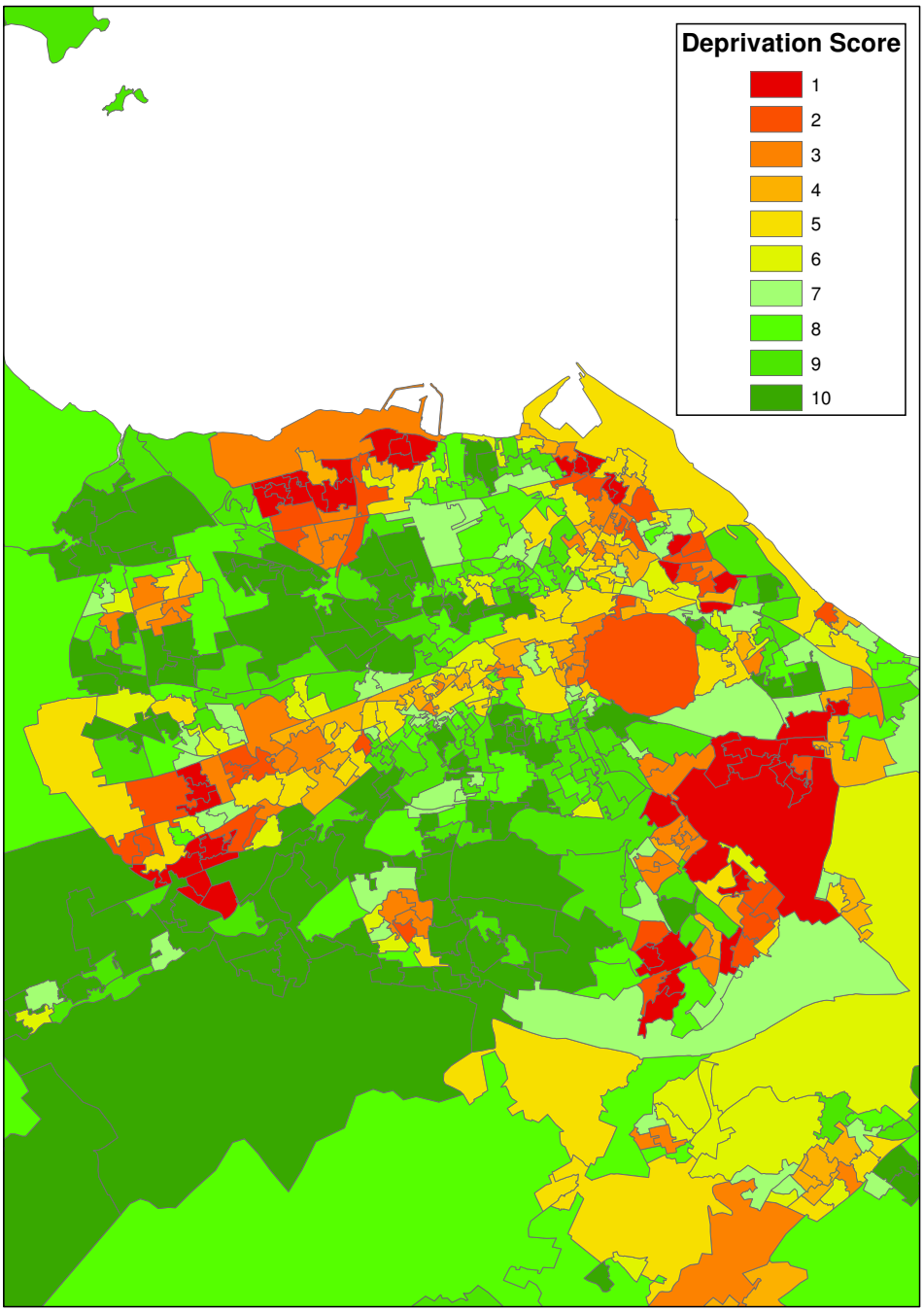


Figure 2.6: Map of Edinburgh Data Zone Deprivation Scores

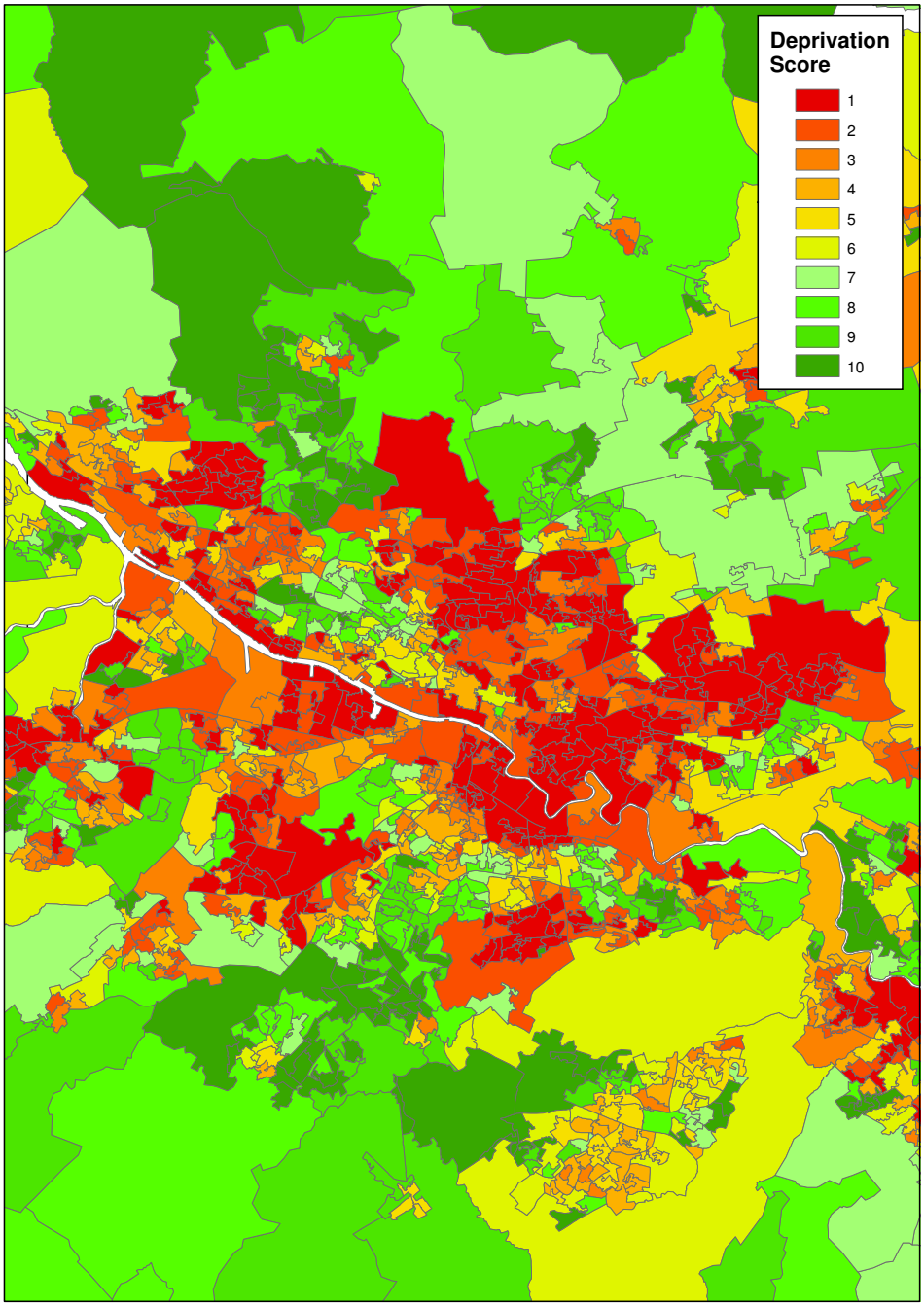


Figure 2.7: Map of Glasgow Data Zone Deprivation Scores

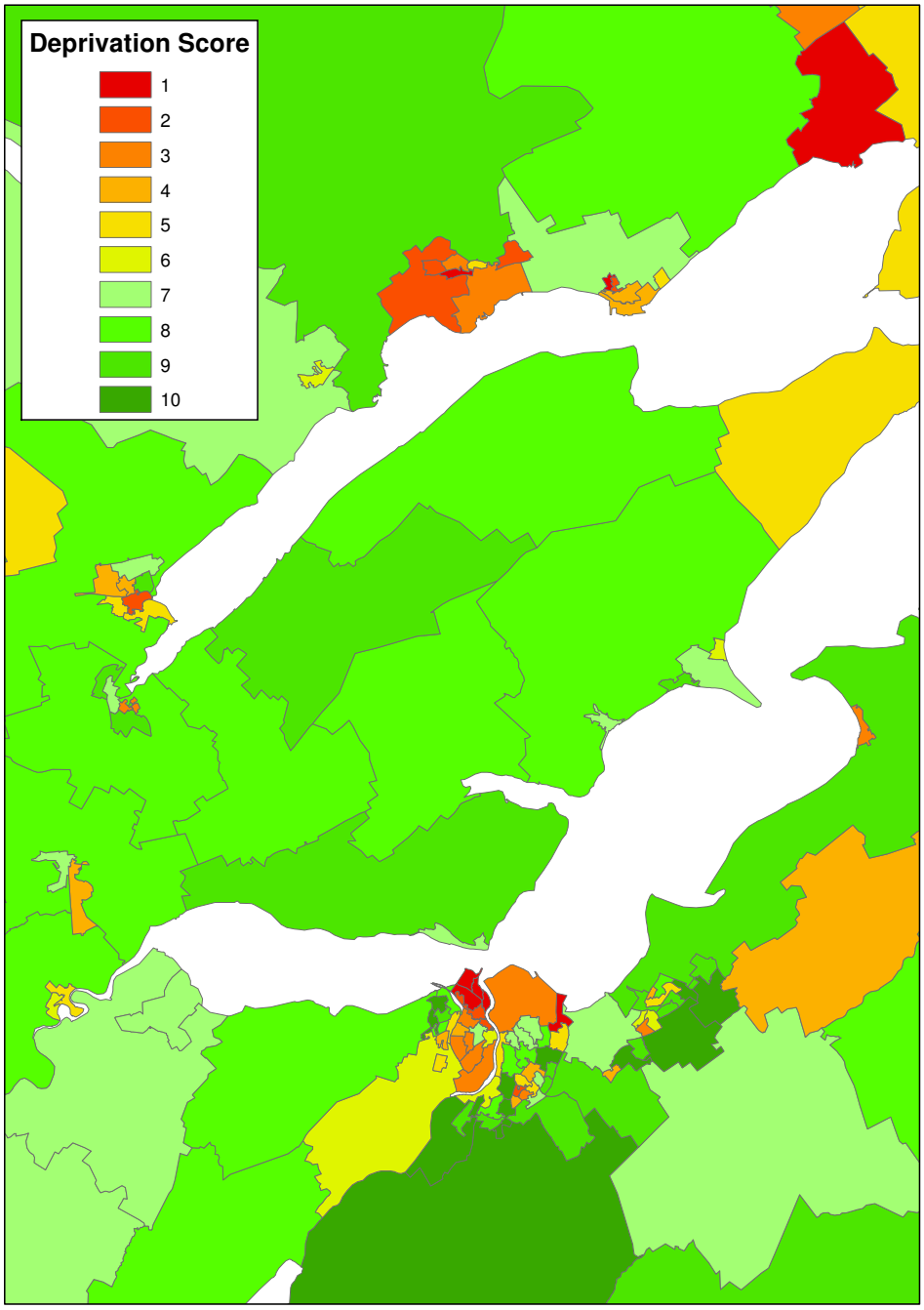


Figure 2.8: Map of Inverness & the Highlands Data Zone Deprivation Scores

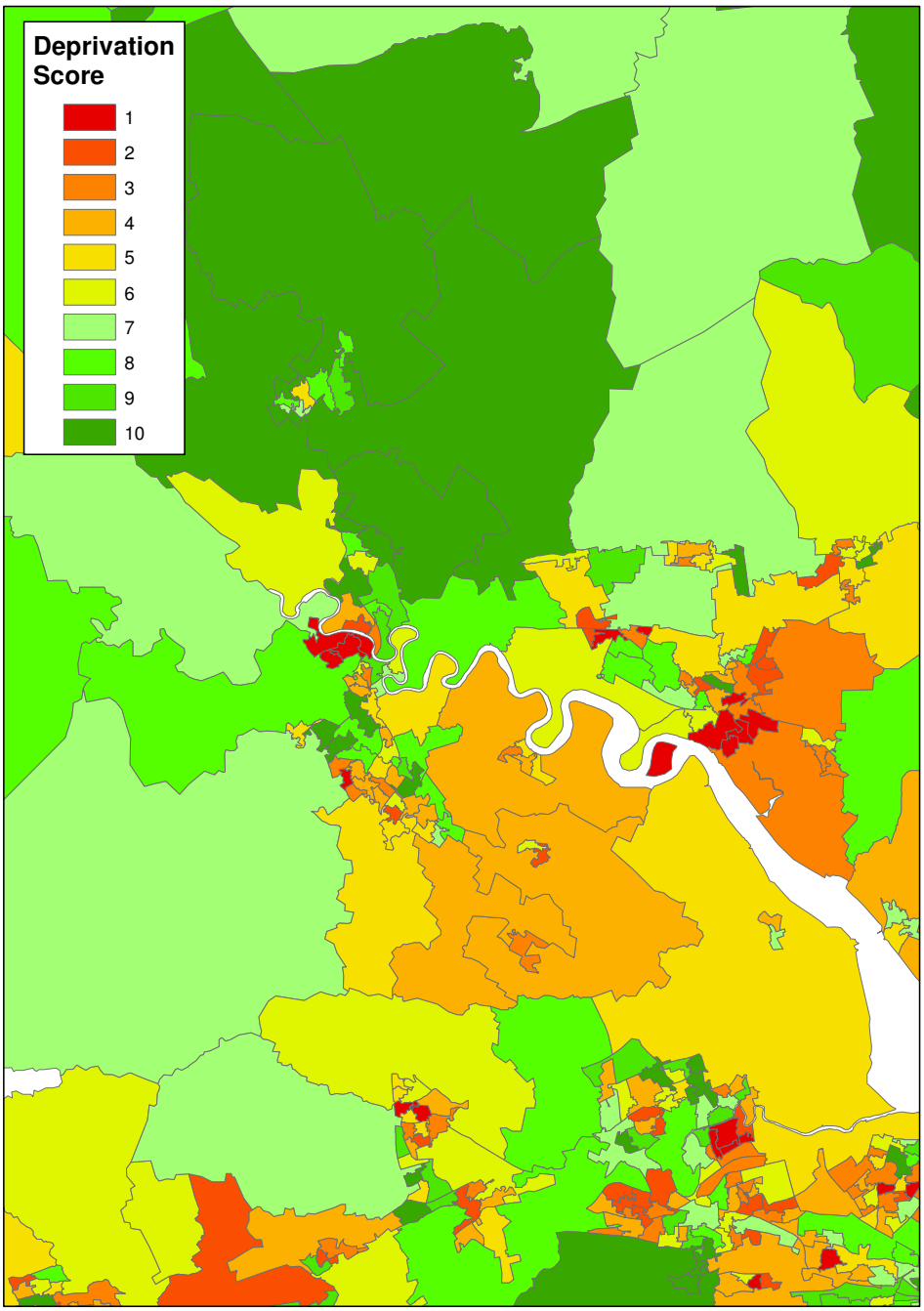


Figure 2.9: Map of Stirling Data Zone Deprivation Scores

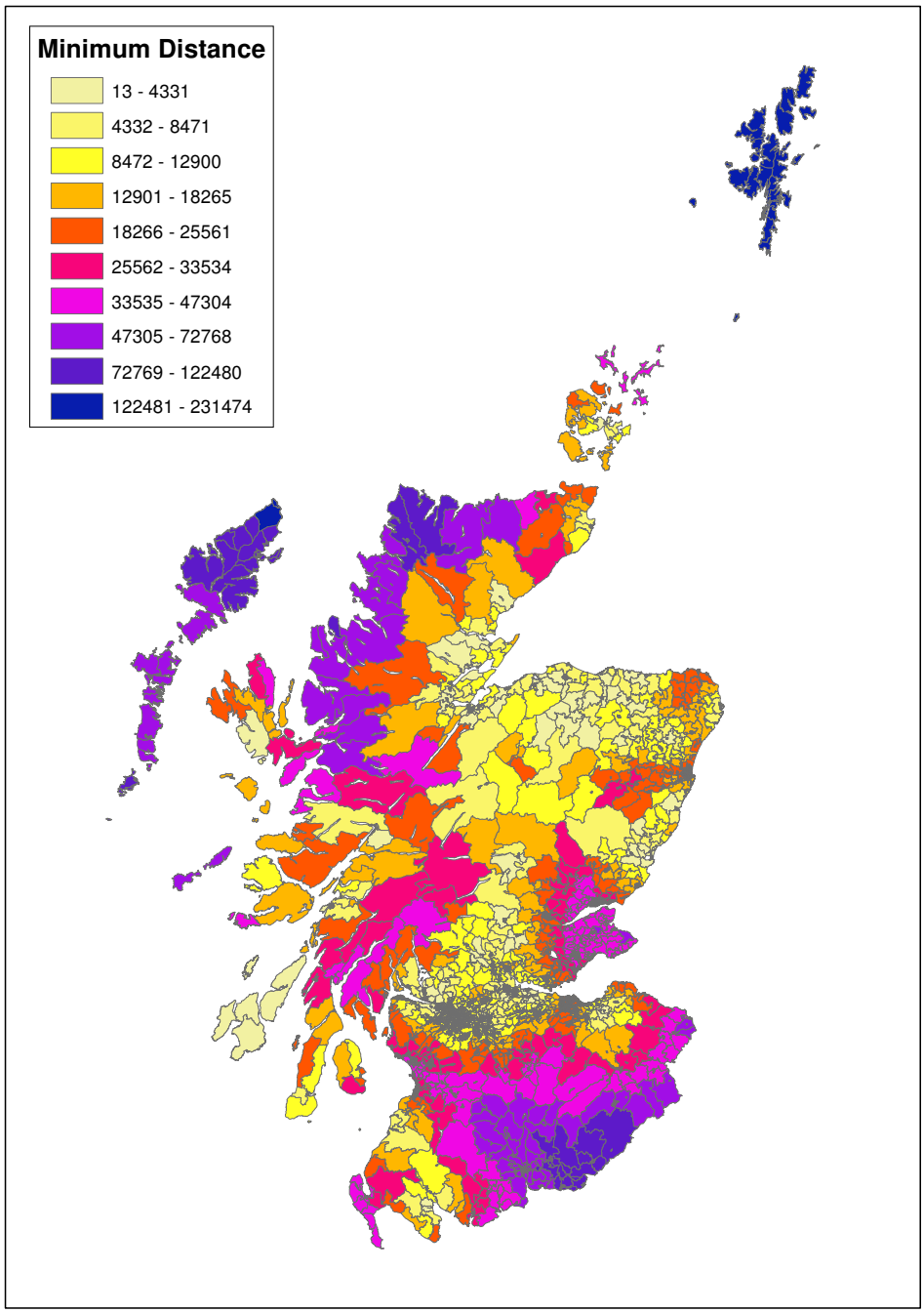


Figure 2.10: Scotland Map of Proximity to a Single Malt Whisky Distillery

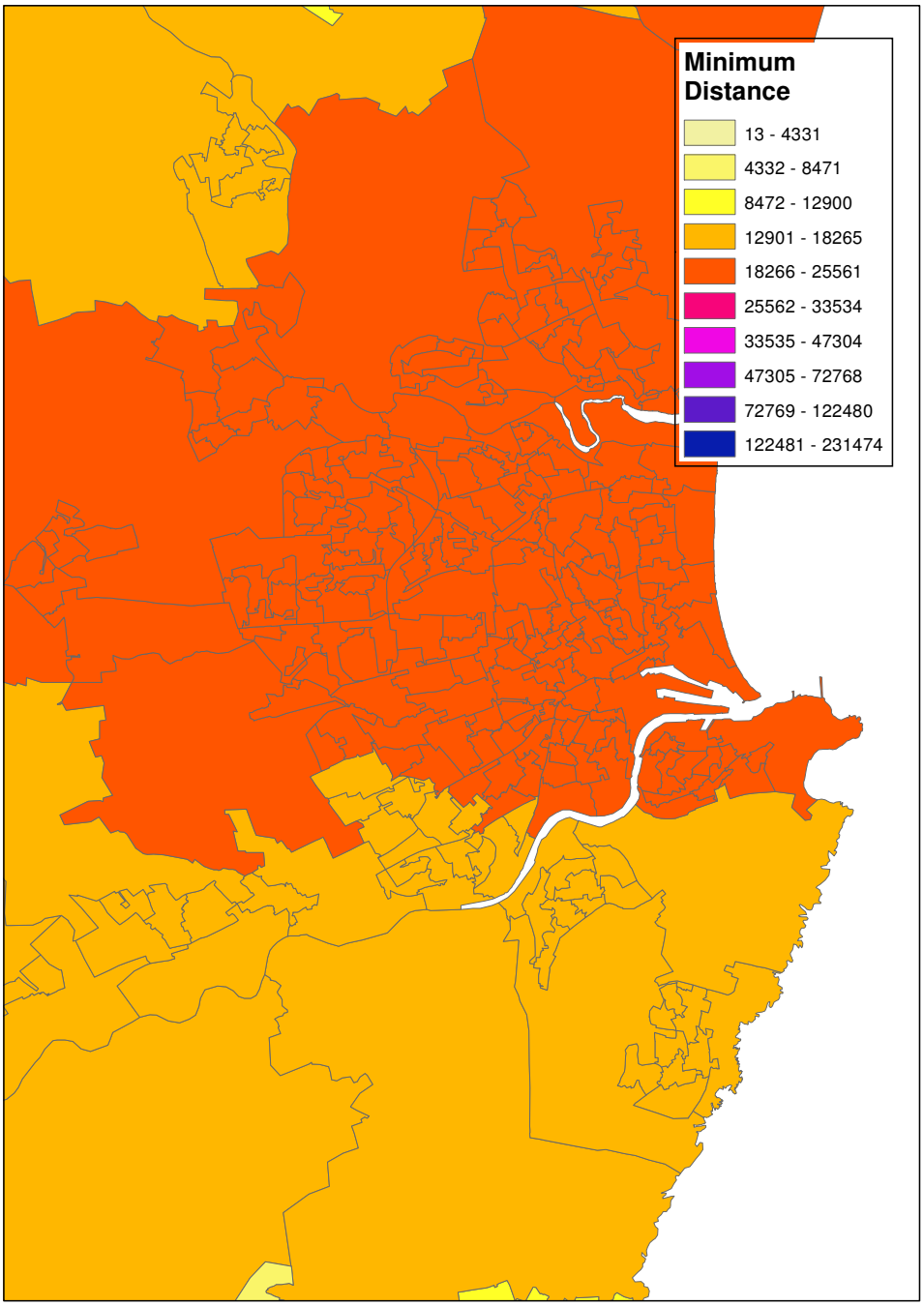


Figure 2.11: Aberdeen Map of Proximity to a Single Malt Whisky Distillery

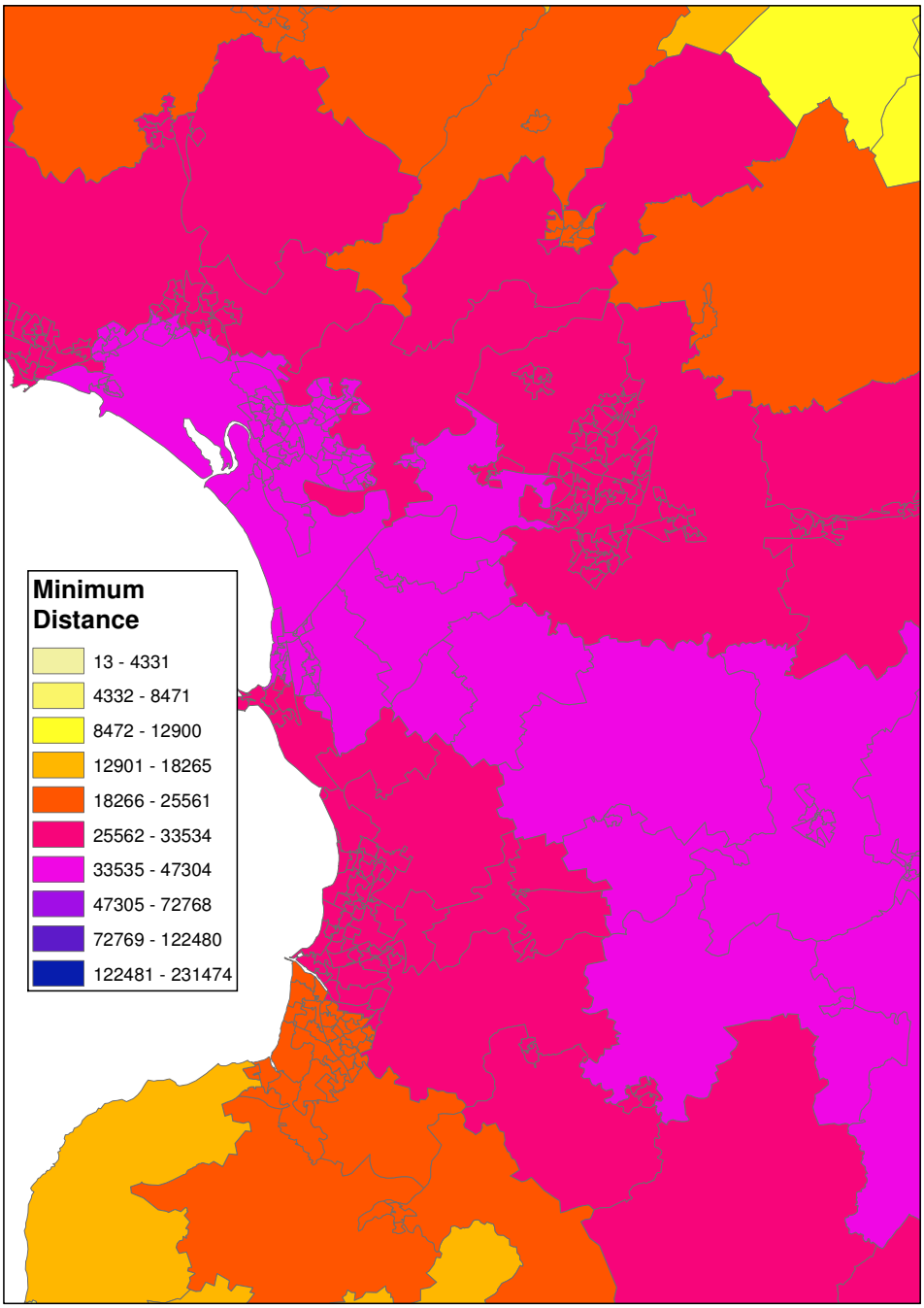


Figure 2.12: Ayrshire Map of Proximity to a Single Malt Whisky Distillery

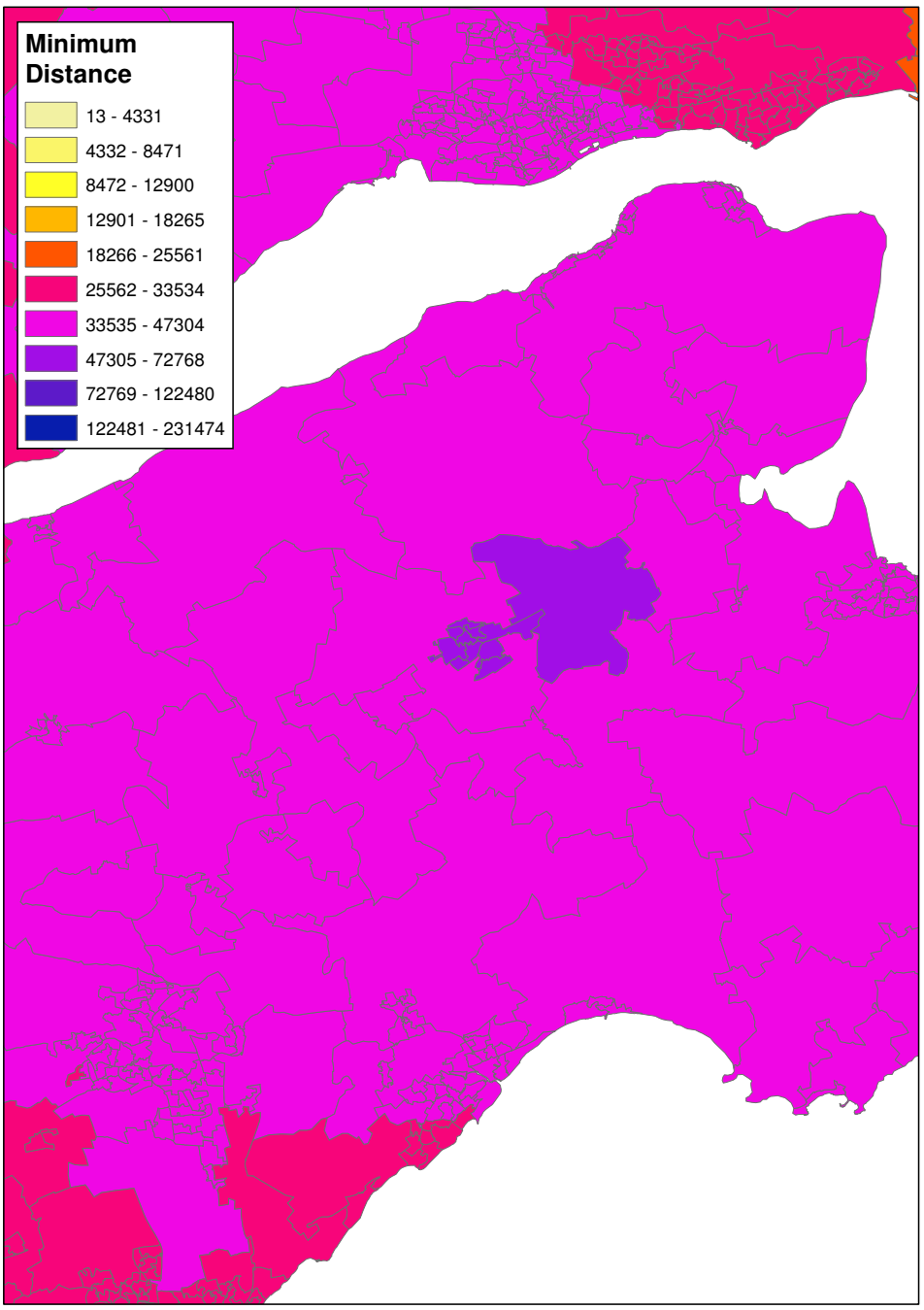


Figure 2.13: Dundee Area Map of Proximity to a Single Malt Whisky Distillery

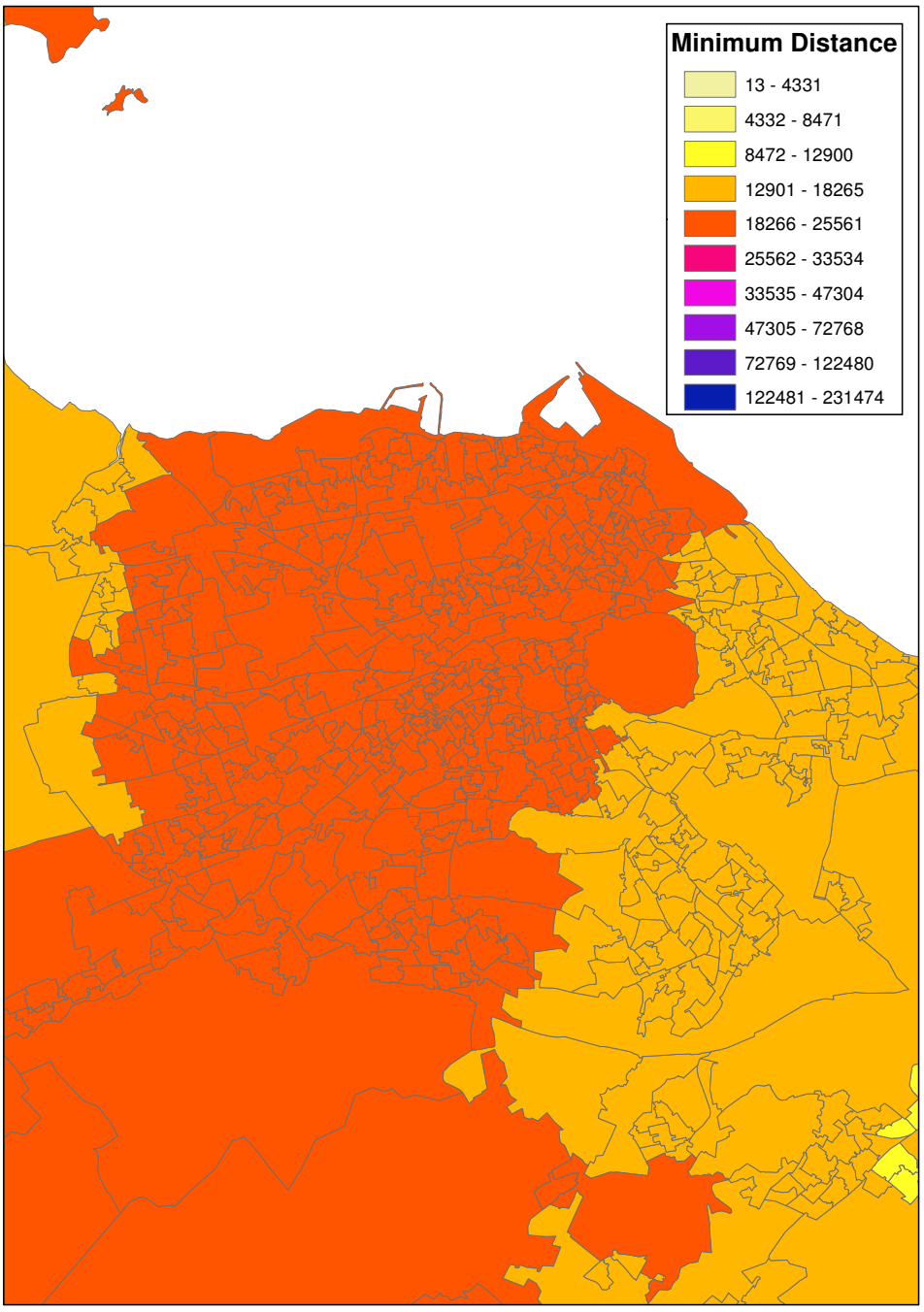


Figure 2.14: Edinburgh Map of Proximity to a Single Malt Whisky Distillery

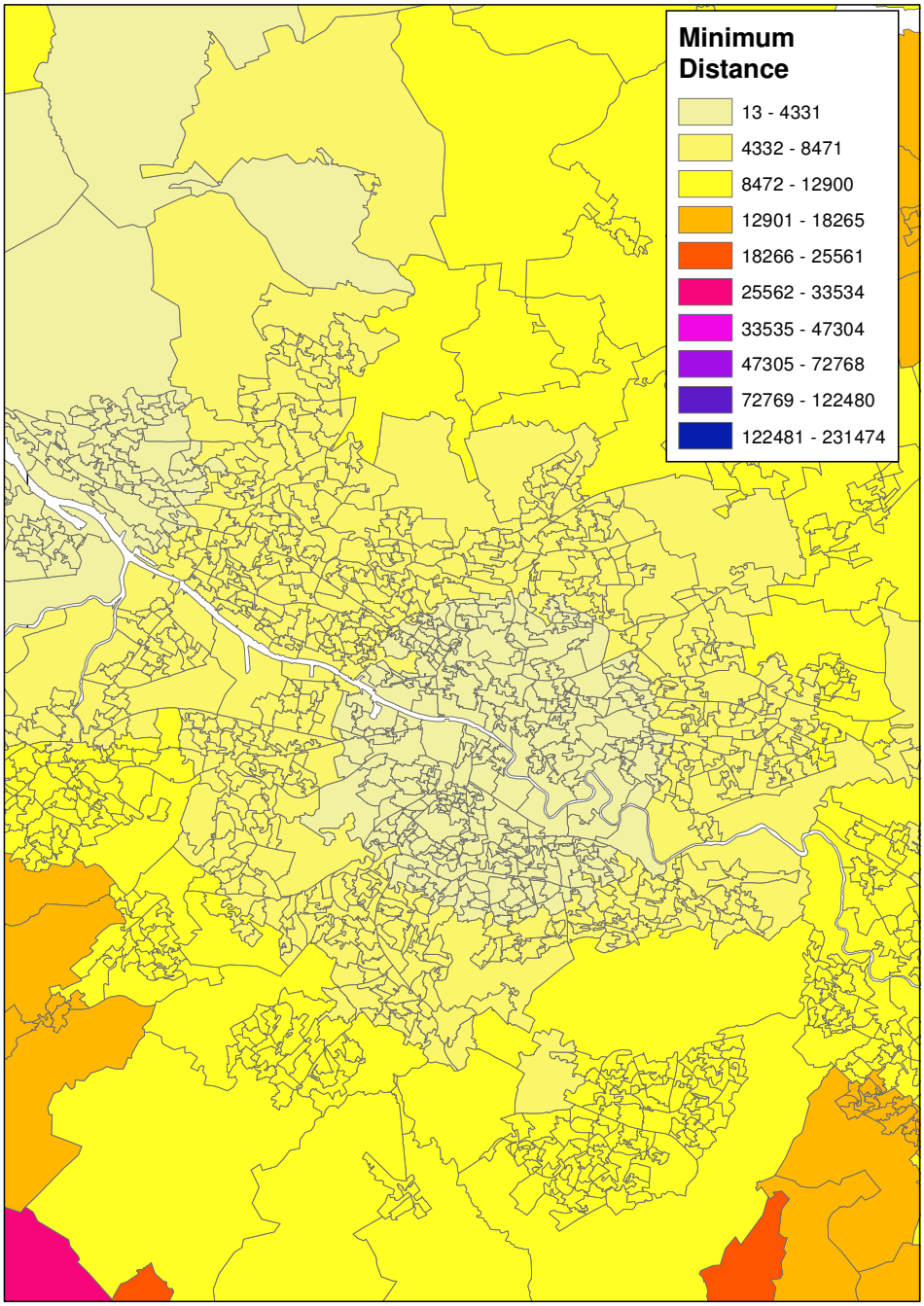


Figure 2.15: Glasgow Map of Proximity to a Single Malt Whisky Distillery

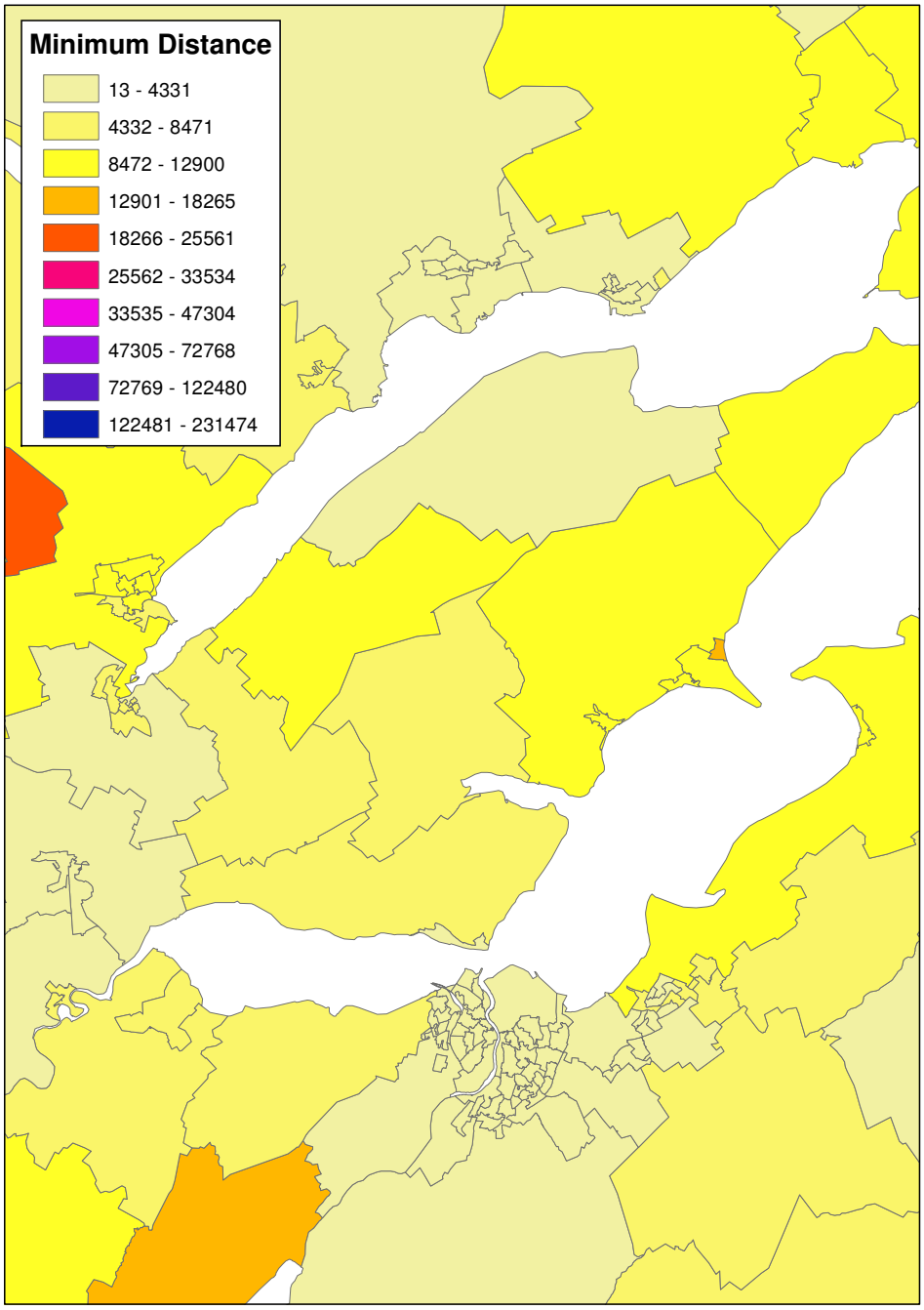


Figure 2.16: Inverness Area Map of Proximity to a Single Malt Whisky Distillery

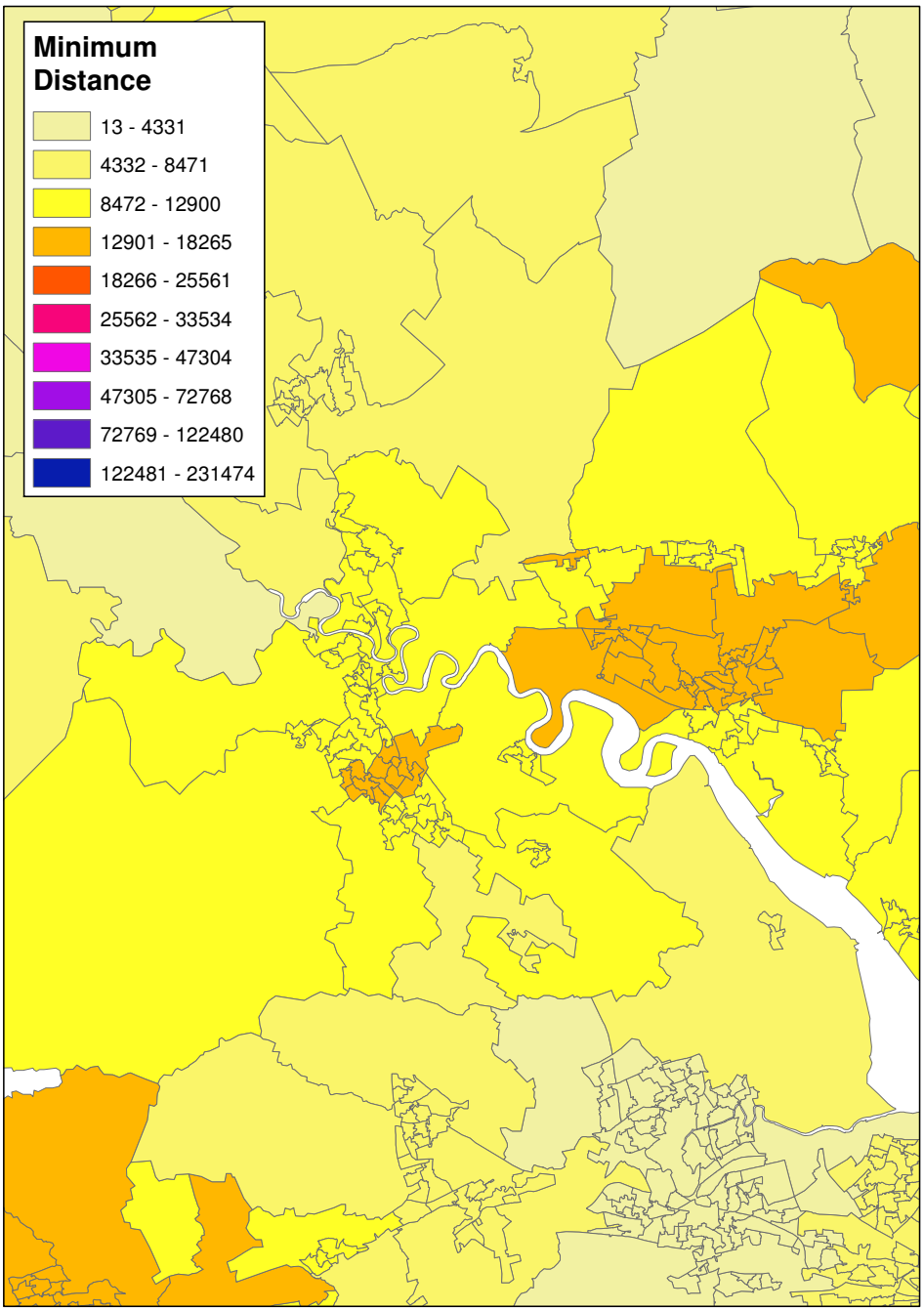


Figure 2.17: Stirling Map of Proximity to a Single Malt Whisky Distillery

Chapter 3

Review of Disease Mapping Methods

The study of the geographical distribution of disease is facilitated by the use of disease maps. Such maps have clear advantages over tables and have a number of uses, often considered regarding public policy, medical research and public health.

3.1 Introduction to Disease Mapping

Geographic monitoring of disease is paramount to understanding spatial patterns that identify differences in disease prominence between different regions or communities. There are two classes of disease maps: those showing maps of individual cases and those showing maps of aggregated counts or rates. The first requires the availability of individual addresses, the locations of which are then mapped. Often information at this level of accuracy is not publicly available due to privacy issues. Creating and analysing maps of disease rates/incidence is carried out extensively in modern public-health studies. In this chapter various mapping quantities and methods will be discussed, including the SMR and model-based risk estimates.

3.1.1 Age and Sex Standardisation

In order to assess whether or not the number of disease cases that occur in an area is high or low it is useful to work out first how many cases are expected to arise in that area given the at-risk population structure. It is then possible to compare the observed number of cases in area i (O_i) to the expected number of cases in area i (E_i).

To account for the population structure indirect age and sex standardisation is normally used. This first calculates an expected number of cases in age-and-sex stratum j in area i (E_{ij}), given by

$$E_{ij} = N_{ij}\frac{\sum_i O_{ij}}{\sum_i N_{ij}},$$

where N_{ij} represents the number of people in age-and-sex stratum j in area i and O_{ij} represents the observed number of cases among people in age-and-sex strata j in area i. The expected number of cases in area i is then calculated by summing the expected number of cases in that area in each age-and-sex stratum using the following method:

$$E_i = \sum_j E_{ij}.$$

3.1.2 Standardised Mortality and Incidence Ratios

The standardised mortality ratio (SMR) and standardised incidence ratio (SIR) are commonly used when creating disease maps as they give a simple and efficient estimate of the relative risk of disease-related risk in a given area. The SMR relates specifically to disease-related deaths and so gives an estimate of the risk of death by a given disease in an area, whereas the SIR relates to the number of new disease cases within an area.

Standardised Mortality Ratio

The SMR is calculated as the ratio of observed disease-related deaths within an area to the expected number of disease-related deaths in that area. Mathematically, the SMR for area i is given by

$$SMR_i = \frac{O_i}{E_i}. \tag{3.1}$$

If the observed number of disease deaths in an area is greater than the expected number, this results in an SMR greater than 1, which indicates that this area has a higher risk of death due to the disease in question relative to the study region as a whole.

Although it is an easy-to-understand and convenient risk estimate there are several drawbacks associated with the SMR and its use has been criticised by several researchers including Lawson et al. (2003*a*), Clayton & Kaldor (1987) and Tsai & Wen (1986). The measure is very sensitive to zero values of observed counts and expected counts close to zero. The following example illustrates the latter point. If we are looking at a fairly rare disease while using geographical areas with low populations, it is possible that the expected number of disease deaths in a given area is 0.5. If this area had an observed count of zero then it would result in a relative risk estimate of zero. However, if the observed count was 1 this would result in a relative risk estimate of 2. The fact that a single death has the ability to affect risk estimates so much is a very undesirable property. A further weakness of the SMR is that it does not have the capacity to incorporate important risk factors in the way that model-based risk estimates can.

Standardised Incidence Ratio

The standardised incidence ratio (SIR) is concerned with new disease cases rather than disease deaths and can be used to estimate the prevalence of disease in different areas. It is calculated in the same manner as the SMR,

but uses the expected number and observed number of new disease cases rather than the expected and observed number of disease deaths. It suffers from the same interpretation problems as the SMR. Since this thesis deals with first-hospitalisations and deaths for people who have not been admitted to hospital with alcohol-related illness for at least 10 years, it deals with new cases of alcohol-related disease. The SIR will therefore will be used in place of the SMR in the later analysis chapters.

3.1.3 Mapping Relative Risk Estimates

Relative risk estimates are usually represented using a cloropleth thematic map which provides an easy way to visualise how the values vary across the region and shows the level of variability within a region. This involves splitting the estimated risk values into different interval classes and assigning to each class a shade, colour or pattern which will be used to fill areas of that class on the map. Two common methods for specifying the classification intervals are the equal-interval method and the equal-representation method. The equal-interval method involves splitting the estimated relative risks into a fixed number of classes each of which represents an equal range of values but relates to differing numbers of areas. One pitfall of this method is that for a highly skewed or uneven relative risk distribution some classes will cover many more observations than others do. The map will therefore be dominated by these classes and will show little spatial variation when this may not be the case. For the equal-representation method percentiles of the estimated relative risk distribution, such as quartiles or quintiles, are used as cut-points for the class intervals, which results in each class representing an equal number of areas but having different ranges. Although this method ensures that each class is represented equally on the map it can also be deceptive because risk estimates with similar values may often be assigned to different classes causing some areas of the map to appear more heterogeneous than they really are (Davies (2005)).

Another common method for choosing categories when mapping values is the Jenks natural breaks classification [1]. This is the method used by the ArcGIS geographical mapping software which has been used to produce later maps in this study. The Jenks classes aim to reflect natural groupings inherent in the data. Category boundaries are chosen as the breakpoints that best group similar values together, maximising the difference between the classes.

Even when suitable interval classes have been created the colour or pattern scheme used to represent these classes on the map can affect how the values are interpreted. Some alternative graphical presentation methods are discussed in Marshall (1991).

The mapping issues discussed above should be considered when dealing with every relative risk estimate discussed in this thesis. This includes the SMR, SIR and model-based risk estimates.

3.2 Basic Disease Mapping Models

3.2.1 Likelihood Models

Using parametric modelling methods allows the relative risks to be estimated and mapped using maximum likelihood methods.

Poisson Model

The classical model adopted in many disease mapping studies assumes that the counts of disease cases follow Poisson distributions with different expectations for each area, as discussed by Lawson et al. (2003*b*). For area i the observed count of disease cases (O_i) is assumed to follow a Poisson distribution with a mean which is a multiplicative function of the expected count (E_i) and a relative risk (θ_i). Mathematically speaking the distribution

[1] http://webhelp.esri.com/arcgisdesktop/9.2/index.cfm?TopicName=Natural$_b reaks_{\%}28 Jenks\%29$

of the area counts for area i is assumed to be

$$O_i \sim Poisson(E_i\theta_i)$$

with probability mass function

$$P(O_i|\theta_i) = \frac{(E_i\theta_i)^{O_i}e^{-(E_i\theta_i)}}{O_i!}.$$

Apart from an additive constant the log-likelihood for this model can be derived to be

$$l(\boldsymbol{\theta}|\mathbf{O}) = \sum_{i=1}^{m} O_i log(E_i\theta_i) - \sum_{i=1}^{m} E_i\theta_i$$

where m is the number of areas in the study region. If we differentiate the log-likelihood and follow the usual steps the maximum likelihood estimate of the relative risk $\widehat{\theta_i}$ is found to be $\frac{O_i}{E_i}$, which is equal to the SMR for area i, with an estimated standard error of $ese(\widehat{\theta_i}) = \frac{\sqrt{O_i}}{E_i}$.

This method shares some drawbacks with the SMR; the most extreme relative risk estimates will be those based on only a few cases and, on the other hand, the most extreme p-values from tests comparing relative risk estimates to 1 or those confidence intervals excluding 1 may simply identify regions with larger populations and hence more information (Mollié (1999)). These issues are more likely to occur when dealing with small study regions or rare diseases and mean that, although the relative risks have been estimated using a model when mapped, they may still be misleading. Mapping issues, including choosing specification intervals and colour schemes for cloropleth maps, are still relevant and can affect how the risk estimates are interpreted.

A further negative feature of the SMR/SIR and the basic Poisson model discussed here is that none of them considers the spatial structure of the study region. Using these methods it is not possible to account for the fact that areas which are in close proximity to one another often have similar levels

for many factors. This may affect the risk of disease, such as environmental factors or social views towards alcohol or drug use.

Extra-Poisson variation is a common problem when using this model and it occurs when the observed counts within the regions fluctuate around the mean for each region more than is expected for a Poisson model. The existence of such extra variation can give rise to unrepresentative geographic variation in the disease relative risks (Davies (2005)).

The level of over-dispersion present can be reduced by considering any available confounding variables during the standardisation stage, on top of age and sex. Many possible confounding variables relating to socioeconomic status, which can indicate local deprivation levels or lifestyle choices, could be included at this stage. Including such variables should result in a map which is a much better representation of the true underlying risk surface.

3.2.2 Fixed Effects

Although confounding variables can be included in the standardisation stage this may not be the best approach. For example, a categorical deprivation score could be included in the standardisation stage, but this will not allow the effect of deprivation to be estimated or its significance tested. Including deprivation score in a model for relative risk as a fixed effect will allow an estimate of the effect of deprivation on the relative risk to be quantified.

Since the relative risk (θ_i) must always be positive it is common to model $\log(\theta_i)$. Deprivation score could be included as a confounding variable in many ways in such a model, but in the following example we will model the logarithm of the relative risk as a linear function of deprivation score:

$$\theta_i = exp\{\beta_0 + \beta_1 x_i\},$$

where $\exp(\beta_0)$ represents the background risk across the entire study region, β_1 is a linear parameter and x_i represents the deprivation score in area i.

An example of fitting a spatial trend using the spatial coordinates of the tract centroids is shown in the fixed-effects section of Lawson et al. (2003*b*).

Many types of fixed-effects model for the relative risk (θ_i) can be fitted using conventional statistical packages which allow Poisson regression or log-linear modelling, such as R and S-Plus.

3.2.3 Random Effects

The modelling techniques mentioned so far assume that once all confounding variables are included in the model the resulting risk estimates will convey the true disease risk structure. Unfortunately, it is rarely the case that every confounding variable is measured or even thought of in such studies. There are almost always believed to be some unobserved factors, known as random effects, which affect the risk of disease as well as any observed factors. These random effects should be included in the risk-modelling process and the method for doing so has been the topic of much literature.

The consideration of random effects in disease mapping studies has become more common in recent years. In its simplest interpretation a random effect represents an extra component of variation which can be estimated within the study region and assigned a probability distribution. A possible source of this additional variation could be if a spatial covariate is interpolated to region centroids. When this happens there will be some degree of error in the estimated values and hence in any analysis which uses these values. Also, there may be some extra variation attributable to the regions themselves; for example, if local authority boundaries are used as tracts, there may be differences in council intervention programs for some diseases that the researcher is unaware of. When observed counts that are thought to follow a Poisson distribution exhibit a higher variance than expected, i.e. the variance is greater than the mean, it is known as overdispersion. Sparseness and clustering of disease cases can both cause overdispersion.

For spatial mapping models it is possible to break this extra variation

down into uncorrelated heterogeneity and correlated heterogeneity. Uncorrelated heterogeneity is simply a kind of independent and spatially uncorrelated additional variation, whereas correlated heterogeneity arises from a model which assumes that each spatial tract is correlated with the neighbouring geographical units. The correlated type implies that there is spatial autocorrelation between the tracts, which can arise if the disease cases are clustered throughout the study region or if there are unobserved factors at work in the data.

Some further discussion on random effects can be found in Lawson et al. (2003*b*).

3.3 Hierarchical Bayesian Disease Mapping

The development of Bayesian disease mapping models has helped to overcome the problem of over-dispersion and provide a means to include existing spatial information about the geographical distribution of disease risk across the study region.

Hierarchical Bayesian models, using which a problem is broken down into a series of levels linked by simple rules of probability, take on a very flexible framework capable of accommodating uncertainty and prior scientific knowledge while retaining many advantages of earlier likelihood methods (Arab et al. (2007)).

Since the introduction of the Bayesian hierarchical model and the development of Markov Chain Monte Carlo methods (discussed below) there has been a vast amount of research, both theoretical and applied, in this area. Good introductions and discussions on Bayesian hierarchical methods can be found in Congdon (2010), Congdon (2003), Carlin & Louis (2009) and Gelman et al. (2004*a*), while a good introduction to using these methods in a disease mapping context is given in Lawson et al. (1999) and Lawson et al. (2003*a*).

Bayesian disease mapping methods utilise two sources of information, using both the observed disease data together with prior knowledge about how the disease rates vary within the study region. The information that dominates depends on the study; in cases where there are a large number of disease cases in each area the abundance of data will lead the disease observations to dominate the analysis, whereas, when we are either looking at a rare disease or using small area tracts with low populations, the sparseness of events observed often leads to the prior information having a larger influence on the relative risk estimates.

3.3.1 Bayesian Approaches to Relative Risks

Bayesian methods incorporate the observed data through the likelihood of observed values (O_i) given the relative risk parameters (θ_i). Any prior beliefs about the geographic variation of the relative risks are catered for by assigning an appropriate probability distribution to $\boldsymbol{\theta}$ which is known as the joint prior distribution and denoted by $\boldsymbol{g}(\boldsymbol{\theta}|\boldsymbol{\delta})$, where $\boldsymbol{\delta}$ are hyperparameters.

This prior distribution for $\boldsymbol{\theta}$ explains all that is known about the relative risks before the study data has been collected. It is possible to use informative priors, weakly informative priors or non-informative priors. However, care must be taken since some seemingly uninformative prior distributions can prove to be quite informative. An informative prior is used when some information which is available before data collection is incorporated into the analysis; a non-informative prior is a common choice and is used to express that there is no knowledge of $\boldsymbol{\theta}$ before the data has been observed. A uniform distribution over the sample space is commonly used as a non-informative or diffuse prior.

The likelihood function of the relative risks given the observed disease counts is the product of m independent Poisson distributions, where m is the number of areas in the study region, since the O_i can be considered conditionally independent given $\boldsymbol{g}(\boldsymbol{\theta})$ (Mollié (1999)). The likelihood function

of relative risks given the observed data is therefore

$$L(\boldsymbol{O}|\boldsymbol{\theta}) = \prod_{i=1}^{m} L(O_i|\theta_i).$$

The aim of Bayesian analysis is to estimate the posterior distribution for $\boldsymbol{\theta}$, on which inference about the relative risks is based. This distribution describes the behaviour of the risk parameters when the data is observed and prior assumptions have been made. If we assume for now that all that is unknown are the relative risks then the posterior is given by

$$p(\boldsymbol{\theta}|\boldsymbol{O}, \boldsymbol{\delta}) \propto L(\boldsymbol{O}|\boldsymbol{\theta})\boldsymbol{g}(\boldsymbol{\theta}|\boldsymbol{\delta}).$$

It is unusual in practice to consider a completely specified prior distribution with known hyperparameters $\boldsymbol{\delta}$. The Empirical Bayes approach assumes that the hyperparameters are unknown and drawn from an unspecified probability distribution whereas the fully Bayesian approach uses a three-stage hierarchical model in which the hyperparameters are said to follow a specified probability distribution, known as the hyperprior distribution.

3.3.2 Empirical Bayes

It is common to distinguish between empirical Bayes methods and fully Bayesian methods on the basis that any method which seeks to approximate the posterior distribution is regarded as empirical Bayes and all others are regarded as fully Bayesian (Bernardo & Smith (1994)).

Using the empirical Bayes approach involves assuming that the hyper-parameters are unknown and are drawn from some unspecified distribution, and estimates of these hyper-parameters are used to work out the posterior distribution. Often, but not always, these estimates are obtained using maximum marginal likelihood or generalised least squares methods.

Informative discussions of several empirical Bayes methods can be found in Leyland & Davies (2005), Davies (2005) and Lawson et al. (2003*b*) along with useful references to research in this area.

3.3.3 Fully Bayesian

The fully Bayesian approach differs from the empirical Bayes approach, in that now the prior distribution is defined before the observed data is considered. The fully Bayesian approach involves fitting a hierarchical model where the distribution of the hyper-parameters ($\boldsymbol{\delta}$) is fully specified. This is distribution is known as the hyper-prior distribution $g_\delta(\boldsymbol{\delta})$ and is incorporated into the modelling process. If we now consider that neither the relative risks nor the hyperparameter values are known, the joint posterior distribution of the relative risks $\boldsymbol{\theta}$ and the hyper-parameters $\boldsymbol{\delta}$ given the observed data $\boldsymbol{O}$ is

$$p(\boldsymbol{\theta}, \boldsymbol{\delta}|\boldsymbol{O}) \propto L(\boldsymbol{O}|\boldsymbol{\theta}) g_\theta(\boldsymbol{\theta}|\boldsymbol{\delta}) g_\delta(\boldsymbol{\delta})$$

where $g_\theta(\boldsymbol{\theta}|\boldsymbol{\delta})$ is the prior distribution of $\boldsymbol{\theta}$.

The marginal posterior distribution for $\boldsymbol{\theta}$ given the observed data can be found by integrating out the hyperparameters as follows:

$$p(\boldsymbol{\theta}|\boldsymbol{O}) = \int p(\boldsymbol{\theta}, \boldsymbol{\delta}|\boldsymbol{O}) d\delta.$$

The use of a hierarchical structure leads the Bayes point estimates to be shrunk towards a value that is related to the distribution of all parameters in the hierarchical structure. It is assumed that the prior closely represents the "truth", and hence different prior choice should lead to different levels of shrinkage.

A comparison of some common Bayesian disease mapping models in terms of goodness-of-fit criteria is given by Lawson et al. (2000) and an in-depth review of the main spatial priors which have been proposed for fitting full Bayesian disease mapping models is given by Best et al. (2005).

The fully Bayesian approach has fairly recently become commonly used due to the increased availability of software which can perform Markov chain Monte Carlo methods of posterior simulation.

3.4 Posterior Inference

For simple likelihood models, like the Poisson model discussed above, often maximum likelihood is used to compute point estimates and associated variability for the parameters. When Bayesian hierarchical models are used the parameters are assumed to arise from a distribution of possible values rather than take on fixed values, meaning that it is no longer possible to provide simple point estimates for the θ_is in this way. In this case the posterior distribution must be found and examined to find point estimates such as the posterior mode or posterior mean for a parameter of interest. For some simple posterior distributions it is possible to find exact forms of these estimates, but in most realistic disease mapping models it is not possible to derive simple estimators for parameters such as the relative risk since a closed form of the posterior is unobtainable. In these situations posterior sampling must be used.

Posterior sampling involves using simulation methods to gain samples from the posterior distribution which are then summarised to get estimates of the desired parameters. The remainder of this section discusses some of the posterior simulation methods which can be used.

3.4.1 Markov Chain Monte Carlo Methods

Markov Chain Monte Carlo (MCMC) methods are efficient and flexible posterior sampling methods which can be applied to a variety of models (see e.g. Lawson et al. (2003*a*)). Such methods have been incorporated into several statistical packages including WinBUGS and OpenBUGS.

Most MCMC methods aim to produce a sample from the joint posterior

distribution. To do this a Markov chain must be constructed such that the proposed distribution is easy to sample from and represents the joint posterior distribution. The parameter values are then iteratively simulated within this Markov chain and the iteration process continues until the chain converges to a stationary distribution. Once a stationary distribution is reached, the chain is assumed to represent the posterior distribution. If the chain has run for a sufficient number of iterations, realised values from this chain can be used to estimate various properties of the posterior distribution of the parameters.

Put simply, MCMC iterations involve using only the most recent values of the parameters, to generate proposed new values from given probability distributions. The posterior probability of the new values is compared with that of the old values and then new values will be accepted according to a certain rule. If the new values are accepted then these values will replace the existing values to become the current parameter values. This process will repeat many times, each time simulating an estimate for each unknown parameter. The idea is that the output from each iteration together will form sample from the joint posterior distribution of unknown parameters.

Some algorithms used to construct the required Markov chain are discussed in the following sub sections.

3.4.2 Sampling Algorithms

It is essential for all MCMC algorithms that the right transition probabilities for a Markov chain which has the joint posterior distribution, $P(\boldsymbol{\theta}|\boldsymbol{O})$, as its equilibrium distribution can be constructed. These transition probabilities will be defined for a Markov chain consisting of $\boldsymbol{\theta}^1$, $\boldsymbol{\theta}^2$, ... $\boldsymbol{\theta}^t$ with state space Θ and equilibrium distribution $P(\boldsymbol{\theta}|\boldsymbol{O})$ (Lawson (2009*a*)) below.

Let $q(\boldsymbol{\theta}, \boldsymbol{\theta}^{'})$ be a transition probability function, sometimes referred to as the proposal density, where $\boldsymbol{\theta}$ represents the current values of the parameters and $\boldsymbol{\theta}^{'}$ represents the new proposed values. The algorithms use this proposal

density, which depends only on $\boldsymbol{\theta}$ the latest chain values for the parameters, to generate new proposed parameter values $\boldsymbol{\theta}^{'}$.

Metropolis Updates

For Metropolis updates a symmetric proposal function, $q(\boldsymbol{\theta}, \boldsymbol{\theta}^{'})$, should be chosen. Then the transition probabilities for a discrete distribution can be defined as

$$p(\boldsymbol{\theta}, \boldsymbol{\theta}^{'}) = \begin{cases} \alpha(\boldsymbol{\theta}, \boldsymbol{\theta}^{'}) q(\boldsymbol{\theta}, \boldsymbol{\theta}^{'}) & \text{if } \boldsymbol{\theta}^{'} \neq \boldsymbol{\theta} \\ 1 - \sum_{\boldsymbol{\theta}^{''}} \alpha(\boldsymbol{\theta}, \boldsymbol{\theta}^{''}) q(\boldsymbol{\theta}, \boldsymbol{\theta}^{''}) & \text{if } \boldsymbol{\theta}^{'} = \boldsymbol{\theta} \end{cases}$$

where $\alpha(\boldsymbol{\theta}, \boldsymbol{\theta}^{'}) = \min \left\{ 1, \frac{P(\boldsymbol{\theta}^{'}|\boldsymbol{O})}{P(\boldsymbol{\theta}|\boldsymbol{O})} \right\}$ and $\boldsymbol{\theta}^{''}$ represents any permitted combination of parameter values which is not the same as the current values $\boldsymbol{\theta}$.

Here $\alpha(\boldsymbol{\theta}, \boldsymbol{\theta}^{'})$ represents the acceptance probability and the proposed values $\boldsymbol{\theta}^{'}$ will be accepted with this probability.

For Metropolis updates the proposal function must be an irreducible and aperiodic transition function.

Metropolis-Hastings Updates

The Metropolis-Hastings algorithm is an extension to the Metropolis algorithm in which the proposal function no longer needs to be symmetric and

$$\alpha(\boldsymbol{\theta}, \boldsymbol{\theta}^{'}) = \min \left\{ 1, \frac{P(\boldsymbol{\theta}^{'}|\boldsymbol{O}) q(\boldsymbol{\theta}^{'}, \boldsymbol{\theta})}{P(\boldsymbol{\theta}|\boldsymbol{O}) q(\boldsymbol{\theta}, \boldsymbol{\theta}^{'})} \right\}.$$

For this algorithm the definition of the proposal function can be quite general and posterior distribution only needs to be known up to a proportionality. Metropolis-Hastings updates also require that the proposal function is irreducible and aperiodic.

Gibbs Updates

The Gibbs sampler is one of the more popular algorithms to use with Bayesian hierarchical models. It is a special case of the Metropolis-Hastings

algorithm where the proposal is generated for each θ_i from the conditional distribution of θ_i given all other elements of $\boldsymbol{\theta}$. The new parameter value which is proposed is always accepted, i.e. the acceptance probability is always 1.

If, say, θ_i is to be updated, then $\theta'_j = \theta_j$ for $j \neq i$ and

$$\theta_i \sim P\left(\theta_i^* | \theta_{-i}\right)$$

in which $p\left(\theta_i^* | \theta_{-i}\right)$ represents the conditional distribution of θ_i given that

$$\theta_{-i} := \left\{\theta_j, j \neq i\right\}.$$

3.4.3 Convergence

When using MCMC methods it is necessary to assess whether the iterative simulations have converged to the equilibrium distribution of the Markov chain. Each chain must be run for a sufficiently long burn-in period to allow convergence to this distribution to occur and all parameter values simulated during this burn-in period should be discarded from further analysis and parameter estimation. The length of this burn-in period can be very different between different problems. The burn-in period also needs to be long enough to allow the full parameter space to be explored and avoid the estimator becoming stuck at a local maxima rather than the global maxima.

It is therefore crucial to check that there has been an adequate burn-in period. There are several methods to check for convergence, although there is no way to be totally sure, and most methods are at least slightly subjective.

Several convergence diagnostics are discussed in Cowles & Carlin (1996). These diagnostics can be split into those which require multiple chains to be run in parallel with different starting values and those which can be applied to single chains. Obviously those methods which can be applied to single chains can be applied individually to each chain in multiple-chain examples.

Single-Chain Methods

Various diagnostic methods for assessing convergence of single chains have been suggested, including monitoring the stability of functions of the posterior probability across the iterations, the Brooks-Draper diagnostic (Brooks & Draper (1999)) and the Raftery-Lewis diagnostic (Raftery & Lewis (1992)).

The most common method to visually check for convergence of a single chain is to look at a history graph which plots the simulated parameter value at each iteration against the iteration number. When this plot shows no obvious patterns or trends and looks roughly like a horizontal band across the plot then it is likely that the chain has converged. However, even when the history plot does look like this it does not necessarily mean that the whole parameter space has been explored.

Multi-Chain Methods

The most popular multi-chain convergence diagnostic is the Gelman-Rubin diagnostic plot, which is produced by WinBUGS. This plot uses a green line to show the width of the central 80% interval of the pooled chains, a blue line to show the average width of the 80% intervals within the individual runs and red to represent their ratio R=(pooled/within). When checking for convergence one should be looking for R to settle at a value of 1 as well as for the pooled and within-interval widths to reach stability.

Checking whether multiple chains have converged can also be done visually using history plots as described for single-chain methods. In this case separate lines will be shown for each chain and when convergence is reached these lines should form consistently overlapping horizontal bands across the plot. If the lines for each chain form horizontal bands that do not overlap then this can indicate that some or all of the chains have become "stuck" at local maxima as described above.

3.5 Goodness-of-Fit

If the MCMC algorithms converge this does not necessarily mean that the model is a good fit to the data. Many issues relating to model goodness-of-fit should be considered.

The deviance is a measure often used in Bayesian statistics when looking at the goodness-of-fit of a model. One disadvantage of using the estimated deviance directly is that it does not incorporate the level of parameterisation in the model; it is always possible to improve the fit of a model by adding in further parameters, unless the model is already saturated.

Commonly used methods such as the Akaike information criterion (AIC) and Bayesian information criterion (BIC) aim to penalize for model complexity according to the number of parameters in the model.

In hierarchical Bayesian disease mapping studies the most common measure of goodness-of-fit is the Deviance information criterion (DIC).

3.5.1 DIC

Like the AIC and BIC methods DIC aims to penalize more complex models. The DIC was proposed by Spiegelhalter et al. (2002) and has the basic principle of being a measure of goodness-of-fit plus a penalty for model complexity (Spiegelhalter (2006)).

As the name suggests, the goodness-of-fit element is based on the deviance, which is given by

$$D(\boldsymbol{\theta}) = -2logL(\boldsymbol{O}|\boldsymbol{\theta}).$$

The effective number of parameters in the model, p_D, is estimated and used as a measure of model complexity. Spiegelhalter et al. propose that

$$p_D = E_{\boldsymbol{\theta}|\boldsymbol{O}}[D] - D(E_{\boldsymbol{\theta}|\boldsymbol{O}}[\boldsymbol{\theta}]),$$

often denoted by

$$p_D = \overline{D} - D(\overline{\boldsymbol{\theta}}),$$

where $\overline{D}$ is the posterior mean deviance and $D(\overline{\boldsymbol{\theta}})$ is the deviance calculated at the posterior mean of the unknown parameters. These quantities are easily monitored when using MCMC methods in OpenBUGS.

An alternative estimate of the effective number of parameters, proposed by Andrew Gelman and discussed in Gelman et al. (2004*b*) and Lawson (2009*b*), is half the posterior variance of the deviance,

$$p_D^* = \frac{1}{2} var\{D\}.$$

It should be noted that Gelman's p_D^* tends to over-estimate the effective number of parameters in a model, meaning that more complex model structures may be over-penalized. This is likely to be more of an issue when dealing with complicated hierarchical models. The measure is, however, invariant to parameterisation and easy to calculate (discussion given on the DIC BUGS website[2]). A very interesting discussion into some potential pitfalls of the p_D^* measure of model complexity is given on Andrew Gelman's website [3].

The posterior variance of the deviance can also be easily estimated by working out the variance of the deviance values simulated in the MCMC chain.

The DIC statistic is then calculated as either

$$DIC = \overline{D} + p_D$$

or

$$DIC = \overline{D} + p_D^*.$$

When comparing models it is believed that those with a lower DIC are a better fit to the data.

[2]http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/dicpage.shtml

[3]http://andrewgelman.com/2006/07/number_of_param/

3.6 Besag, York and Mollié Model

As discussed in section 3.2.3 when modelling relative risks of disease it is possible to include random effects to account for any extra variation or overdispersion present. Also discussed was that for spatial models this extra variation can be broken down into uncorrelated heterogeneity and correlated heterogeneity, where the former is just independent and spatially uncorrelated additional variation and the latter assumes that each area is correlated with its neighbouring geographic units.

The Besag, York and Mollié model is a fully Bayesian disease mapping model which does just this. The area-specific random effects are decomposed into an element which takes into account the effects that vary in an unstructured manner between areas (correlated heterogeneity) and an element which models the effects which vary in an unstructured manner across the study region (uncorrelated heterogeneity).

The model was initially established by Clayton & Kaldor (1987), further developed by Besag et al. (1991) and has been used in several disease mapping studies. If we continue to let O_i and E_i represent the observed and expected number of disease cases in area i respectively, and let θ_i stand for the relative risk in area i, then this model can be written as

$$\begin{aligned} O_i &\sim Poisson(E_i\theta_i) \\ log(\theta_i) &= \alpha + u_i + v_i \end{aligned}$$

where α is a baseline or overall level of relative risk, u_i represents correlated heterogeneity and v_i the uncorrelated heterogeneity. The log of θ_i is modelled as opposed to θ_i to ensure that any estimated relative risks are not negative.

Since this is an example of a Bayesian model, prior distributions must be specified for these random effects. The prior for the uncorrelated heterogeneity is a normal distribution with mean 0 and precision τ_v^2,

$$v_i \sim N(0, \tau_v^2).$$

The correlated heterogeneity is said to follow a spatial correlation structure, where estimation of the relative risk in each area depends on its neighbouring areas. The specific prior used is the conditional autoregressive (CAR) model introduced by Besag et al. (1991). This prior states that

$$[u_i|u_j, i \neq j, \tau_u^2] \sim N(\overline{u_i}, \tau_i^2)$$

where $\overline{u_i}$ is the mean of the areas bordering area i,

$$\begin{aligned} \overline{u_i} &= \frac{1}{\sum_j \omega_{ij}} \sum_j u_j \omega_{ij}, \\ \tau_i^2 &= \frac{\tau_u^2}{\sum_j \omega_{ij}}, \end{aligned} \tag{3.2}$$

and ω_{ij}=1 if area i and area j are adjacent, or ω_{ij}=0 if they are not.

The hyperparameters τ_v^2 and τ_u^2 control the variability of random effects v and u. If this is to be a fully Bayesian example then these hyperparameters need to be assigned hyperpriors. These are both often assigned gamma(ε, ρ) priors with some appropriate set values for ε and ρ.

3.7 Alternatives to the Besag, York and Mollié Model

Although the Besag, York and Mollié model seems to be the most popular disease mapping model, there are several alternatives to the conditional autoregressive prior structure. One such alternative specification involves only a single random effect which covers both correlated and uncorrelated heterogeneity. This can be done in practice by specifying a prior distribution which has two parameters which govern these affects. An example is given by Diggle et al. (1998), in which the covariance matrix of a multivariate normal prior distribution is parametrically modeled using such terms.

This approach is related to universal Kriging (Cressie (1993)), which involves covariance models that use variance and covariance range parameters.

These methods are commonly known as 'generalised linear spatial modelling.' It is common for these parameters to define a multiplicative relationship between the correlated and uncorrelated heterogeneity. The fully Bayesian analysis of this model also requires the use of posterior sampling algorithms similar to those discussed above.

In comparisons of CAR models and such fully-specified covariance models there appears to be differing opinions about which are most useful in estimating relative risk in disease maps (Best et al. (2005) and Henderson et al. (2002)).

Further disease mapping methods have been suggested by Leroux (2000) which use maximum likelihood estimation for a generalised linear mixed model. This model allows for log-linear covariate adjustments and localised smoothing of rates through the estimation of correlated random effects. The covariance structure of the random effects is based on a recently proposed model which parameterises spatial dependence through the inverse covariance matrix. Markov chain Monte Carlo simulation methods are also required to fit this model.

Chapter 4

Standardised Incidence Ratio

As discussed in Chapter 3 the SMR and SIR have some serious drawbacks. However, it is still of interest to look at these results to gain an initial impression of the risk surface. Doing so also gives the ability to compare these disease maps to model-based disease maps in terms of the degree of smoothness and general pattern. All SIR values discussed here have been calculated at the Scottish Data zone level using the data discussed in Chapter 2 and the methods described in Chapter 3.

This chapter will firstly discuss the SIR values for the combined male and female data and then go on to look at the SIR values for each gender individually. Due to the extremely small area of many inner-city data zones several regions of the Scotland maps produced will need to be magnified. Each time a map is discussed a complete map of Scotland will be shown, along with further magnified maps of the Aberdeen, Ayrshire, Dundee, Edinburgh, Glasgow, Inverness and Stirling areas. The magnified areas will always use the same risk ranges and colour key as the full Scotland map so that they are comparable.

4.1 Combined Male and Female SIR

The SIR values discussed in this section have been calculated using the combined male and female data. Age and sex standardisation has been used as described in Chapter 3.

Table 4.1 below identifies the data zones which exhibit ten of the lowest and the ten highest SIR values. The 10 zero SIR values shown in Table 4.1 are just sample of the 63 zero SIR values observed. This highlights one of the main disadvantages of using the SIR as a risk estimate: since there were no observed deaths or hospitalisations in these data zones the SIR indicates that there is no risk in these areas, which obviously cannot be true. Of the 63 data zones for which the combined SIR value is zero, all of which have a deprivation score of 6 or more, with 32 of these areas having the least-deprived deprivation score of 10.

In contrast, the ten highest combined SIR values shown in Table 4.1 range from 4.259 to 6.308 and all relate to data zones in the most deprived category. These values clearly show how unevenly alcohol-related risk is distributed across Scotland, with a data zone within Parkhead West and Barrowfield experiencing over 6.3 times the number of deaths and hospitalisations due to alcohol that was expected. It definitely appears that there is a strong association between deprivation score and SIR value. However, this association appears to be particularly strong for the worst deprivation score of 1. Of the 100 highest combined SIR values in Scotland, 81 correspond to areas with a deprivation score of 1, and all but 1 relate to a score of 3 or less.

Given below in Figure 4.1 is a violin plot for the combined SIR values for every data zone in Scotland grouped by deprivation score. The violin plot used here was created using the vioplot package in R, which combines a boxplot and a (doubled) kernel density plot. These plots show that the median SIR values increase from deprivation score 10 through to deprivation score 1. The difference in SIR values between consecutive deprivation scores

appears to be greater for the more deprived scores, and is largest between deprivation scores 1 and 2. This indicates that, if there does prove to be a relationship between deprivation score and combined SIR, it may not be linear with deprivation treated as a bona fide numerical score.

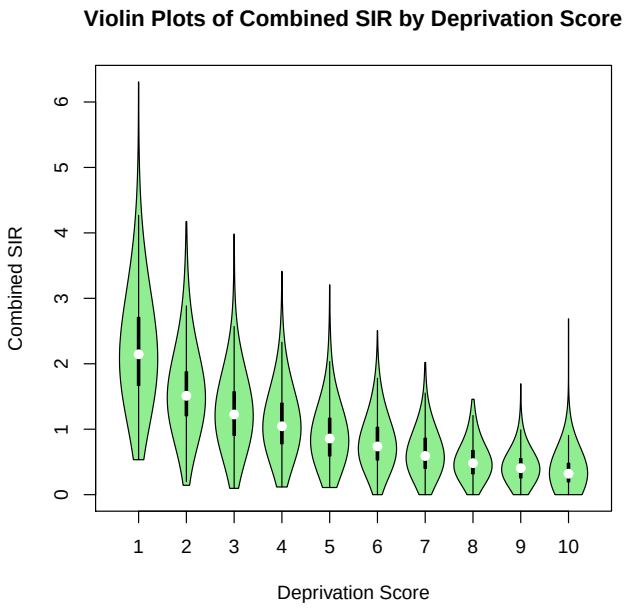


Figure 4.1: Violin Plots of Combined SIR by Deprivation Score

On top of the indication of a relationship between deprivation score and combined SIR value, Table 4.1 also suggests that there may be spatial clustering in alcohol-related deaths and hospitalisations in Scotland. Of the 10 data zones with the highest combined SIR values, 5 fall within the Glasgow City local authority. When looking at the 100 highest combined SIR values there is some fairly strong evidence of spatial clustering, as 51 of these areas fall within Glasgow City.

4.1.1 Combined Male and Female SIR Maps

A data zone choropleth map of Scotland showing the combined SIR values is shown in Figure 4.2. Magnified sections of this map are given for the Aberdeen, Ayrshire, Dundee, Edinburgh, Glasgow, Inverness and Stirling

areas in Figure 4.3, Figure 4.4, Figure 4.5, Figure 4.6, Figure 4.7, Figure 4.8 and Figure 4.9 respectively.

Firstly, if we compare these maps to the corresponding deprivation score data zone maps in Chapter 2 the similarities in patterning are striking. This data definitely seems to agree with previous studies in that alcohol-related health risks appear to be much higher in more deprived areas.

Figure 4.2 indicates that the SIR values are lower in the East of Scotland. However, it must be remembered that the choice of specification intervals and colour scheme can affect the interpretation of such maps. Two further plots have therefore been created: a scatter plot of combined SIR value against data zone centroid easting coordinate using hexagonal binning (Figure 4.10a) and a scatter plot of combined SIR against easting coordinate (Figure 4.10b). The lowess (locally weighted scatterpoint smoothing) line has been imposed on to Figure 4.10b using R. Both plots in Figure 4.10 provide further evidence that the SIR values do tend to be lower in the East of Scotland, although the relationship does not appear quite as strong as Figure 4.2 suggests.

There also seems to be some indication in Figure 4.2 that SIR values are higher in the north of Scotland. Similar plots have been created to objectively look at how the combined SIR values relate to how far north the data zones are; Figure 4.11a shows a scatter plot of combined SIR against data zone northing coordinate using hexagonal binning and Figure 4.11b shows a scatter plot of combined SIR against northing coordinate with a superimposed lowess line. Neither plot in Figure 4.11 shows a particularly strong relationship, although, with the exception of the northing range of around 630000 to 700000, there is slight evidence of the combined SIR increasing as you go further north in Scotland.

Looking at the full SIR map of Scotland in Figure 4.2 allows us to gain a picture of the large sparsely populated rural data zones of Scotland. Most of these rural areas have a combined SIR value less than 0.88, so in general, large rural areas observed fewer total deaths and hospitalisations than expected.

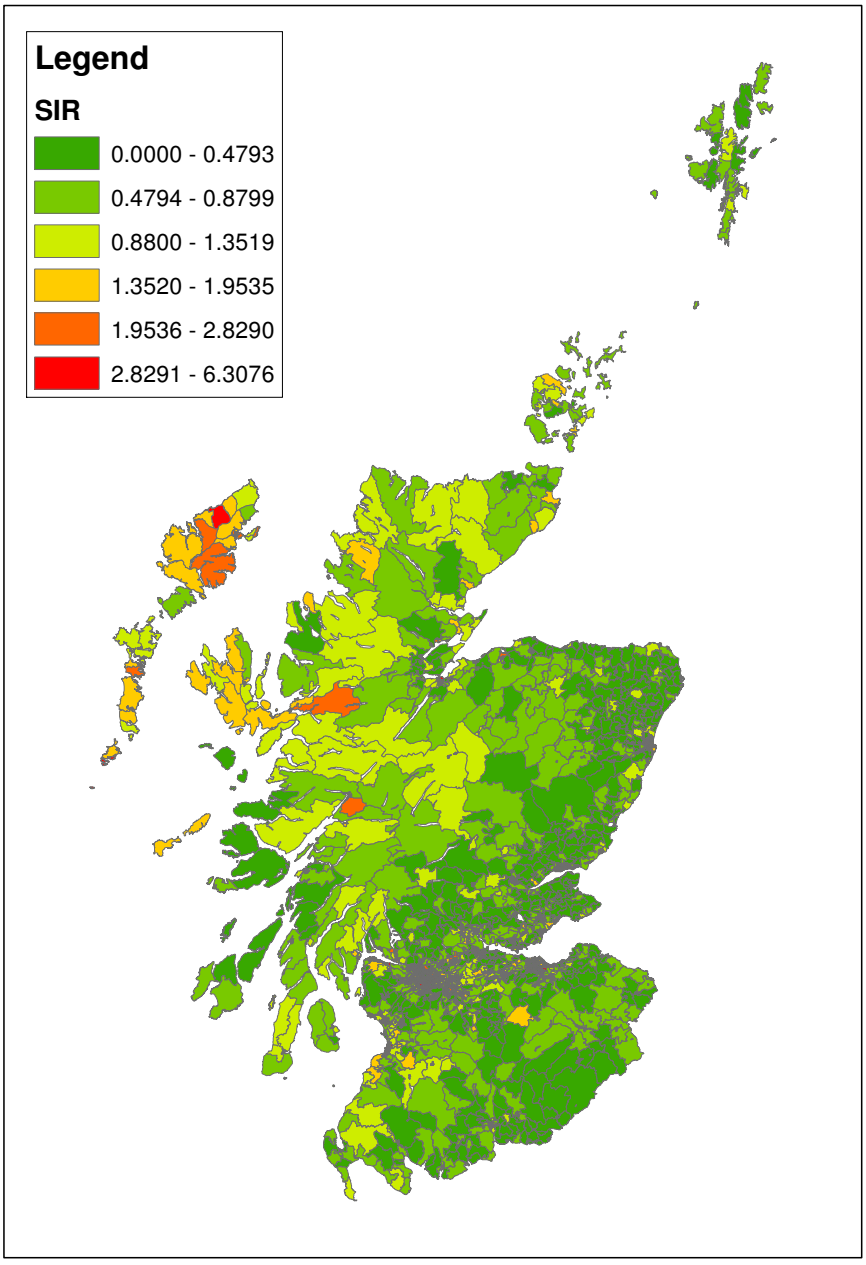


Figure 4.2: Data Zone Map of Alcohol-related SIR

However, in the North West of Scotland and particularly in the regions of the Inner and Outer Hebrides and the Isle of Skye alcohol-related risk appears to be higher than average. There are also two mainland data zones with SIR values between 1.95 and 2.83 which stand out, one in the Ben Nevis area and

the data zone directly north of Loch Alsh.

Now to consider the two largest cities in Scotland and compare the combined SIR maps for Glasgow and Edinburgh in Figure 4.7 and Figure 4.6 respectively. Even at first glance these maps show that Glasgow has far higher SIR values than Edinburgh on average. The majority of data zones in Edinburgh appear to have an SIR value of less than 0.88 and the relatively few data zones in this city which have a SIR values in excess of 1.9536 seem to lie on the periphery of the city. In contrast, the Glasgow combined SIR map shows a high density of extremely high SIR values in inner city areas. SIR values of over 2.8 are shown across many parts of the city, but appear to be most common in the East of Glasgow and to the South of the Clyde. There are some definite clusters of high SIR values in the East of Glasgow.

All of the magnified areas of the combined SIR map, shown in Figures 4.3 to 4.8, indicate that there is much greater variation in the area of data zones within the lower SIR classes. The data zones with the highest SIRs tend to be small densely populated areas with high deprivation levels. All of these magnified maps also show a strong relationship between SIR value and deprivation as well as highlighting that there is no "norm" pattern in SIR values for towns and cities across Scotland.

4.2 Male SIR

This section considers the SIR values calculated using only the male data. In this case the methods described in Chapter 3 were used to calculate the SIR, except that sex standardisation is obviously no longer needed. This section aims to compare the male SIR results with the combined SIR results.

A table showing the 10 highest and 10 lowest male SIR values in Scotland has been produced and is shown in Table 4.2 below. Obviously, since the 10 lowest combined SIR values are zero, the 10 lowest male SIR values are also zero. In fact, for males there were 196 data zones which experienced no

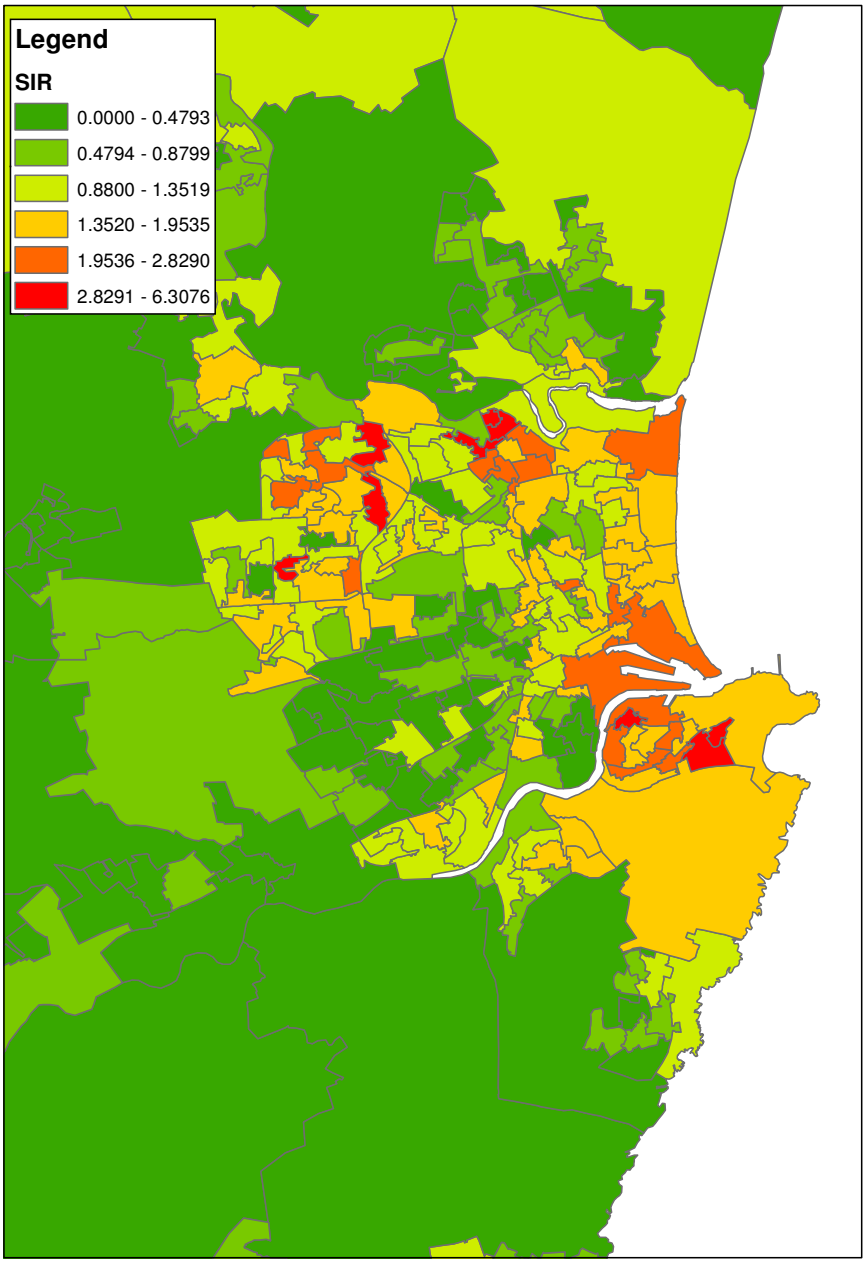


Figure 4.3: Data Zone Map of Aberdeen Alcohol-related SIR

alcohol-related deaths or hospitalisations; of these areas over 39.7% have a deprivation score of 10 and only 9.18% have a deprivation score less than 7.

The 10 highest male SIR values, as shown in Table 4.2, also all relate to areas with the most deprived score of 1. Upon looking at the 100 highest

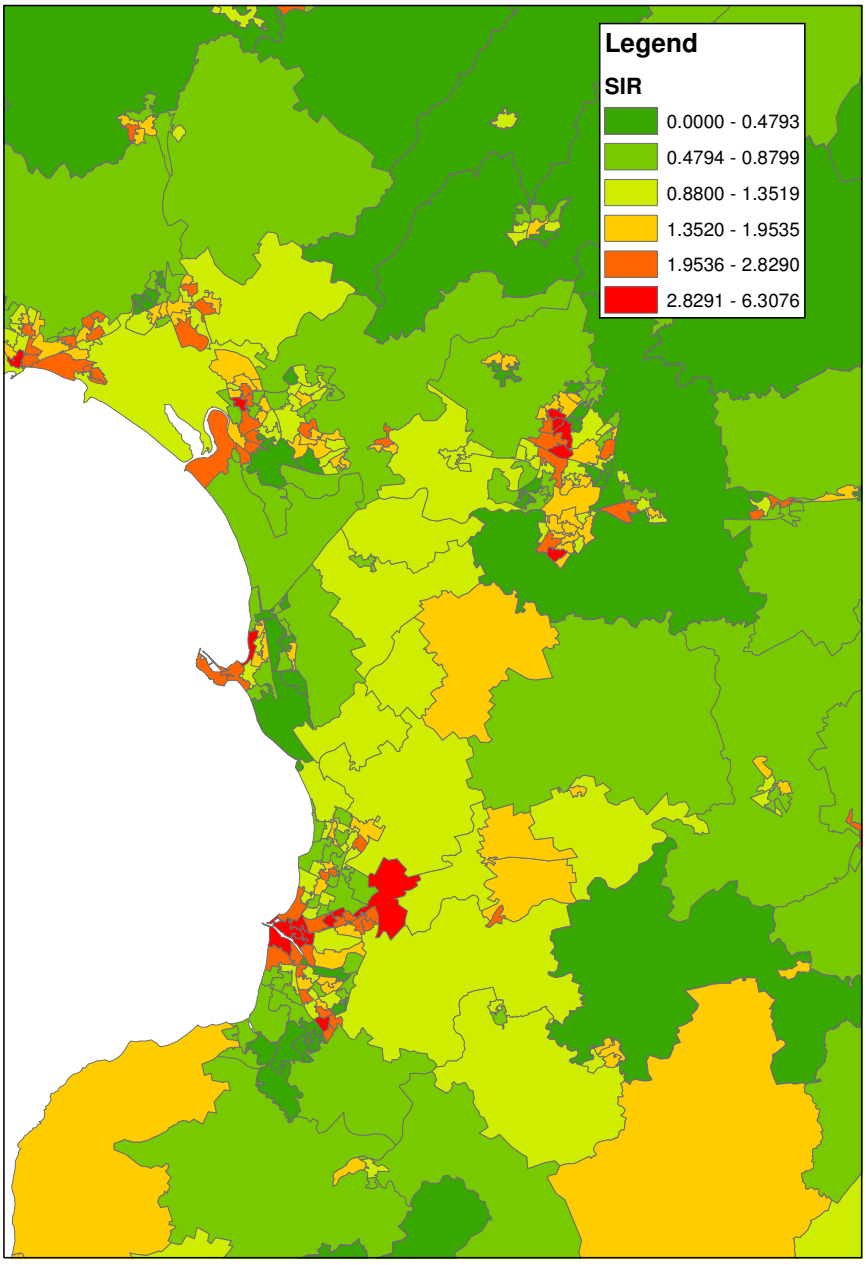


Figure 4.4: Data Zone Map of Ayrshire Alcohol-related SIR

male SIR values, it can be seen that 86 of these correspond to areas with a deprivation score of 1 and 53 fall within Glasgow City. The male SIR results are therefore similar to the combined SIR results. This is to be expected since, as is shown in Chapter 2, there were many more recorded male alcohol-

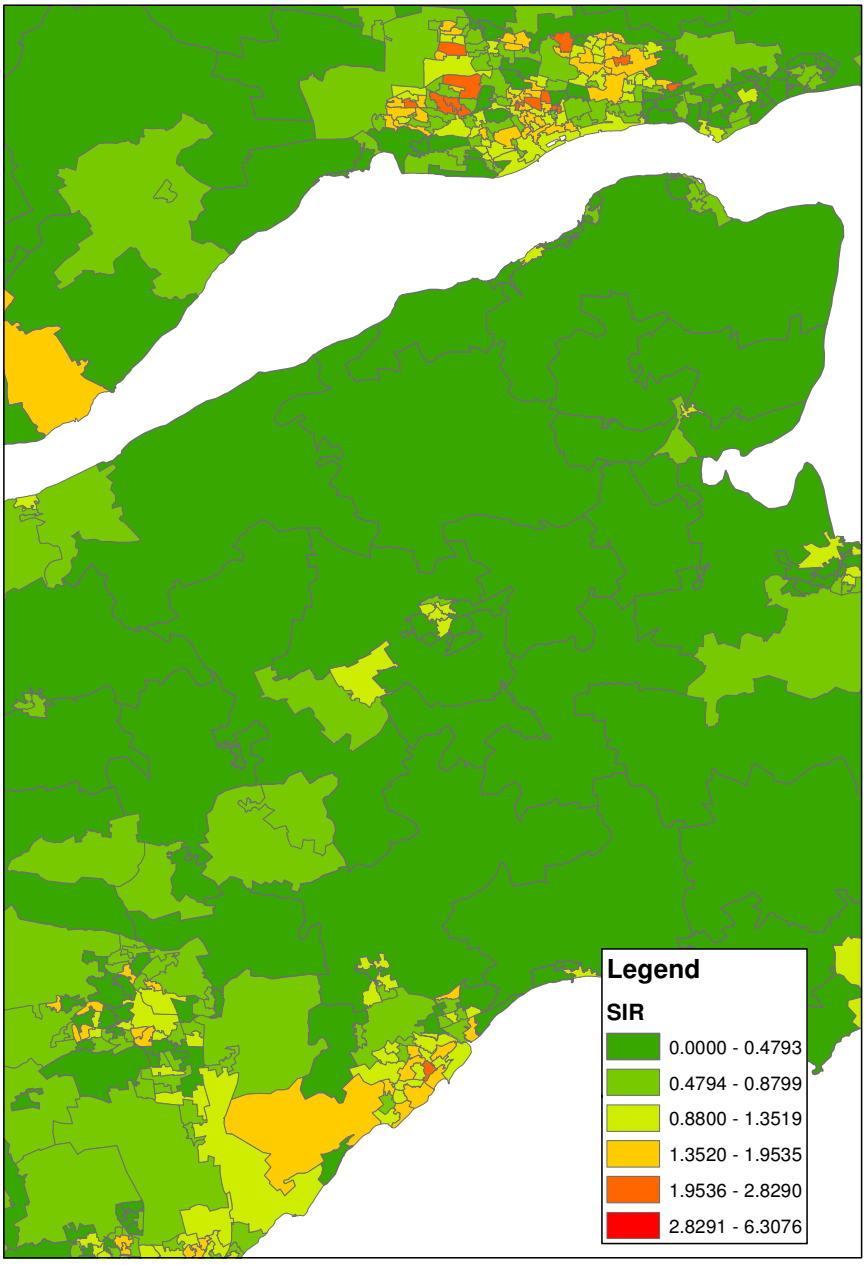


Figure 4.5: Data Zone Map of Fife Alcohol-related SIR

related deaths and hospitalisations during the study period; as a result the male data will have a greater influence on the combined data. The results suggest that there is a strong relationship between deprivation and alcohol-related health risks for males; also that any such relationship may not be

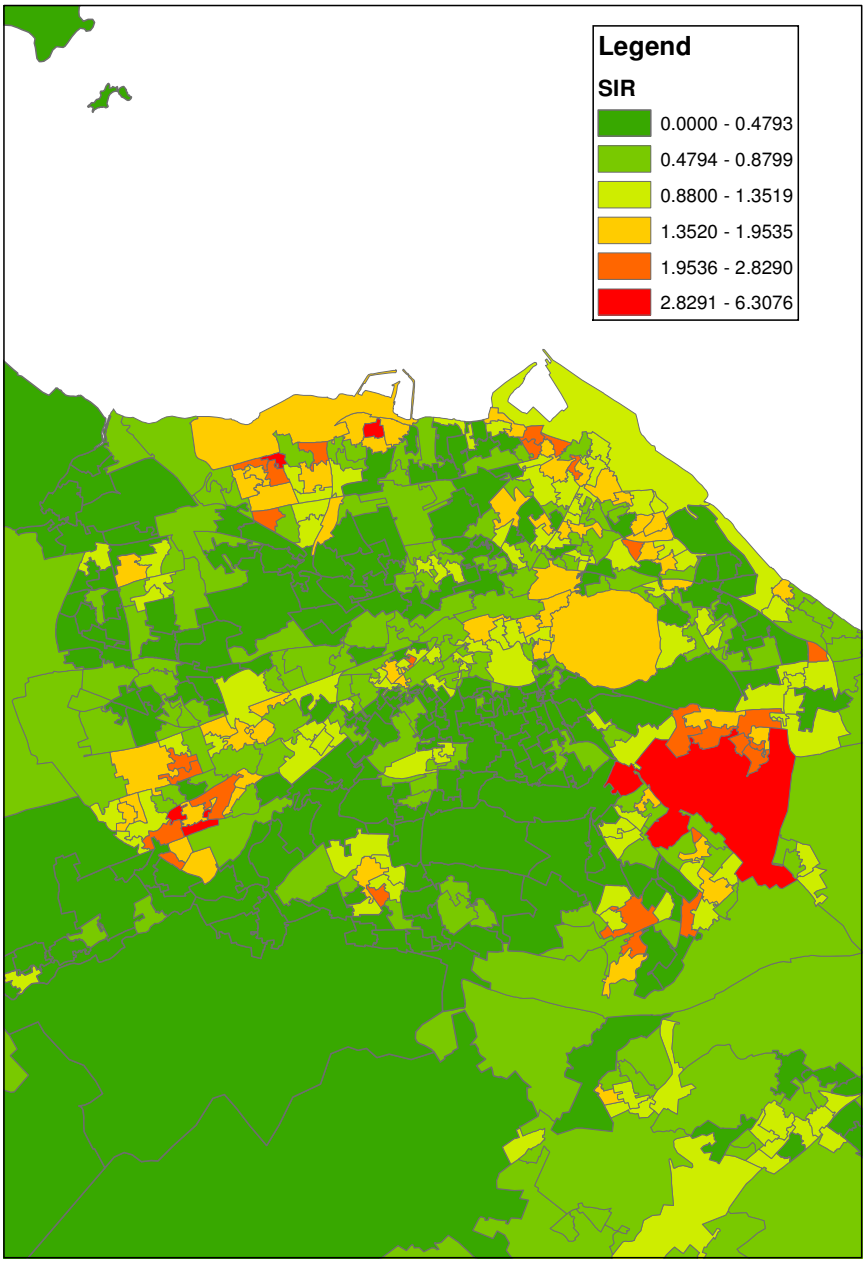


Figure 4.6: Data Zone Map of Edinburgh Alcohol-related SIR

linear, since the worst deprivation score of 1 appears to be more strongly associated with very high male SIR values, than the best deprivation score of 10 is with very low values.

Violin plots have also been produced for the male SIR values and are

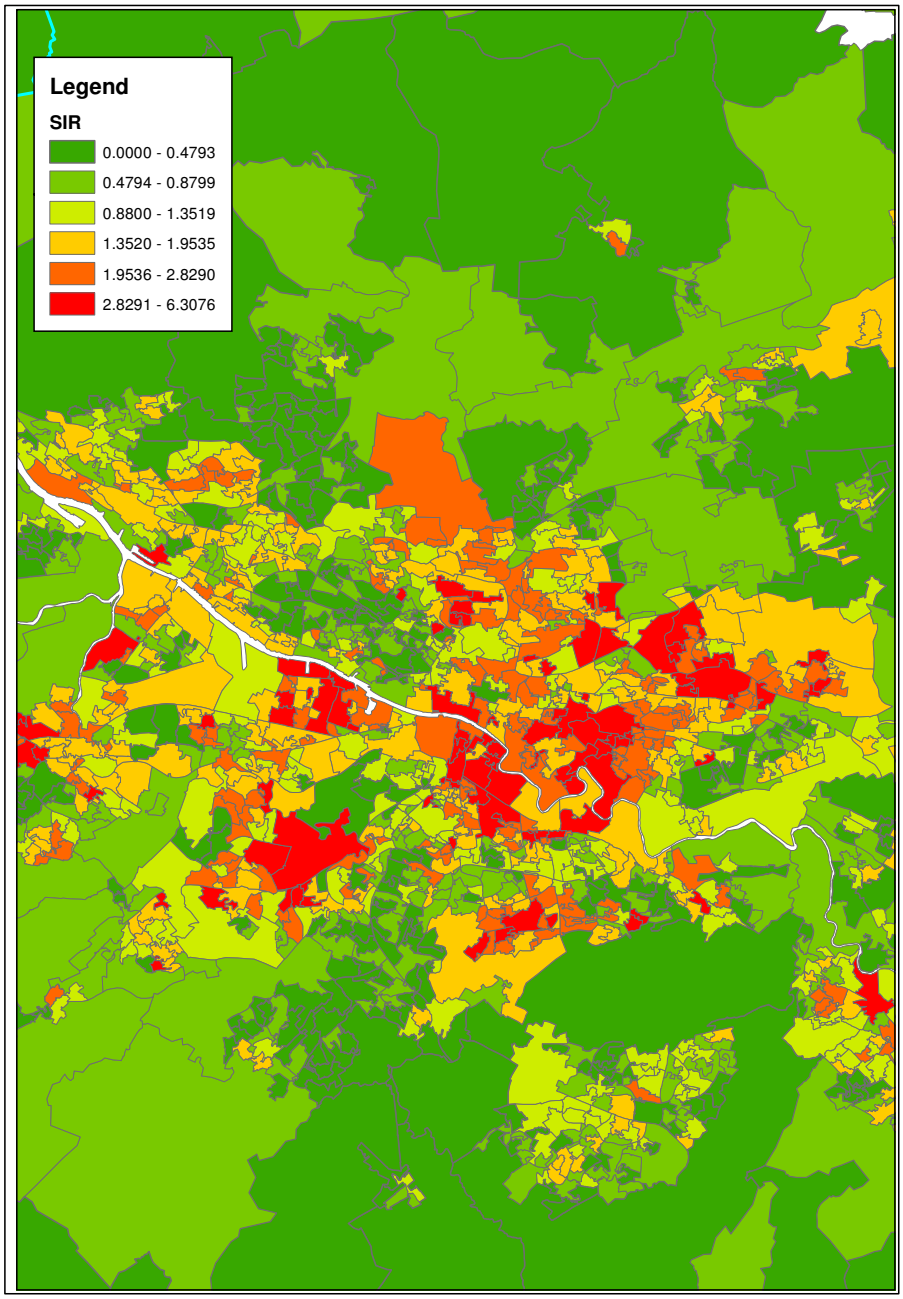


Figure 4.7: Data Zone Map of Glasgow Alcohol-related SIR

shown in Figure 4.12. These show a very similar picture to that in Figure 4.1 above; the male SIR values are greater on average for lower (worse) deprivation scores. Again, the difference in average male SIR value between successive deprivation scores increases as deprivation score decreases, with

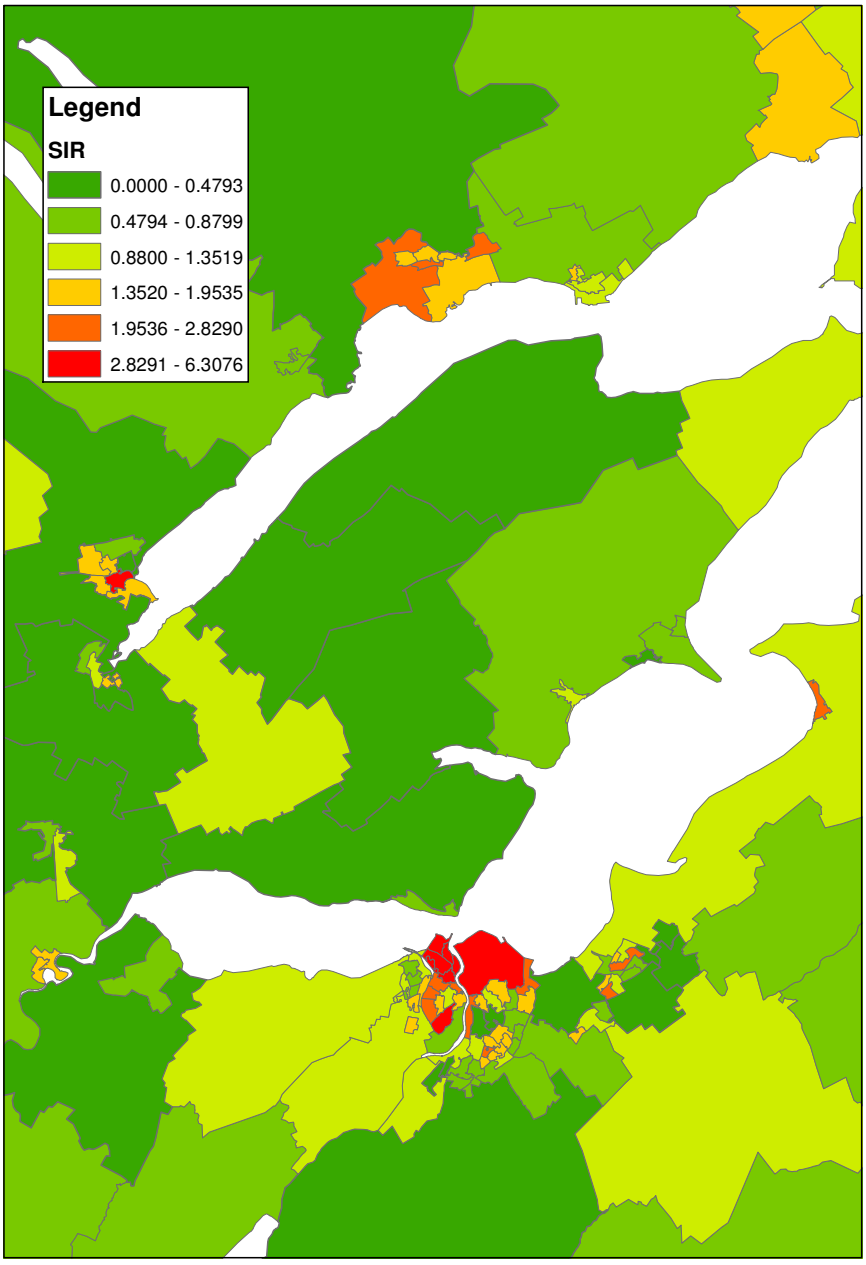


Figure 4.8: Data Zone Map of Inverness Alcohol-related SIR

the greatest difference occurring between deprivation score 1 and 2. It again appears then, that any relationship between deprivation score and alcohol-related risk will not be linear but will be monotonic.

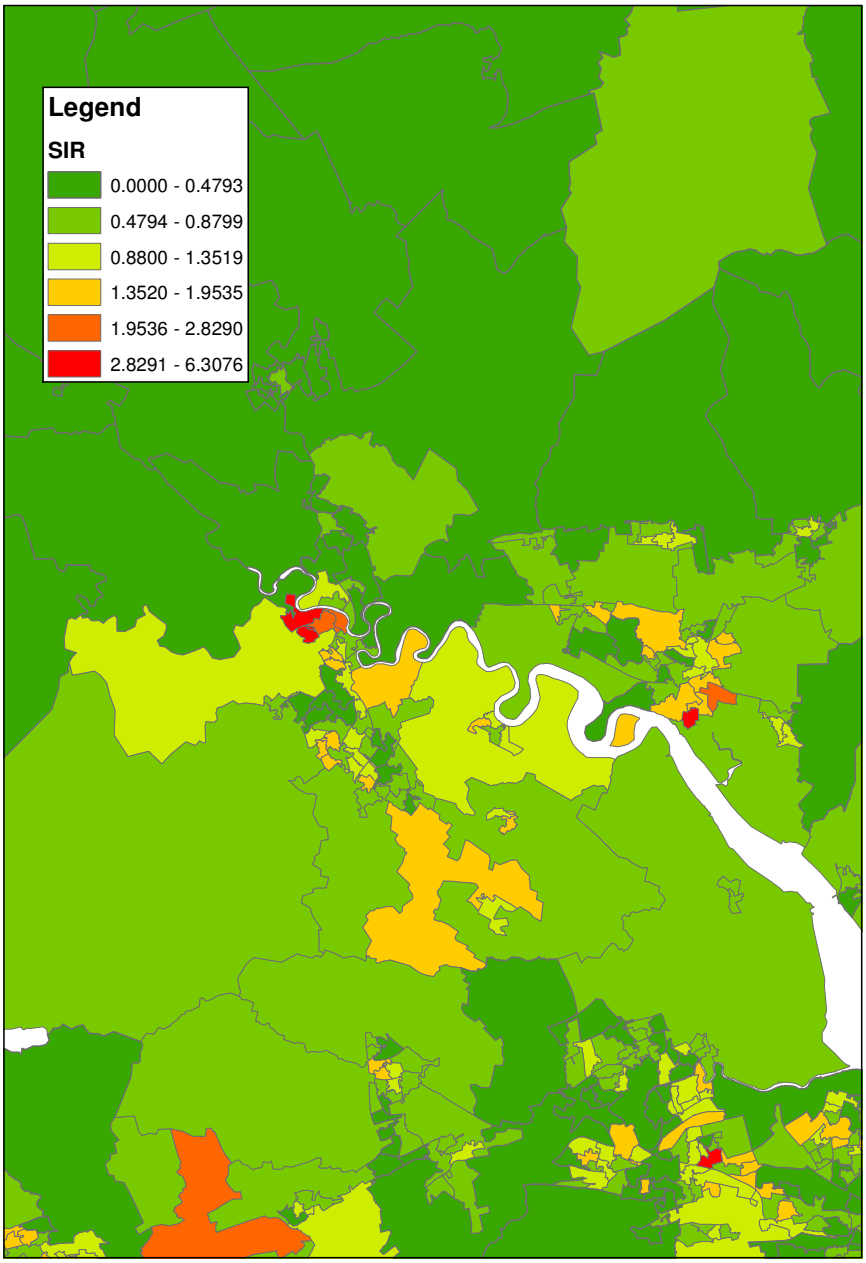
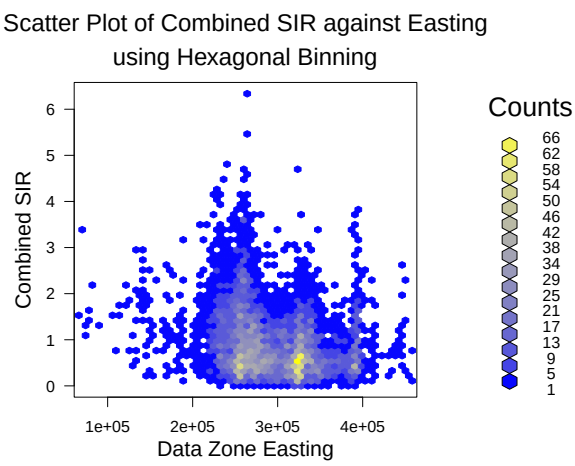


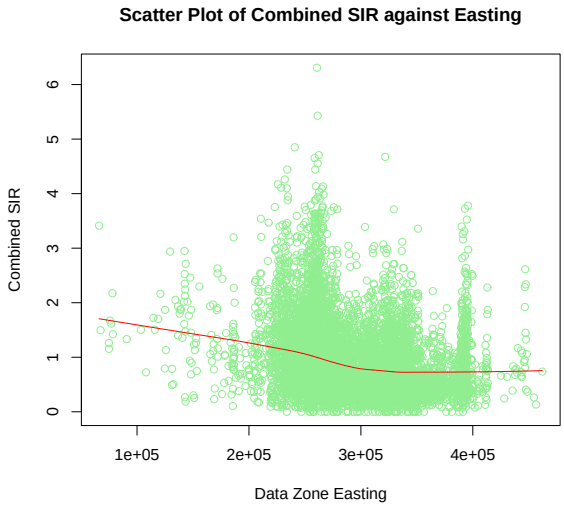
Figure 4.9: Data Zone Map of Stirling Alcohol-related SIR

4.2.1 Male SIR Maps

A data zone map of the male SIR values has been produced for the whole of Scotland (Figure 4.13), along with accompanying magnified sections show-

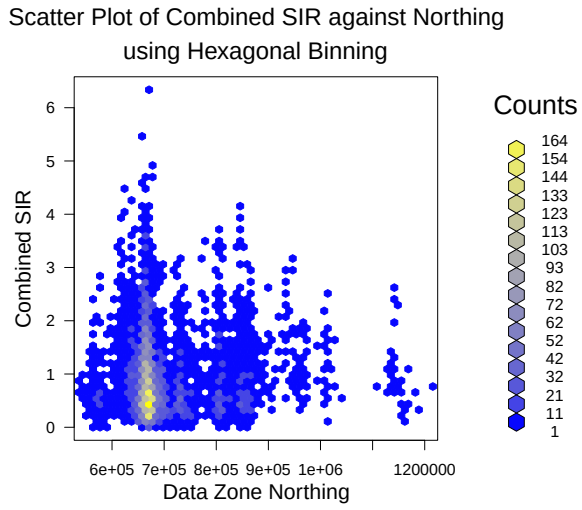


(a) Plot using hexagonal binning

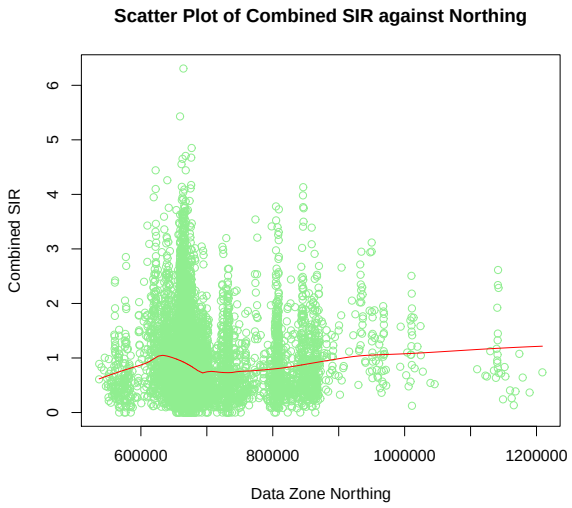


(b) Scatter plot with lowess line

Figure 4.10: Plots of Combined SIR against Easting



(a) Plot using hexagonal binning



(b) Scatter plot with lowess line

Figure 4.11: Plots of Combined SIR against Northing

ing Aberdeen (Figure 4.14), Ayrshire (Figure 4.15), the Dundee area (Figure 4.16), Edinburgh (Figure 4.17), Glasgow (Figure 4.18), the Inverness area (Figure 4.19) and Stirling (Figure 4.20).

The full male SIR map of Scotland, Figure 4.13, exhibits very similar pat-

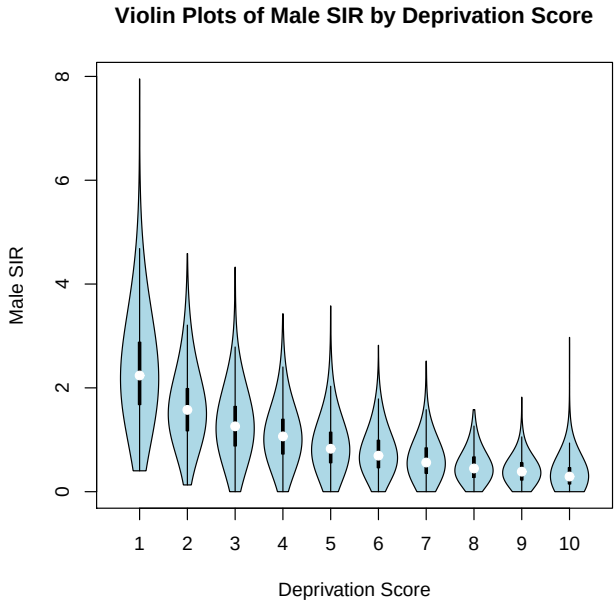


Figure 4.12: Violin Plots of Male SIR by Deprivation Score

terning to both the combined SIR map of Scotland discussed above and the deprivation score map of Scotland shown in Chapter 2. This is further evidence of an association between deprivation score and alcohol-related health risks for males in Scotland.

The full Scottish male SIR map in Figure 4.13, like the combined SIR map, suggests that the values are higher further North and further West in Scotland. Plots similar to those in Figure 4.10 and Figure 4.11 above have been produced for the male SIR values; Figure 4.21 shows two scatter plots of male SIR against data zone centroid easting, one using hexagonal binning and the second with an added lowess line and Figure 4.22 shows the equivalent plots for data zone centroid northings. These plots are extremely similar to those for the combined SIR values. Both plots in Figure 4.21 suggests that the male SIR values do appear to be lower in the East of Scotland, but that any relationship between easting and male alcohol-related risk is likely to be fairly weak. In Figure 4.22 neither plot suggests a particularly strong relationship between male SIR and data zone centroid northing, although,

with the exception of the northing range of around 630000 to 700000, there is slight evidence of the male SIR increasing as you go further north in Scotland.

If we now compare the Glasgow and Edinburgh areas in the male SIR map, shown in Figure 4.18 and Figure 4.17 respectively, they too show very similar patterns to the combined SIR maps for these areas discussed above. For males too the higher SIR values in Edinburgh appear more around the peripheral of the city, where as there are some evident clusters of very high male SIR values just to the East and South of Glasgow city center. In line with the combined results, the male SIRs tend to be much higher on average in Glasgow than in Edinburgh.

The magnified areas showing some smaller cities and towns in Scotland (Figures 4.14 to Figure 4.20) all exhibited patterns so similar to their combined SIR equivaltents that the comments made above still apply. There appears to be no common distribution of high male SIR values throughout the different towns and cities in Scotland. However, the very high values almost always occur in small, densely populated and highly deprived data zones.

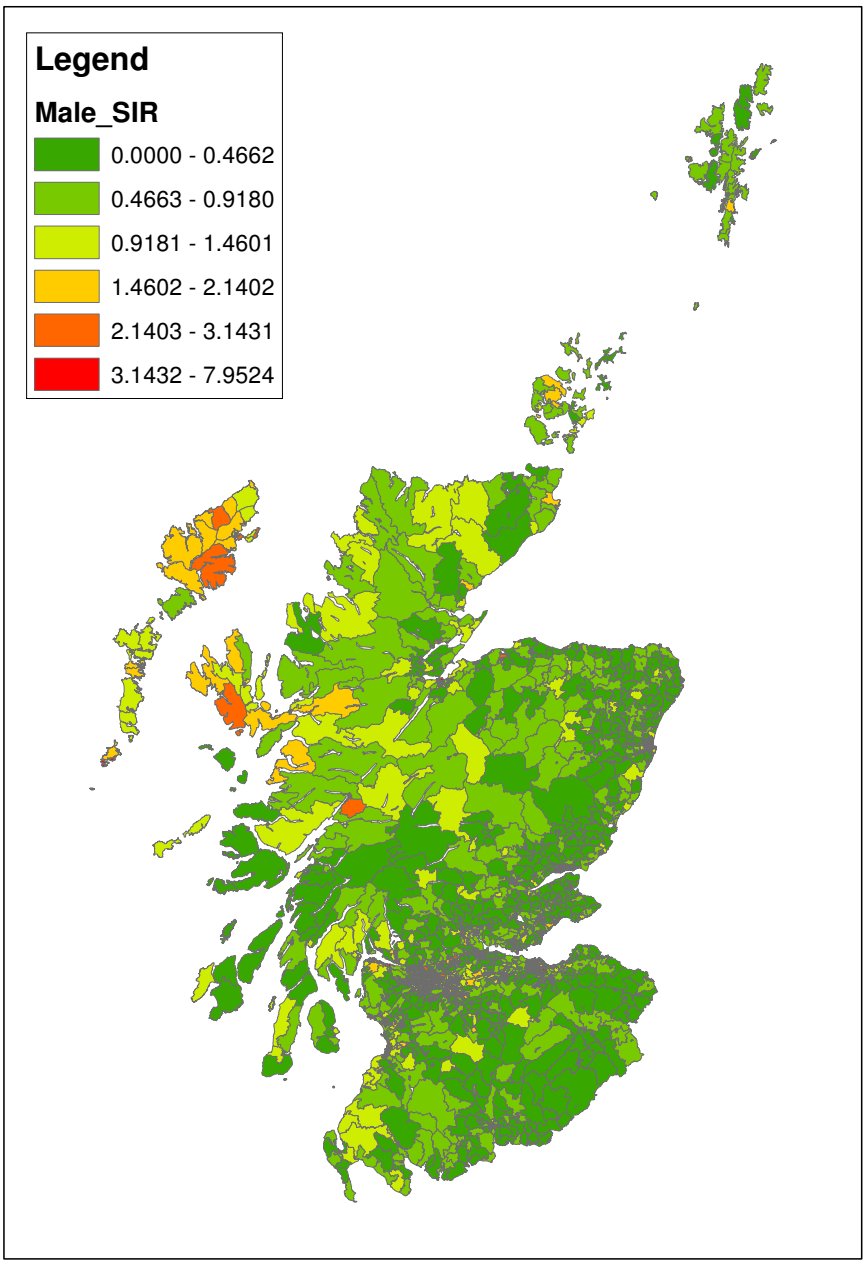


Figure 4.13: Data Zone Map of Male Alcohol-related SIR

Rank	Data Zone	Intermediate Geography	Local Authority	Deprivation	SIR
Sample of lowest SIRs	S01006473	Westfield	West Lothian	7	0
	S01006341	East Calder	West Lothian	8	0
	S01006371	Ladywell	West Lothian	10	0
	S01006486	Linlithgow South	West Lothian	10	0
	S01006154	Dunblane West	Stirling	10	0
	S01006158	Dunblane West	Stirling	10	0
	S01005845	St Leonards South	South Lanarkshire	10	0
	S01005978	Bothwell South	South Lanarkshire	7	0
	S01005224	Paisley West	Renfrewshire	10	0
	S01005154	Renfrewshire Rural South & Howwood	Renfrewshire	8	0
Highest 10 SIRs	S01004388	Irvine Castle Park South	North Ayrshire	1	4.25931
	S01005592	Ayr North Harbour, Wallacetown & Newton South	South Ayrshire	1	4.441144
	S01003582	Possil Park	Glasgow City	1	4.441439
	S01006061	Shawfield & Clincarthill	South Lanarkshire	1	4.553938
	S01003232	Laurieston & Tradeston	Glasgow City	1	4.651525
	S01002296	Muirhouse	City of Edinburgh	1	4.673844
	S01003578	Barmulloch	Glasgow City	1	4.705123
	S01006260	IZ Thirteen	West Dunbartonshire	1	4.850407
	S01003043	Glenwood North	Glasgow City	1	5.429504
	S01003313	Parkhead West & Barrowfield	Glasgow City	1	6.307572

Table 4.1: Combined SIR Table

Rank	Data Zone	Intermediate Geography	Local Authority	Deprivation	SIR
Sample of lowest SIRs	S01000022	Cults, Bieldside & Milltimber East	Aberdeen City	10	0
	S01000045	Braeside, Mannofield, Broomhill & Seafield South	Aberdeen City	10	0
	S01000073	Braeside, Mannofield, Broomhill & Seafield North	Aberdeen City	10	0
	S01000174	Kingswells	Aberdeen City	10	0
	S01000306	Dunecht, Durris & Drumoak	Aberdeenshire	10	0
	S01000330	Crathes & Torphins	Aberdeenshire	8	0
	S01000343	Crathes & Torphins	Aberdeenshire	9	0
	S01000357	Westhill North & South	Aberdeenshire	10	0
	S01000367	Kintore & Blackburn	Aberdeenshire	9	0
	S01000389	Newmachar & Fintray	Aberdeenshire	9	0
Highest 10 SIRs	S01006061	Shawfield & Clincarthill	South Lanarkshire	1	4.719262
	S01003159	Toryglen & Oatlands	Glasgow City	1	4.749119
	S01003578	Barmulloch	Glasgow City	1	4.761058
	S01005315	Renfrew West	Renfrewshire	1	4.831424
	S01003296	Parkhead West & Barrowfield	Glasgow City	1	4.839941
	S01003217	Dalmarnock	Glasgow City	1	5.092317
	S01003299	Parkhead East & Braidfauld North	Glasgow City	1	5.407188
	S01006260	IZ Thirteen	West Dunbartonshire	1	5.502008
	S01003043	Glenwood North	Glasgow City	1	7.010165
	S01003313	Parkhead West & Barrowfield	Glasgow City	1	7.952416

Table 4.2: Male SIR Table

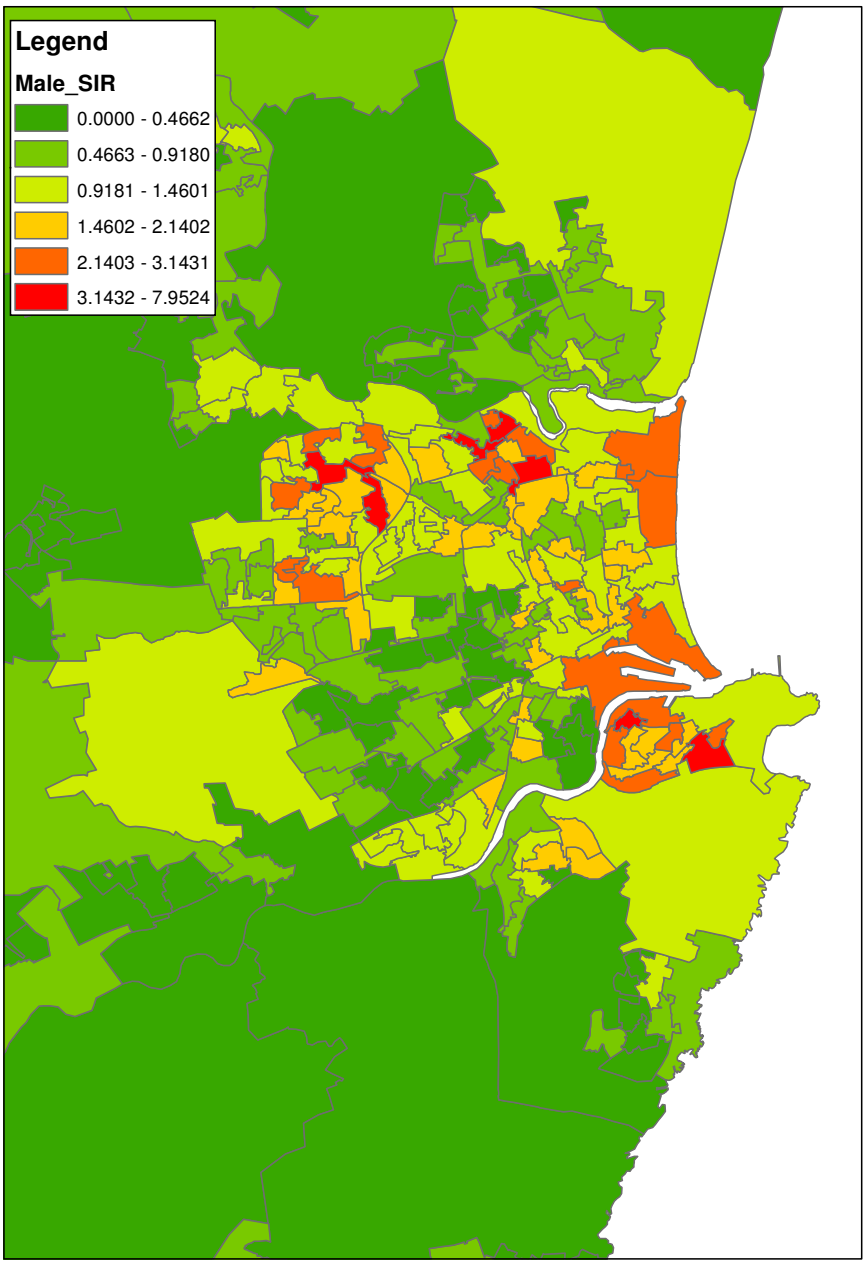


Figure 4.14: Data Zone Map of Aberdeen Male Alcohol-related SIR

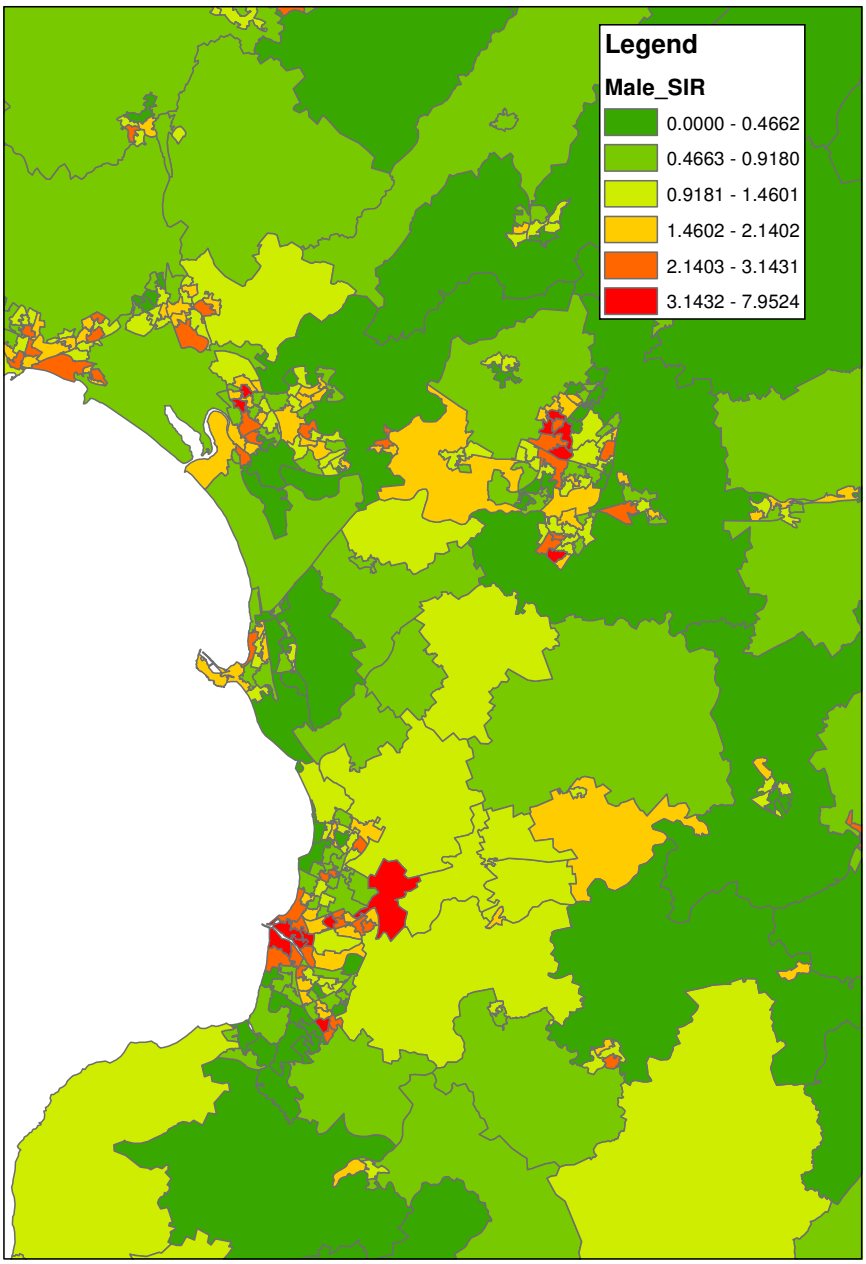


Figure 4.15: Data Zone Map of Ayrshire Male Alcohol-related SIR

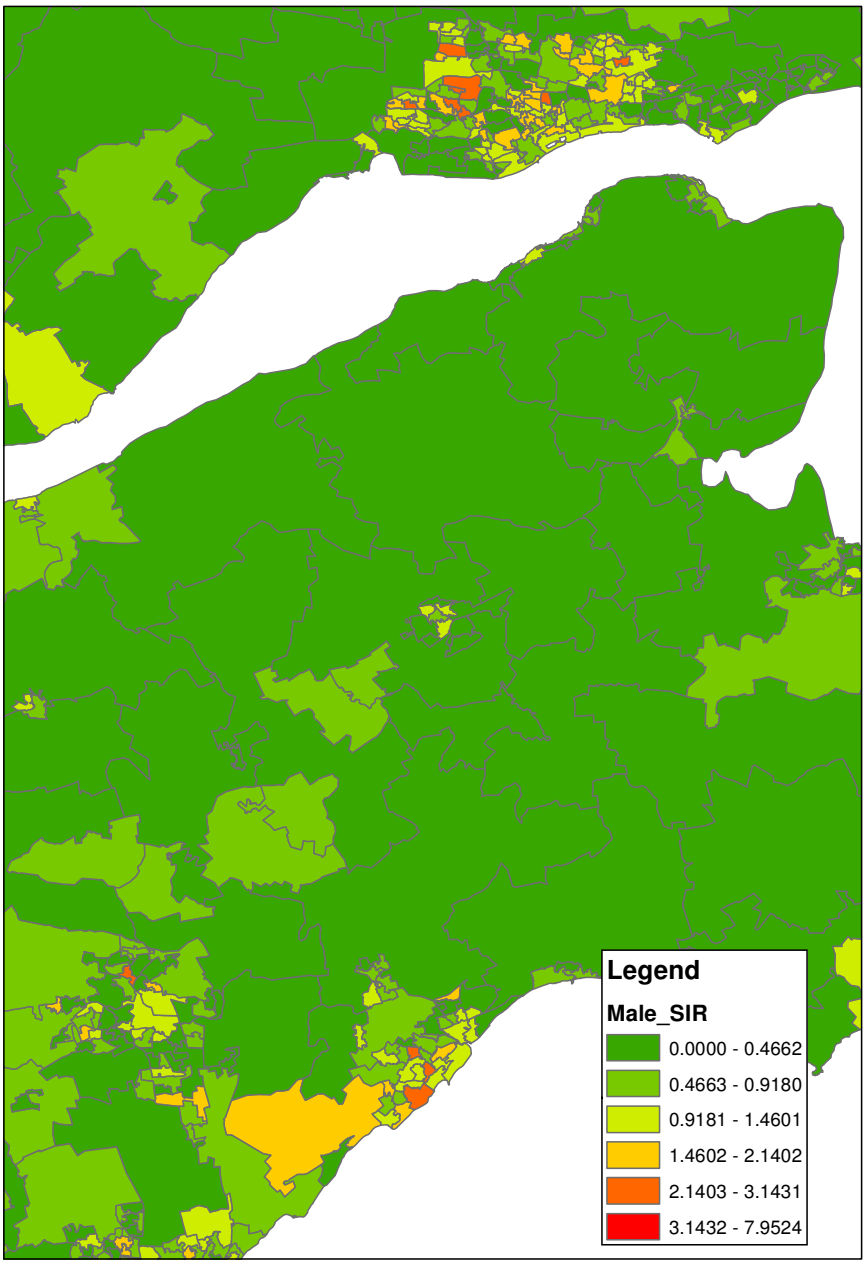


Figure 4.16: Data Zone Map of Fife Male Alcohol-related SIR

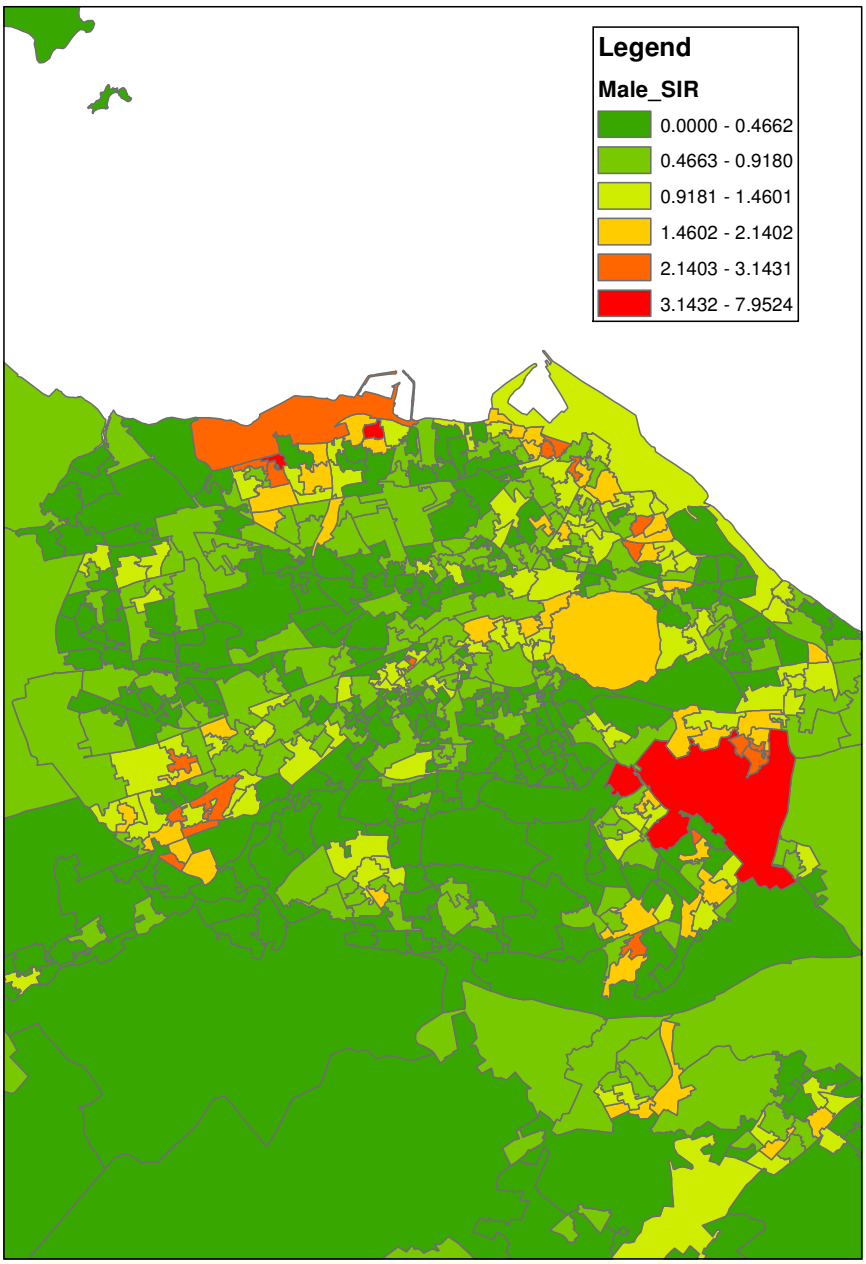


Figure 4.17: Data Zone Map of Edinburgh Male Alcohol-related SIR

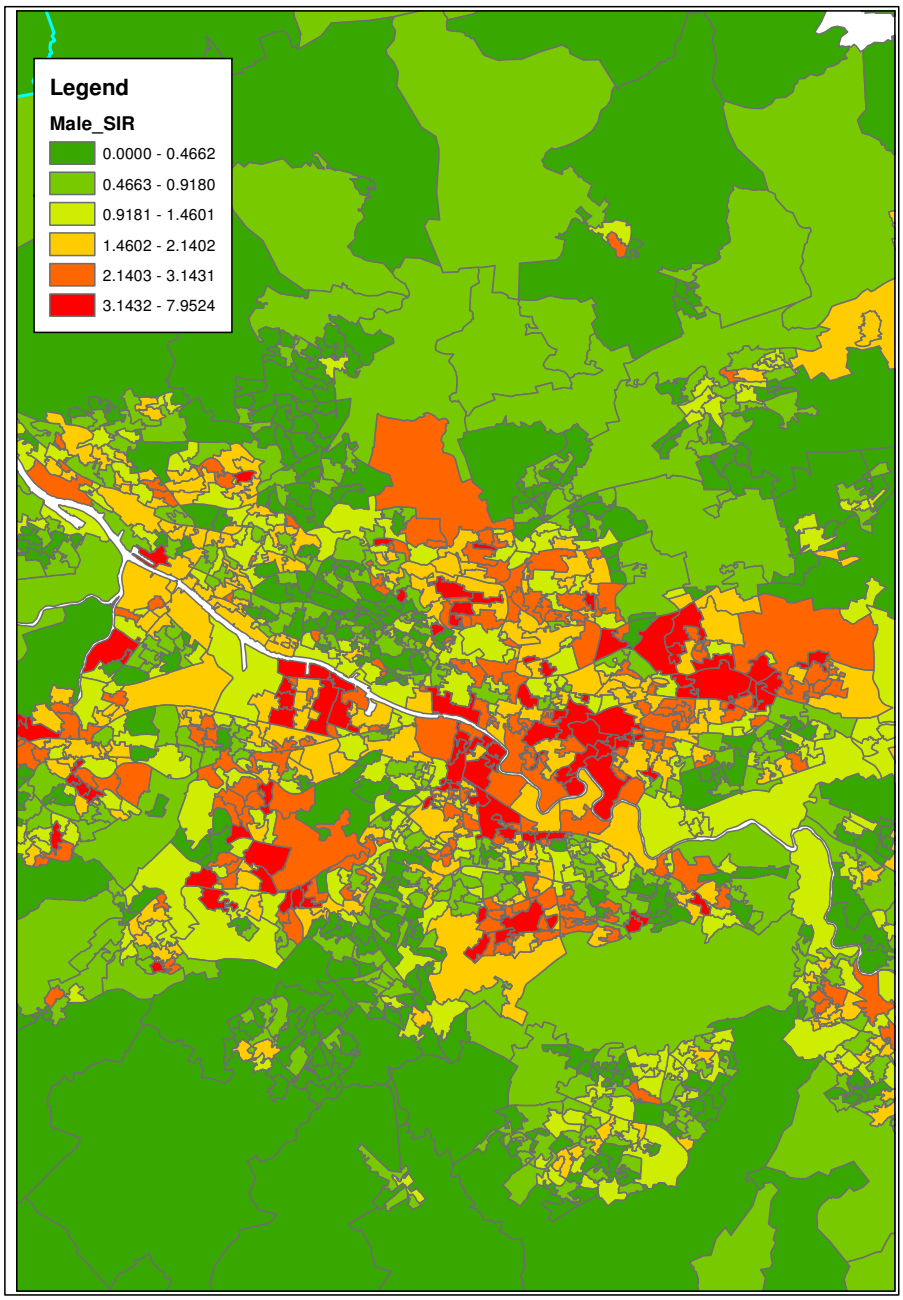


Figure 4.18: Data Zone Map of Glasgow Male Alcohol-related SIR

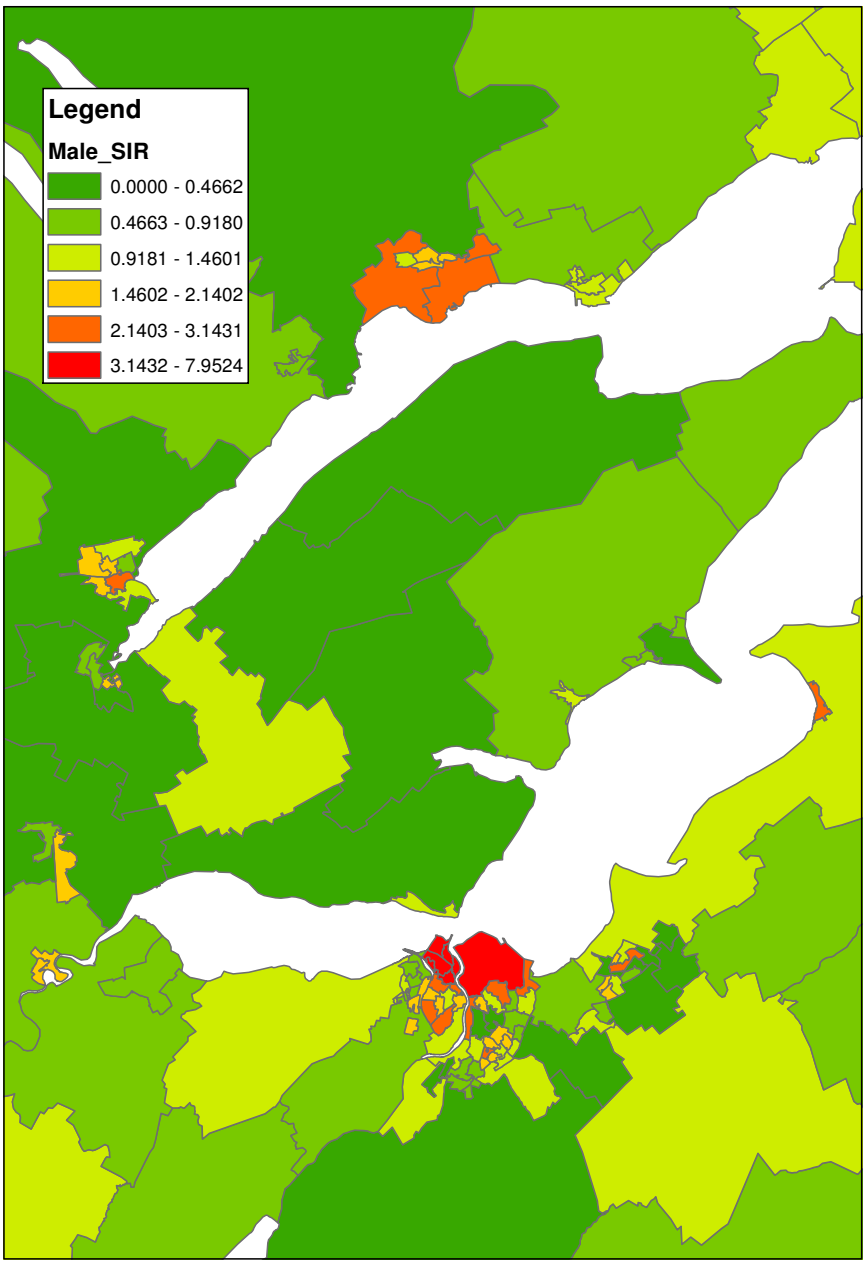


Figure 4.19: Data Zone Map of Inverness Male Alcohol-related SIR

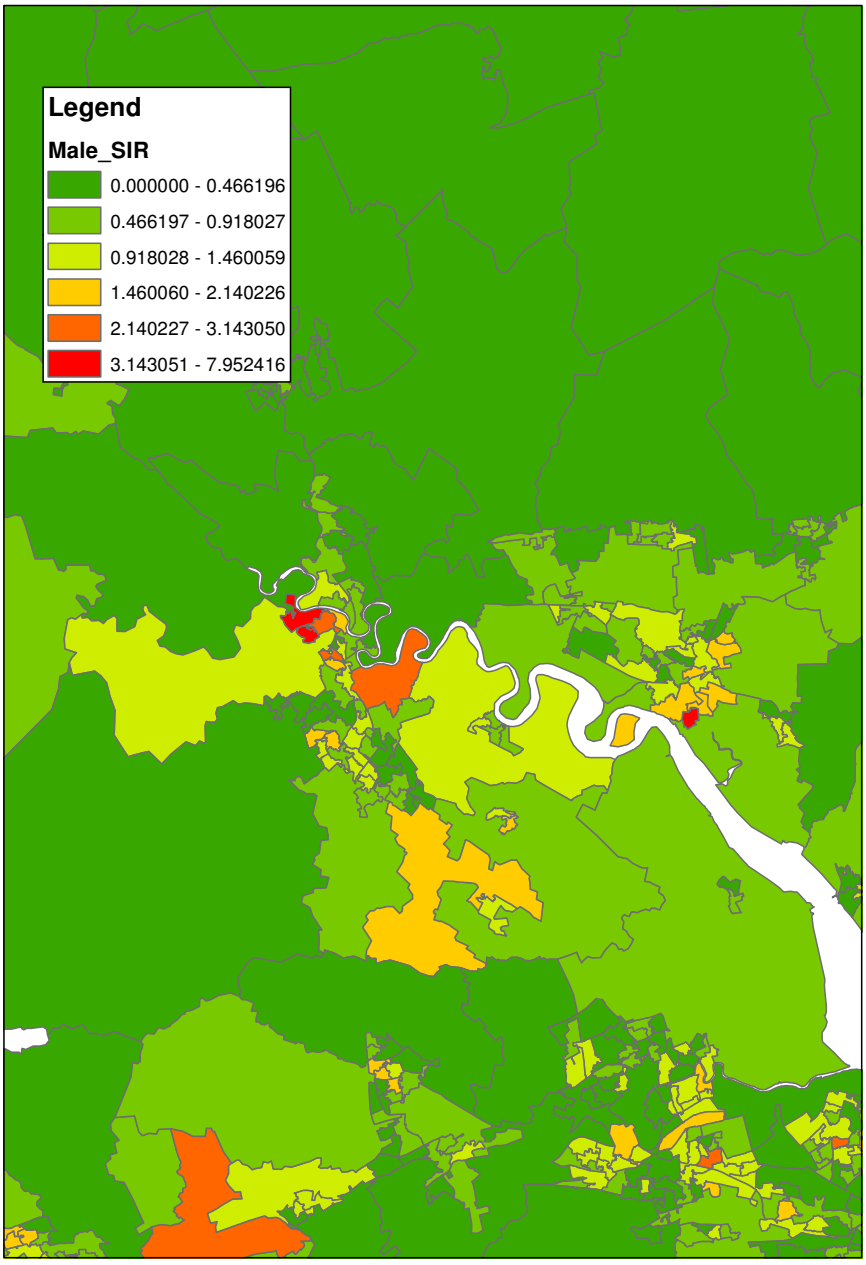
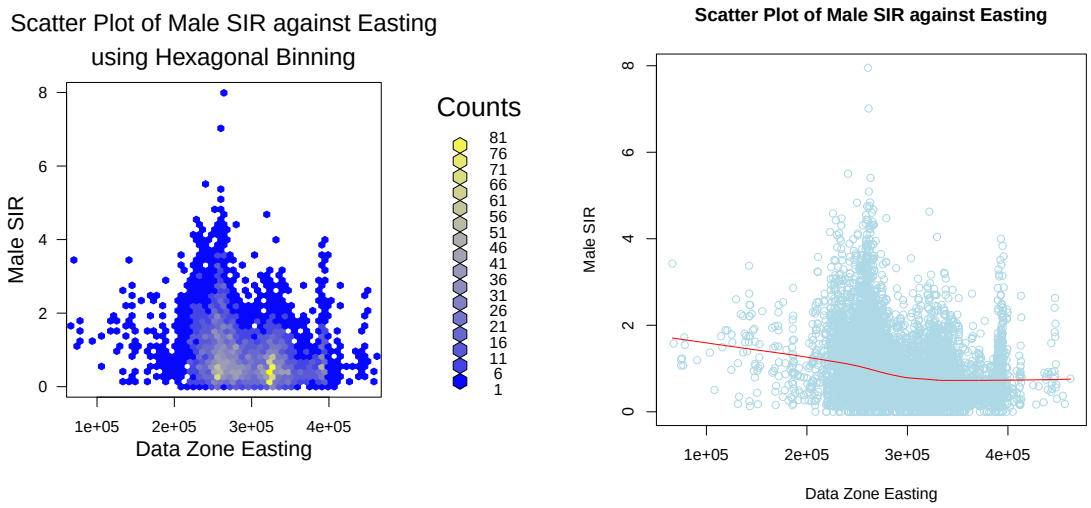


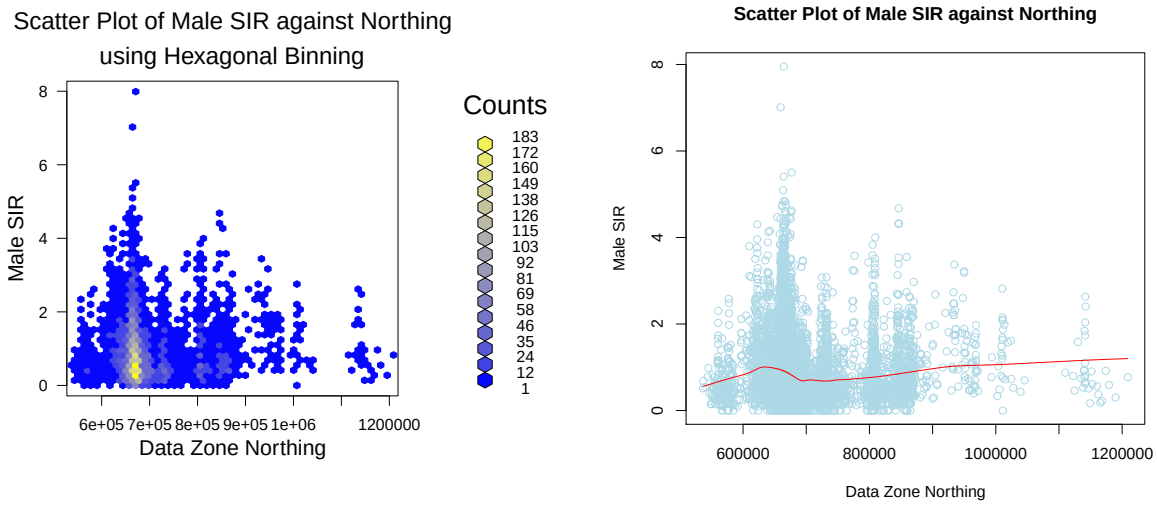
Figure 4.20: Data Zone Map of Stirling Male Alcohol-related SIR



(a) Plot using hexagonal binning

(b) Scatter plot with lowess line

Figure 4.21: Plots of Male SIR against Easting



(a) Plot using hexagonal binning

(b) Scatter plot with lowess line

Figure 4.22: Plots of Male SIR against Northing

4.3 Female SIR

The SIRs calculated using only female data will now be discussed. Like the previous sections, we will start by looking at a table which gives the data zones with the ten highest and 10 lowest female SIRs values in Scotland,

Table 4.3.

Obviously again, if the ten lowest SIR values for the combined data are zero, this must also be the case for the female only data, which is confirmed in Table 4.3. In total there are 761 data zones with a female SIR value of zero. This is considerably more than the male count of just 196. This agrees further with previous research which has shown alcohol abuse to be much greater among males in Scotland than among females. Of the 796 data zones which experience a zero female SIR value, 26.4% have a deprivation score of 10 and more than 16.2% have a deprivation score of 5 or less. This is in contrast to the male SIR results, where over 39% of the zero values were for areas with the least deprived score of 10 and only 9.8% had a score of 7 or less. This suggests that deprivation score may share a greater association with male SIR than with female SIR.

Of the 10 highest female SIR values shown in Table 4.3 7 represent data zones with a deprivation score of 1 and 3 with a score of 2. All of the 10 highest male SIR values were in areas with a deprivation score of 1, so it is of interest to compare the figures for the 100 highest SIR values for males and females. Only 68% of the 100 highest female SIR values correspond to areas with a deprivation score of 1 compared to 86% for males. This adds to the suggestion that deprivation score may have a stronger association with male alcohol-related risk than with female alcohol-related risk. Of the 10 highest female SIR values just three correspond to Glasgow City compared to 8 for males. As a first look this suggests that clustering may be stronger for males, although this is weak evidence and clustering is much better judged by looking at maps of SIR values.

A violin plot of SIR values by deprivation score has also been produced for the female data and is shown below in Figure 4.23. This plot shows that female SIR values tend to be higher in more deprived areas. As the deprivation score worsens from 10 to 2 the median female SIR value appears to increase in a roughly linear fashion. However, there is then a relatively

large jump in median female SIR value moving from deprivation score 2 to 1. It appears, that female SIR values during this period are linked to deprivation score, but probably to a lesser extent than for males. Both the data for males and for females suggest that any relationship between these variables may not be linear; so far a linear relationship between deprivation score and SIR seems less likely for males.

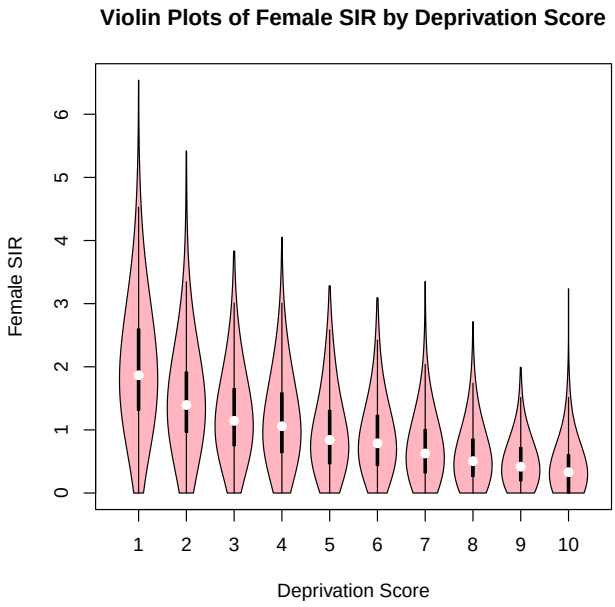


Figure 4.23: Violin Plots of Female SIR by Deprivation Score

Rank	Data Zone	Intermediate Geography	Local Authority	Deprivation	SIR
Sample of lowest SIRs	S01000022	Cults, Bieldside & Milltimber East	Aberdeen City	10	0
	S01000073	Braeside, Mannofield, Broomhill & Seafield North	Aberdeen City	10	0
	S01000409	Inverurie North	Aberdeenshire	10	0
	S01000468	Fyvie-Rothie	Aberdeenshire	8	0
	S01000590	Carnoustie East	Angus	8	0
	S01000603	Monikie	Angus	9	0
	S01000632	Letham & Glamis	Angus	7	0
	S01000662	Letham & Glamis	Angus	8	0
	S01000892	Dollar & Muckhart	Clackmannanshire	10	0
	S01000981	Gretna & Eastriggs	Dumfries & Galloway	9	0
Highest 10 SIRs	S01005592	Ayr North Harbour, Wallacetown & Newton South	South Ayrshire	1	4.728708
	S01005559	Castlehill & Kincaidston	South Ayrshire	2	4.764768
	S01002296	Muirhouse	Edinburgh, City of	1	4.770263
	S01003302	Laurieston & Tradeston	Glasgow City	1	4.829221
	S01005425	Langlee	Scottish Borders	1	4.874045
	S01003202	Toryglen & Oatlands	Glasgow City	1	4.879341
	S01005598	Ayr North Harbour, Wallacetown & Newton South	South Ayrshire	1	4.919099
	S01004055	Bow Farm, Barrs Cottage, Cowdenknowes & Overton	Inverclyde	2	5.143859
	S01000747	Dunoon	Argyll & Bute	2	5.418210
	S01003232	Laurieston & Tradeston	Glasgow City	1	6.539285

Table 4.3: Female SIR Table

4.3.1 Female SIR Maps

A data zone map of Scotland depicting the female SIR values is shown in Figure 4.24, along with magnified sections of this map for Aberdeen (Figure 4.25), Ayrshire (Figure 4.26), the Dundee area (Figure 4.27), Edinburgh (Figure 4.28), Glasgow (Figure 4.29), the Inverness area (Figure 4.30) and Stirling (Figure 4.31).

On a first glance at the female SIR map of Scotland in Figure 4.24 it appears to be less smooth than its male equivalent. In general alcohol-related health risks appear to be higher in the South and East of the country, but possibly less so than for the males.

Two plots of female SIR against data zone centroid easting are shown in Figure 4.32, the first using hexagonal binning, and the second with a superimposed lowess line. In fact these plots show a very similar association between SIR value and easting to that exhibited by the male values. In general there appears to be a decrease in female SIR value as you move from West to East up to around $3e^5$ where it begins to settle. Two similar plots were produced showing female SIR against data zone centroid Northing, shown in Figure 4.33. Again these plots also show an extremely similar pattern to that of the male data; the SIR values tend to increase from South to North, with the exception of a small region which lies between the Northing values of 600,000 and 700,000. There appears to be a large amount of variation around the lowess line, and neither relationship appears to be very strong. This suggests that it may not be worth factoring Easting and Northing into the modelling process, especially since other spatial methods will be explored.

Looking further at Figure 4.24 it can be seen that there are some data zones which have a female SIR of between 1.99 and 3 which also have a male SIR of less than 0.47. There are also some areas which have a lower female SIR than the male equivalent. The female map does exhibit an overall pattern which is similar to, but much less smooth than, that of the males. This relative lack of smoothness may indicate that female alcohol-related

risk varies less smoothly across Scotland. Alternatively, it may simply be due to the fact that alcohol-related deaths and hospitalisations are much less common among women, resulting in lower numbers of observed cases and more erratic/ less reliable risk estimates.

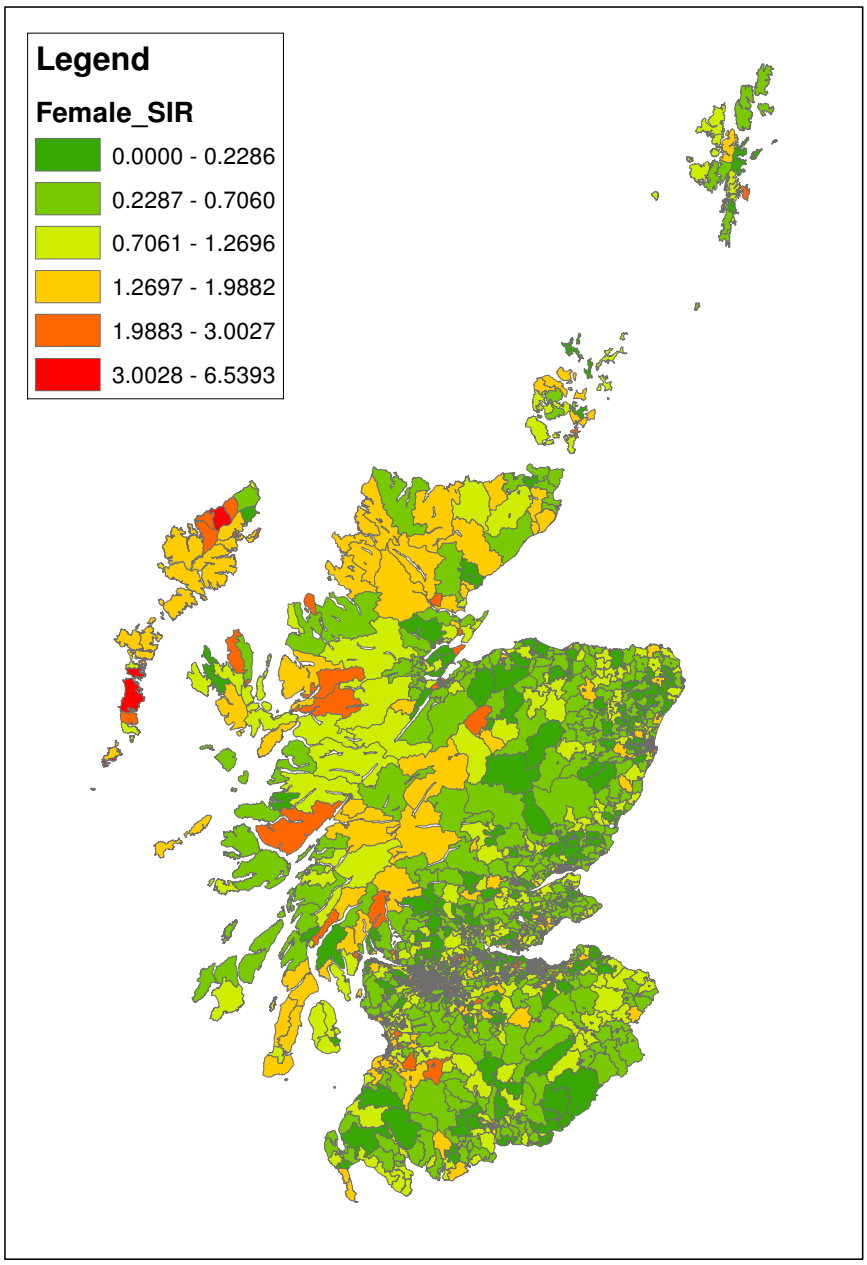


Figure 4.24: Data Zone Map of Female Alcohol-Related SIR

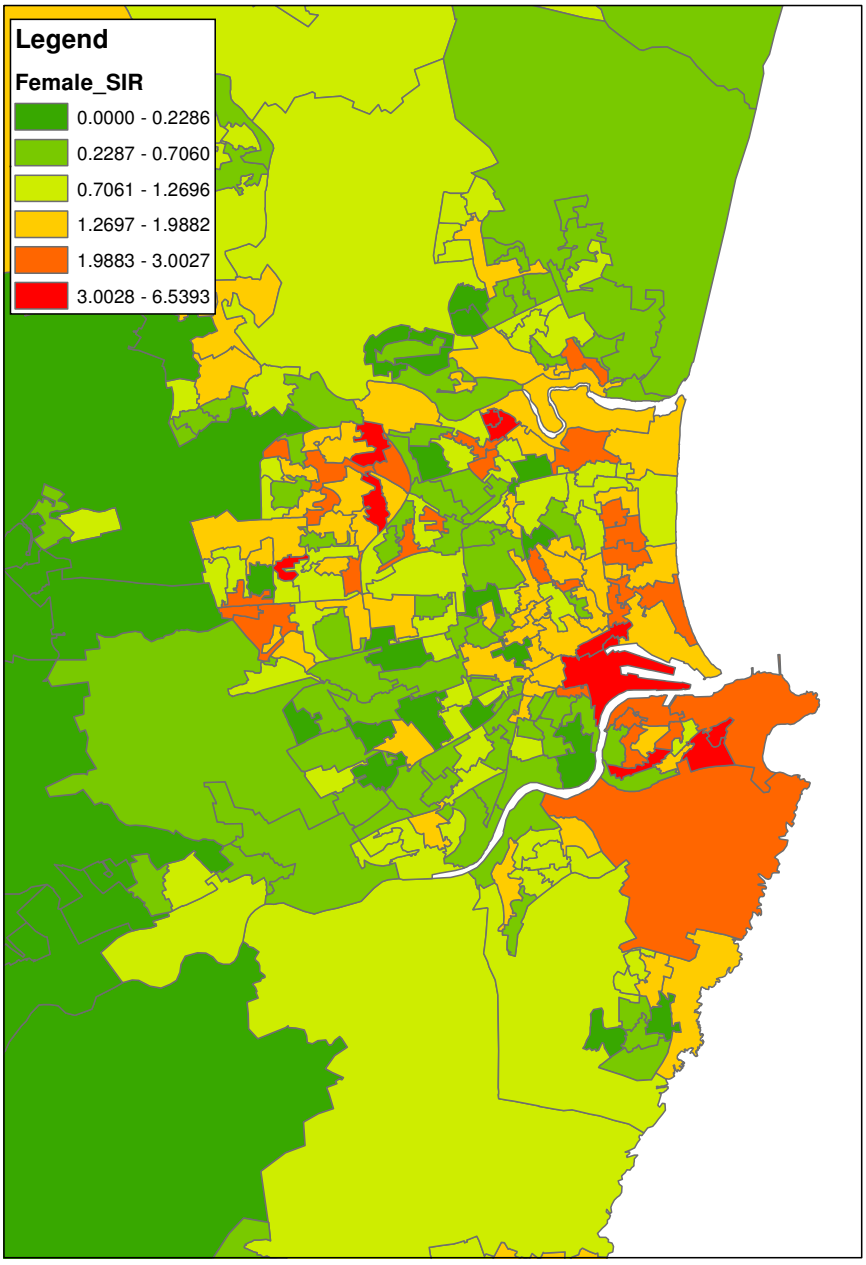


Figure 4.25: Data Zone Map of Aberdeen Female Alcohol-Related SIR

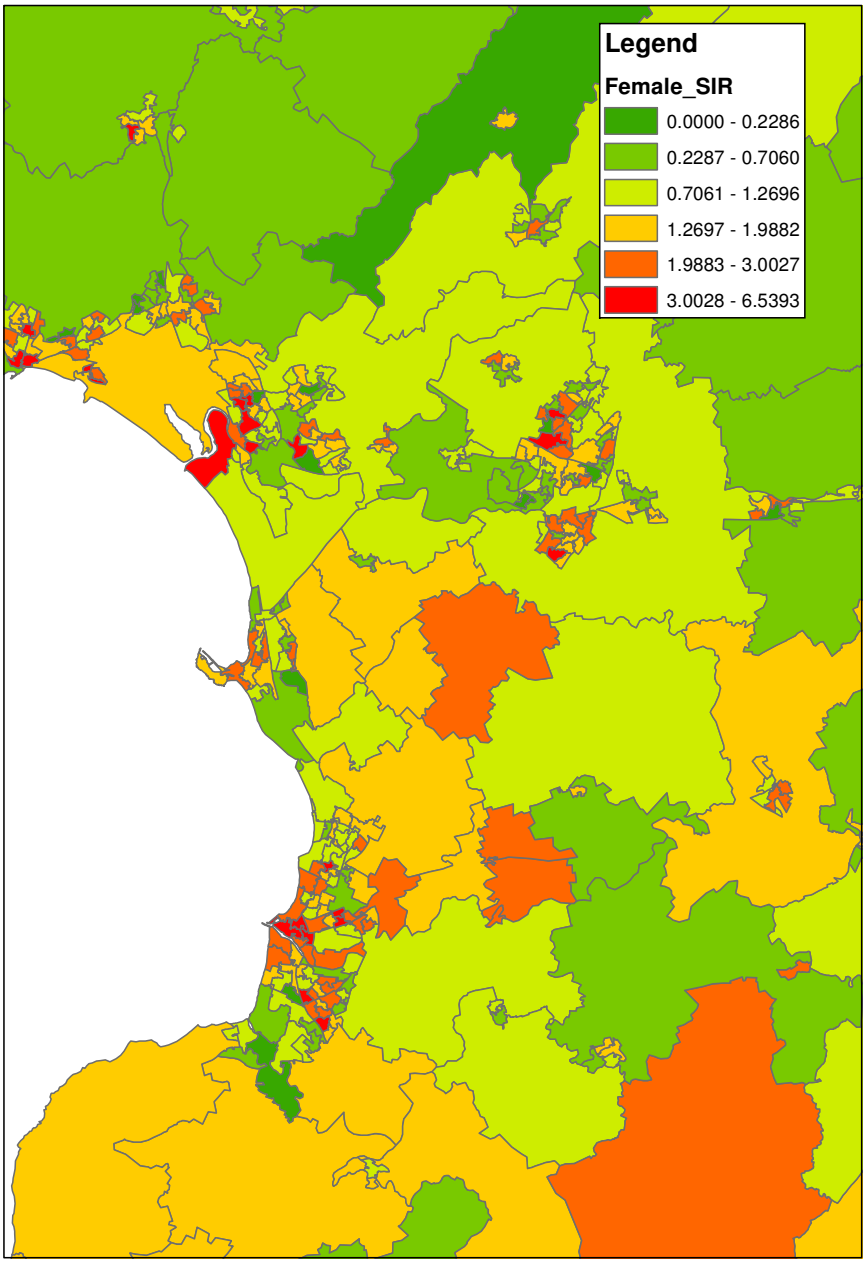


Figure 4.26: Data Zone Map of Ayrshire Female Alcohol-Related SIR

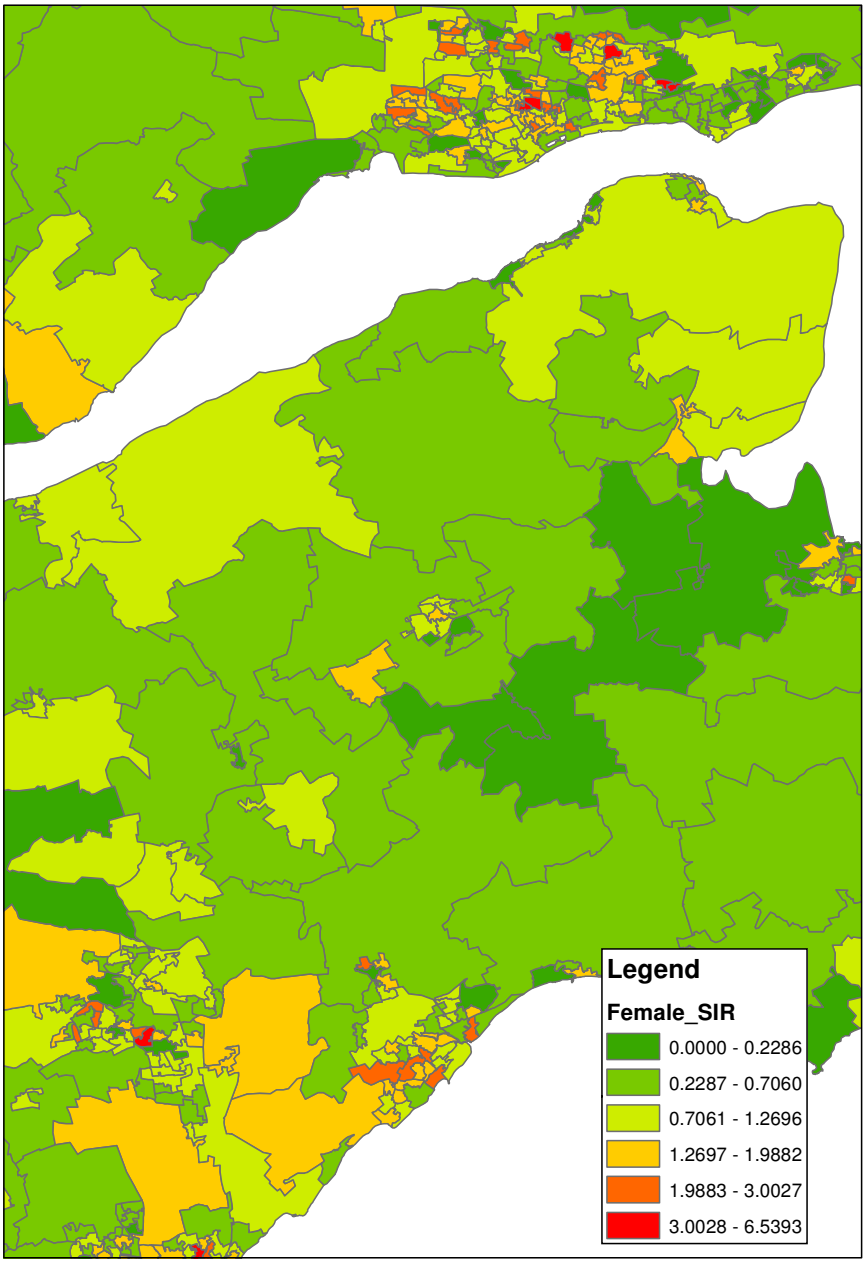


Figure 4.27: Data Zone Map of Fife Female Alcohol-Related SIR

4.4 Comparison of Male and Female SIR Values

Comparing the zoomed-in areas of the female SIR map of Scotland, Figure 4.24, with the male equivalents it can be seen that they all exhibit similar

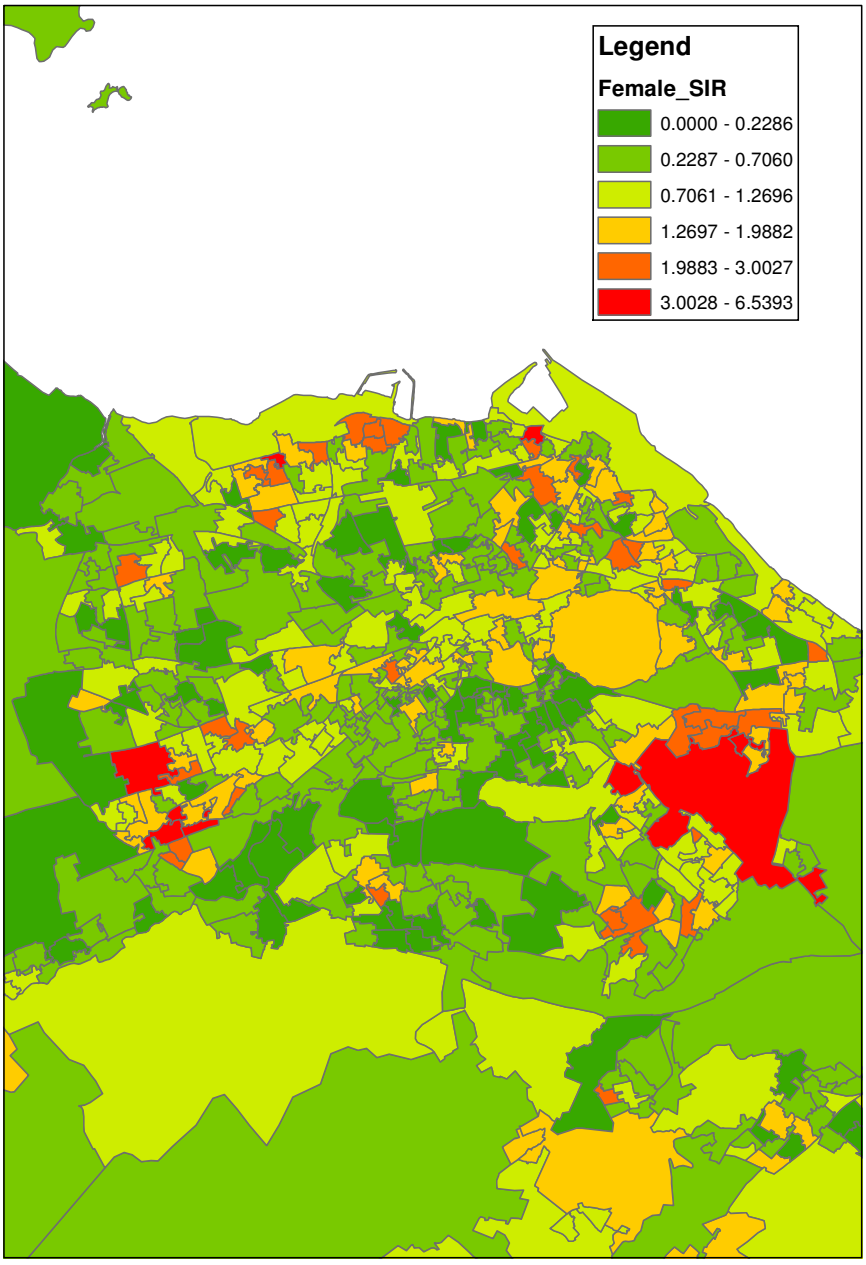


Figure 4.28: Data Zone Map of Edinburgh Female Alcohol-Related SIR

patterns, but that the female values appear to be less smooth than the males in each. However, it should be noted that the SIR map risk level colour definitions are different for the two sexes. In order to make it simpler to compare the estimated level of alcohol-related health risk across Scotland

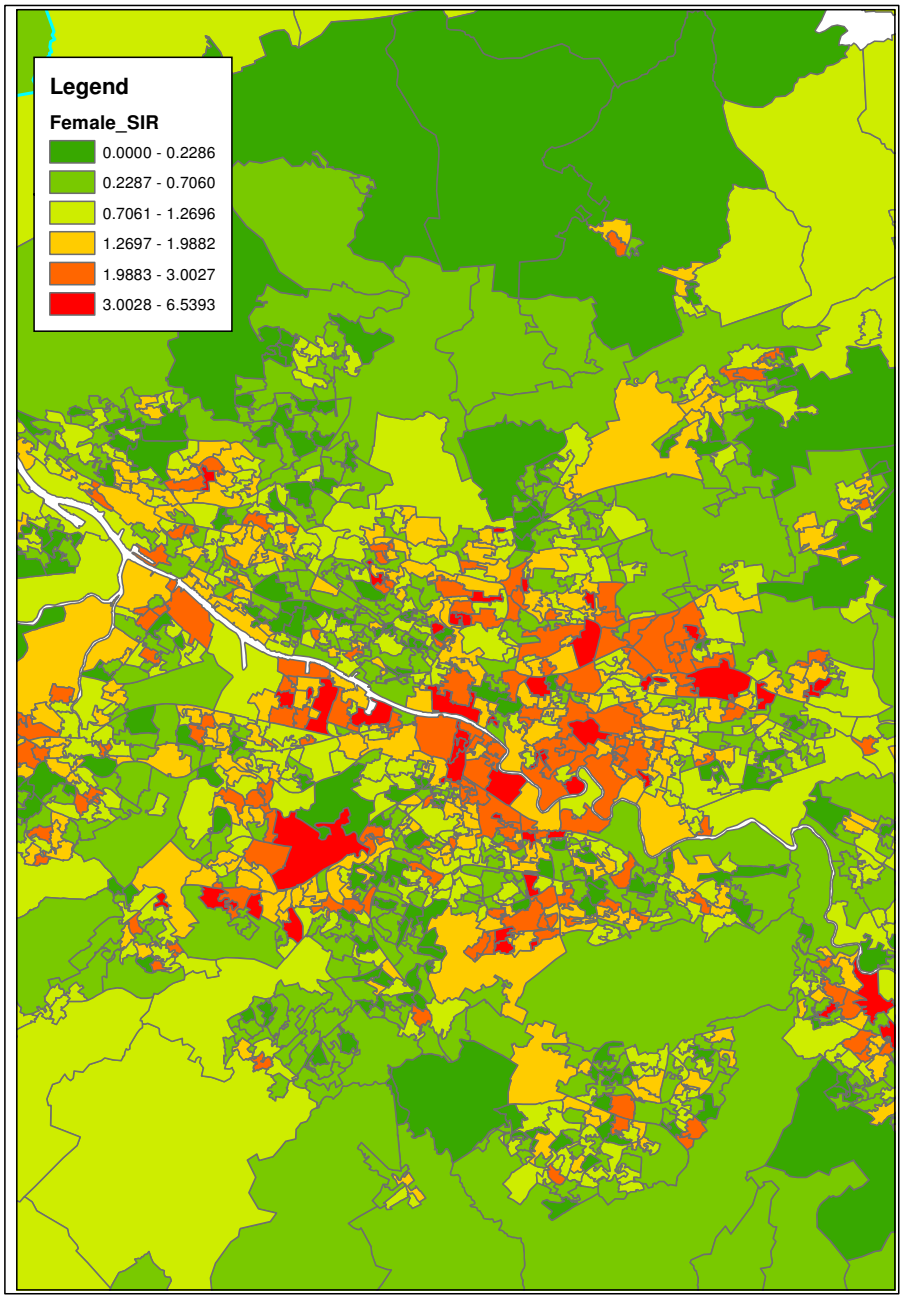


Figure 4.29: Data Zone Map of Glasgow Female Alcohol-Related SIR

between males and females some further plots have been produced. The ratio of female to male SIR in each of the datazones in Scotland has been computed and maps showing these values have been created for the following areas: Scotland (Figure 4.35), Aberdeen (Figure 4.36), Ayrshire (Figure

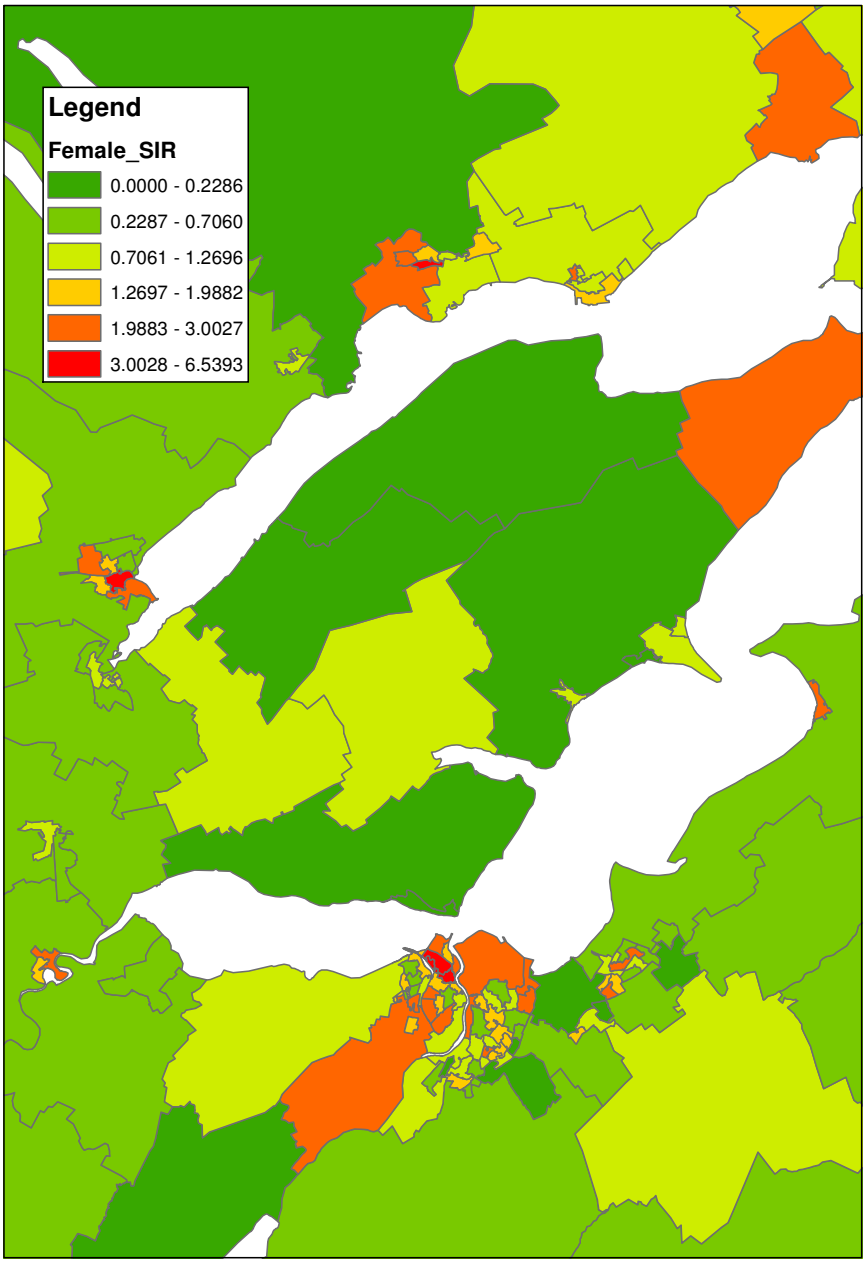


Figure 4.30: Data Zone Map of Inverness Female Alcohol-Related SIR

4.37), Dundee (Figure 4.38), Edinburgh (Figure 4.39), Inverness (Figure 4.41) and Stirling (Figure 4.42).

It should be noted that there are 133 instances where the male alcohol-related SIR is zero but the female equivalent for a data zone is positive. It is

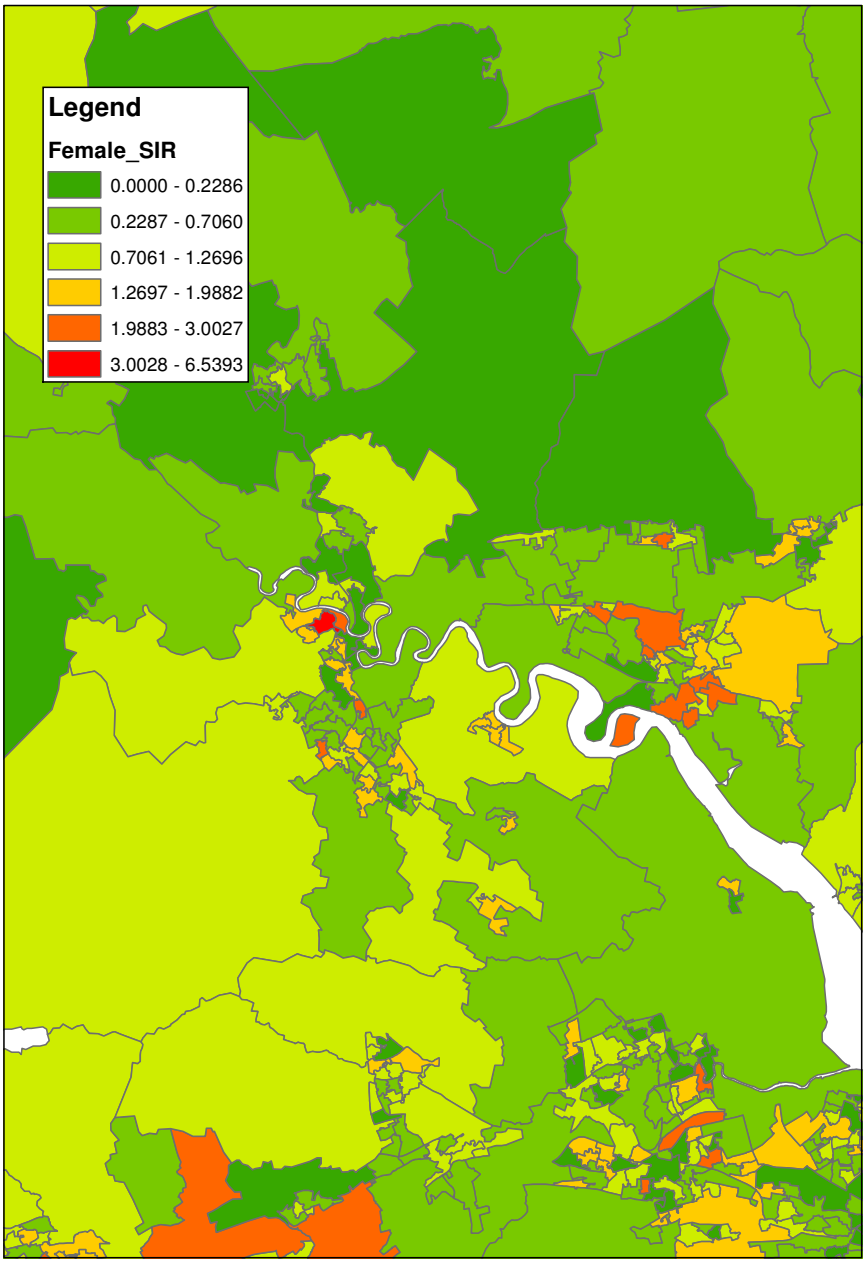
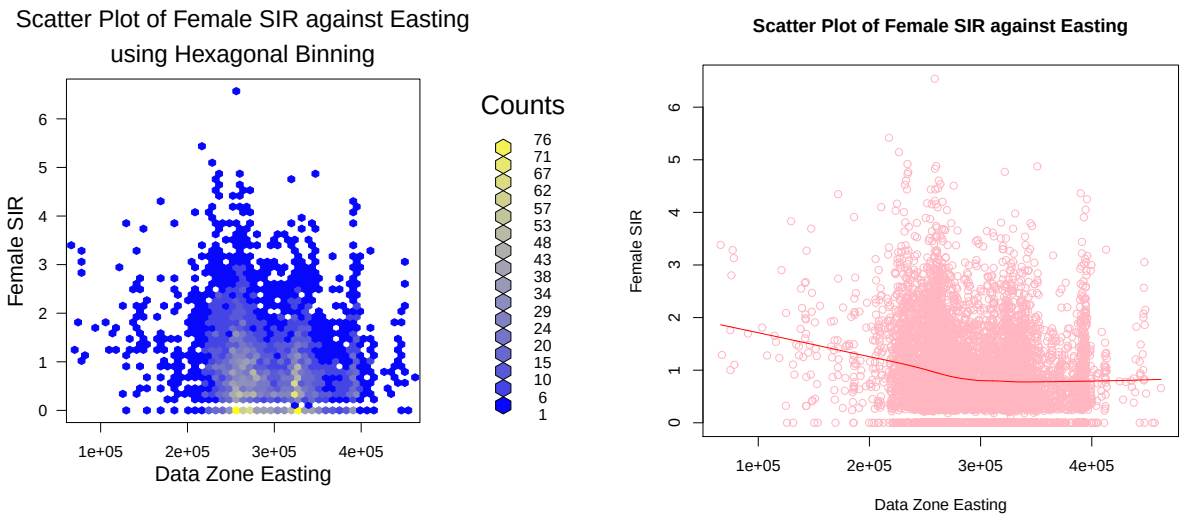


Figure 4.31: Data Zone Map of Stirling Female Alcohol-Related SIR

therefore impossible to estimate a ratio of female to male SIR since it would involve dividing by zero. Such data zones are shaded white in the following ratio maps.

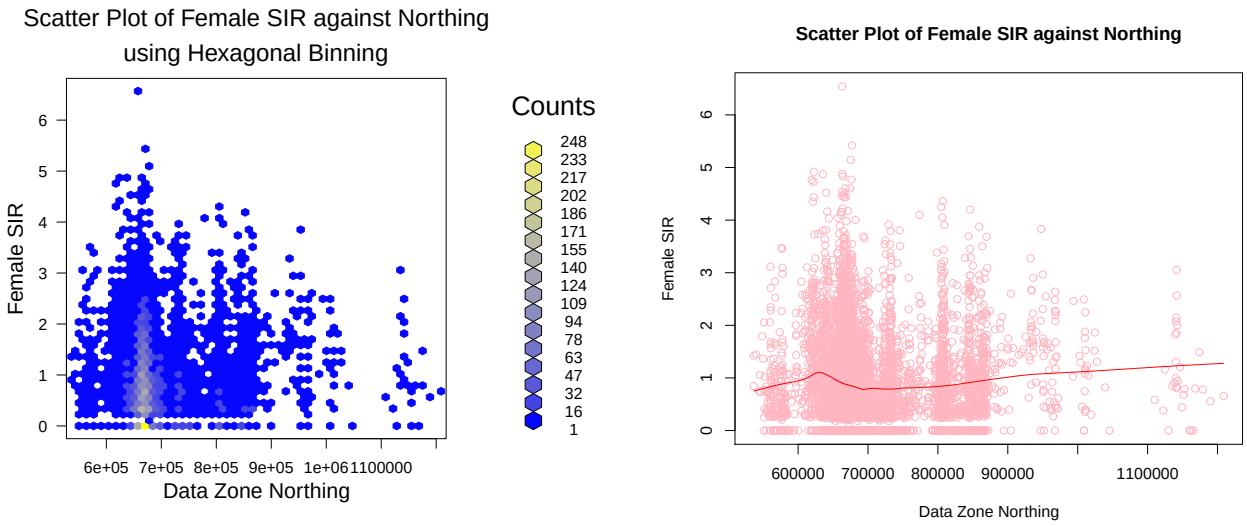
It was found that in just over half of the data zones in Scotland the



(a) Plot using hexagonal binning

(b) Scatter plot with lowess line

Figure 4.32: Plots of Female SIR against Easting



(a) Plot using hexagonal binning

(b) Scatter plot with lowess line

Figure 4.33: Plots of Female SIR against Northing

ratio of female to male alcohol-related SIR is less than 1. In fact, the actual percentage is approximately 50.67% which is around what one would expect if the spatial pattern of alcohol-related risk is the same in both sub-populations.

There were several areas where the female alcohol-related relative risk

was unusually high compared to that for males, with a ratio of greater than 10 in 25 of the data zones. The maximum ratio observed was 21.83 and this was observed for data zone S01001315 which is an area of Mauchline in East Ayrshire. Boxplots of the ratio of female to male SIR values split by deprivation score are shown in below in Figure 4.34; this shows that the median ratio of around 0.93 is very similar across the 10 deprivation scores, but that there is a larger variation in the ratio in less deprived areas.

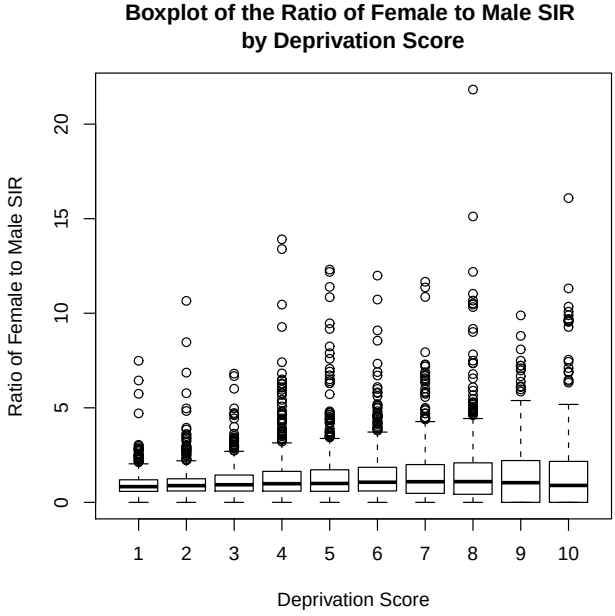


Figure 4.34: Boxplots of the Ratio of Female to Male SIR by Deprivation Score

Firstly, by looking at Figure 4.35 it is apparent that there is no obvious patterns or trends in the ratio of female to male SIR values. There is evidence of data zones with very high ratios both in large rural areas, island areas and in small inner city areas. The majority of the ratios above 3.9541 appear to fall in the central belt of Scotland.

Further, the enlarged Edinburgh area from the ratio map (Figure 4.39) exhibits much higher female to male SIR ratios on the whole than in the equivalent map of Glasgow (Figure 4.40). This ties in with the above com-

ments relating to the variation in ratios differing according to deprivation levels, since from earlier maps it is clear that on average deprivation is much lower in the the Edinburgh area than in Glasgow.

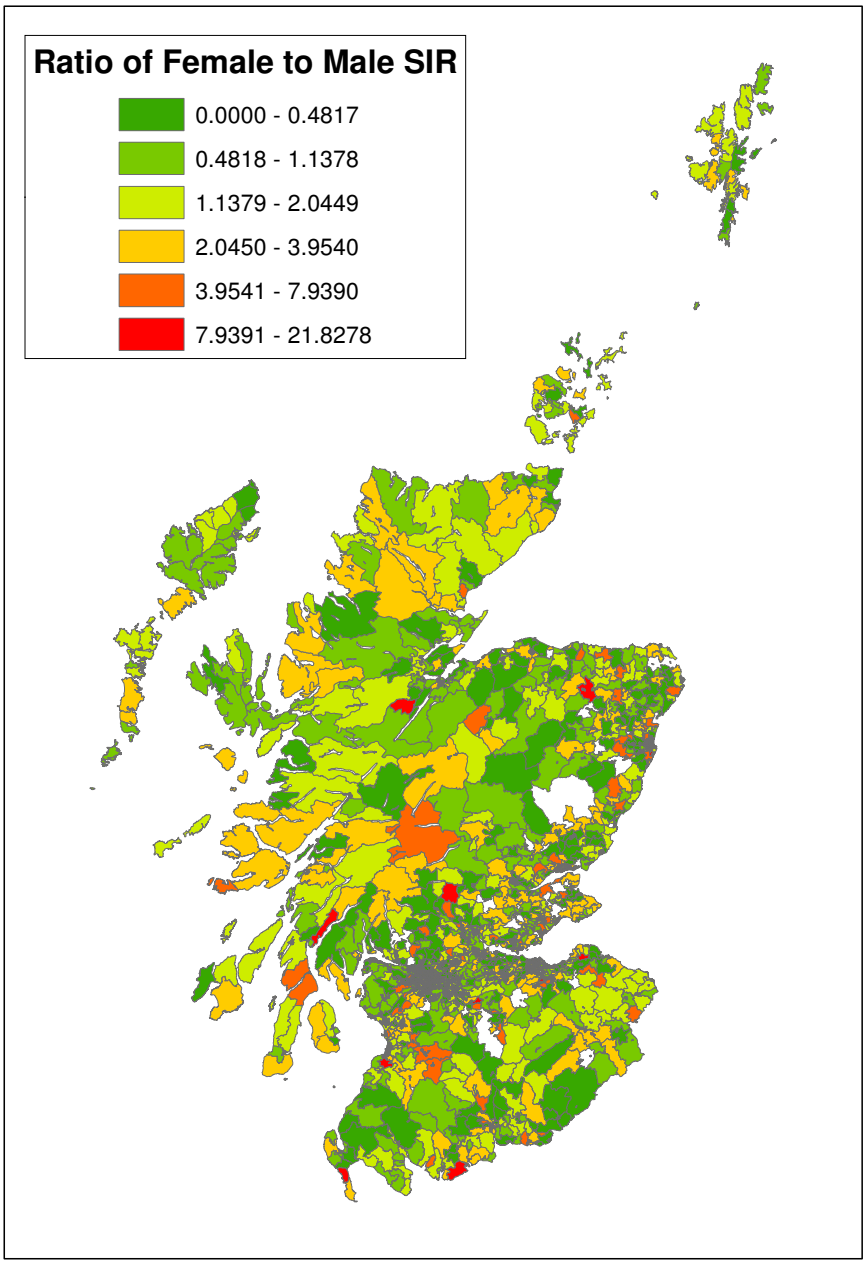


Figure 4.35: Data Zone Map of the Ratio of Female to Male SIR in Scotland Area

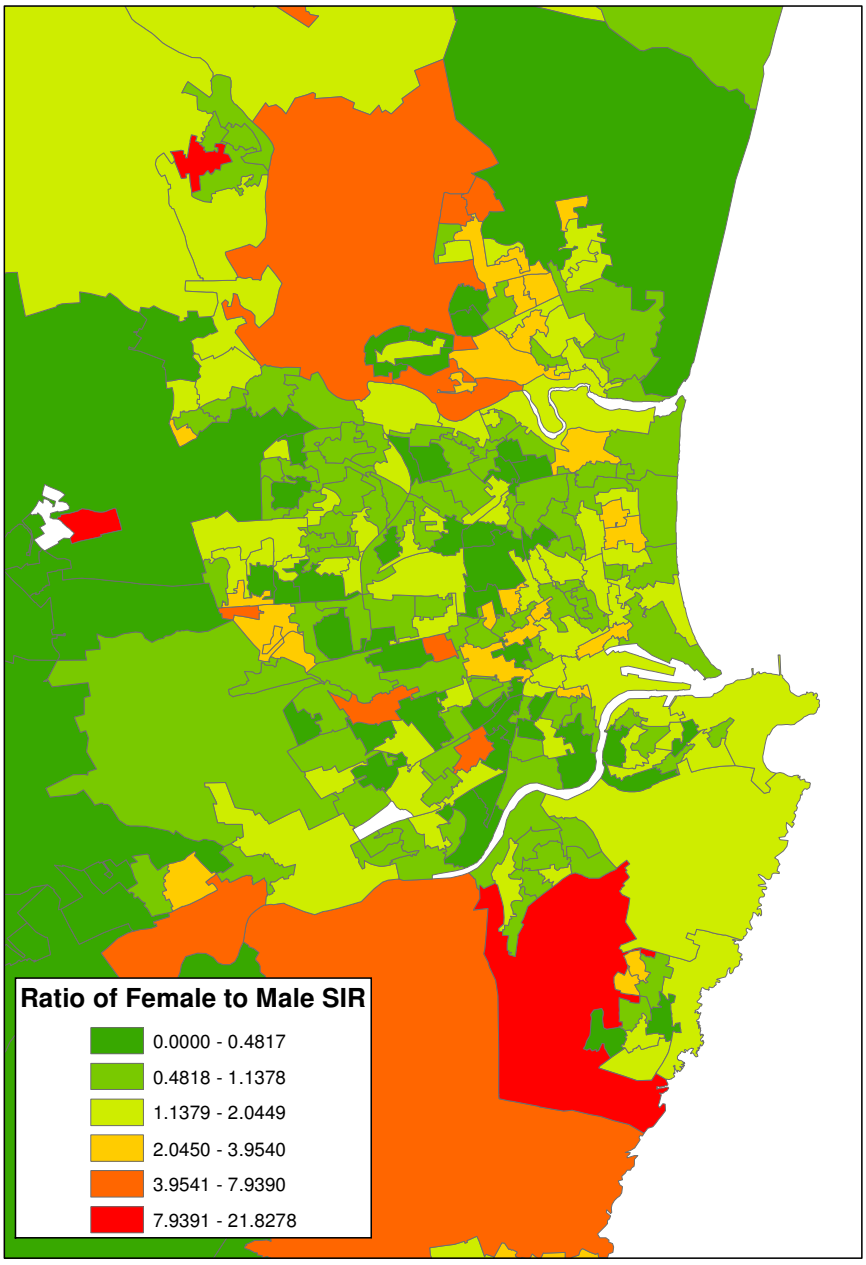


Figure 4.36: Data Zone Map of the Ratio of Female to Male SIR in the Aberdeen Area

Improving this lack of smoothness, along with improving the reliability of the estimates is the aim of the modelling process.

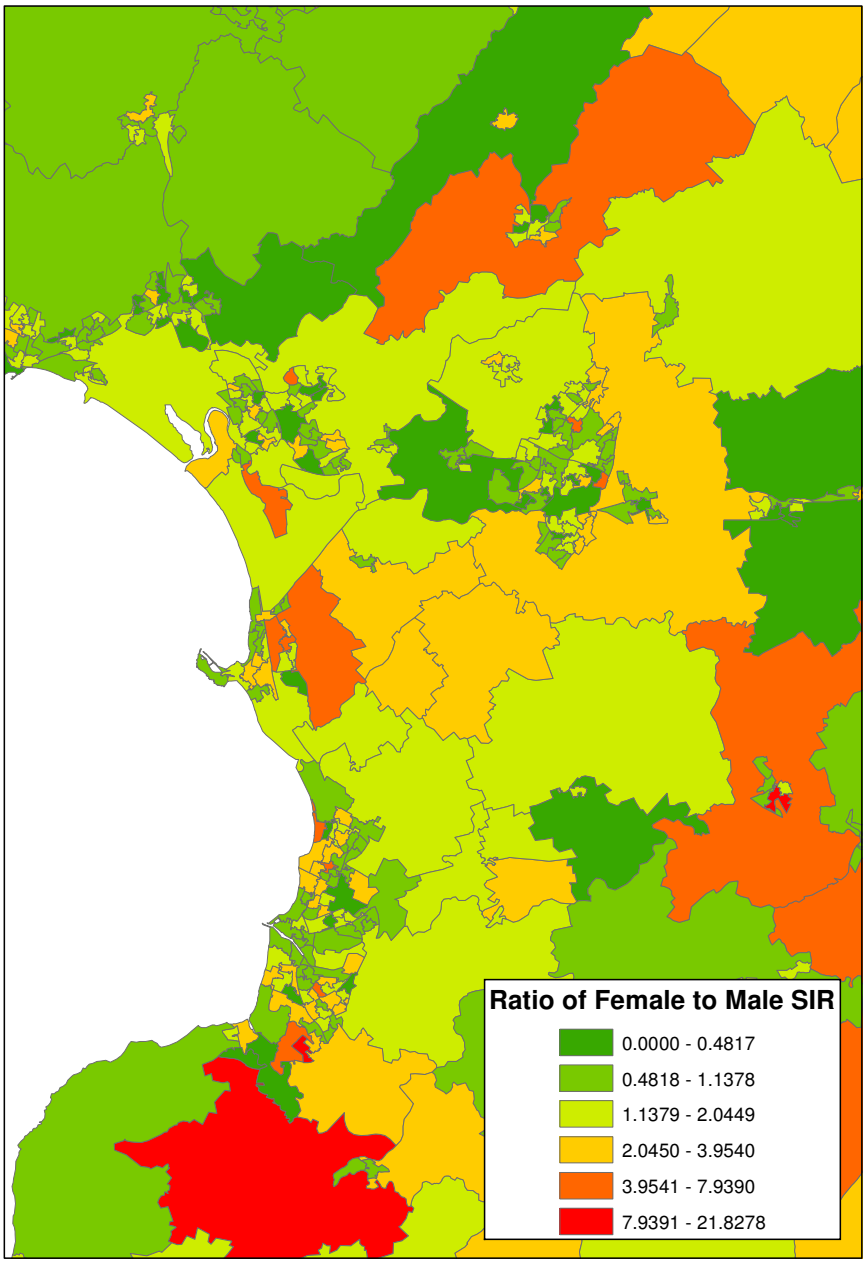


Figure 4.37: Data Zone Map of the Ratio of Female to Male SIR in the Ayrshire Area

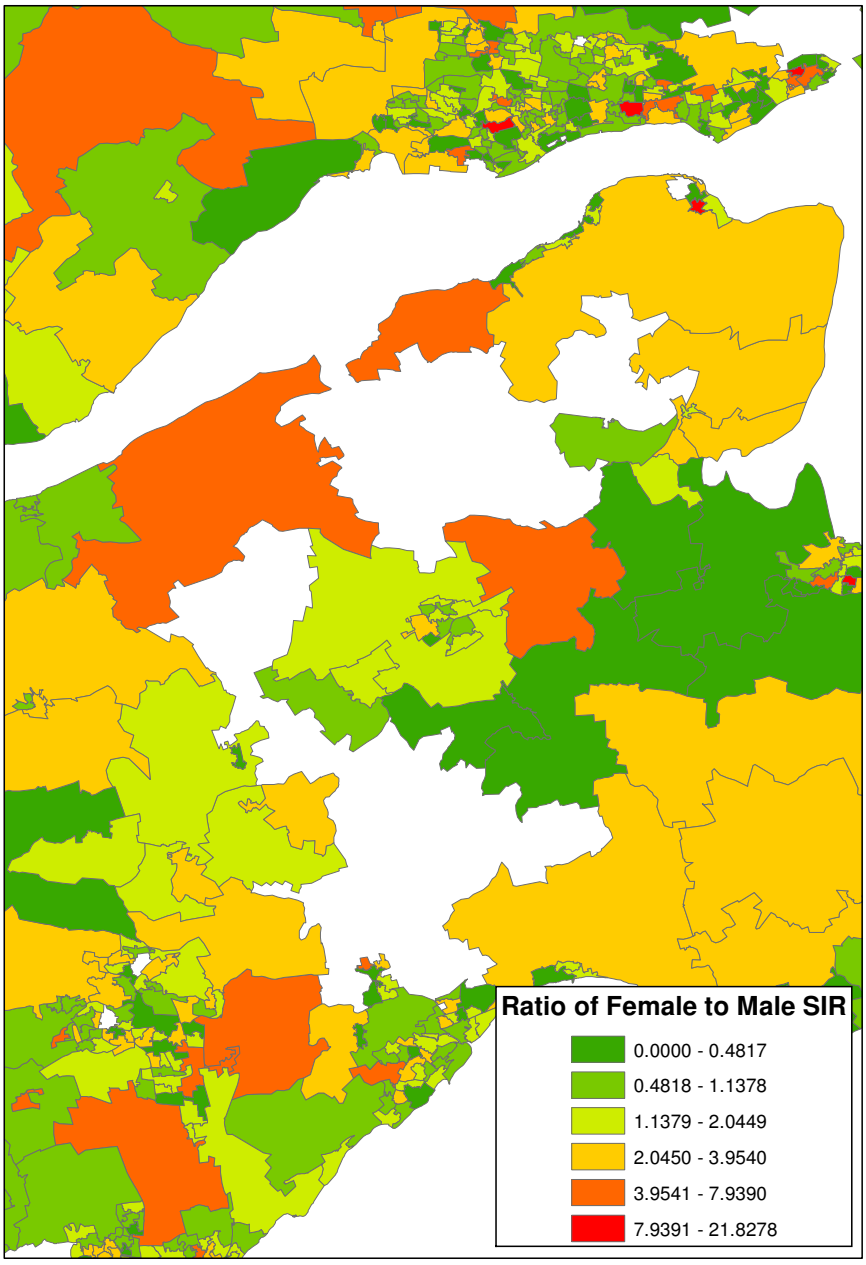


Figure 4.38: Data Zone Map of the Ratio of Female to Male SIR in the Dundee and Fife Area

4.5 SIR and Local Authority

As discussed above, due to the extremely small geographical area of many of the data zones, on A4 paper it is necessary to show the map of Scotland in

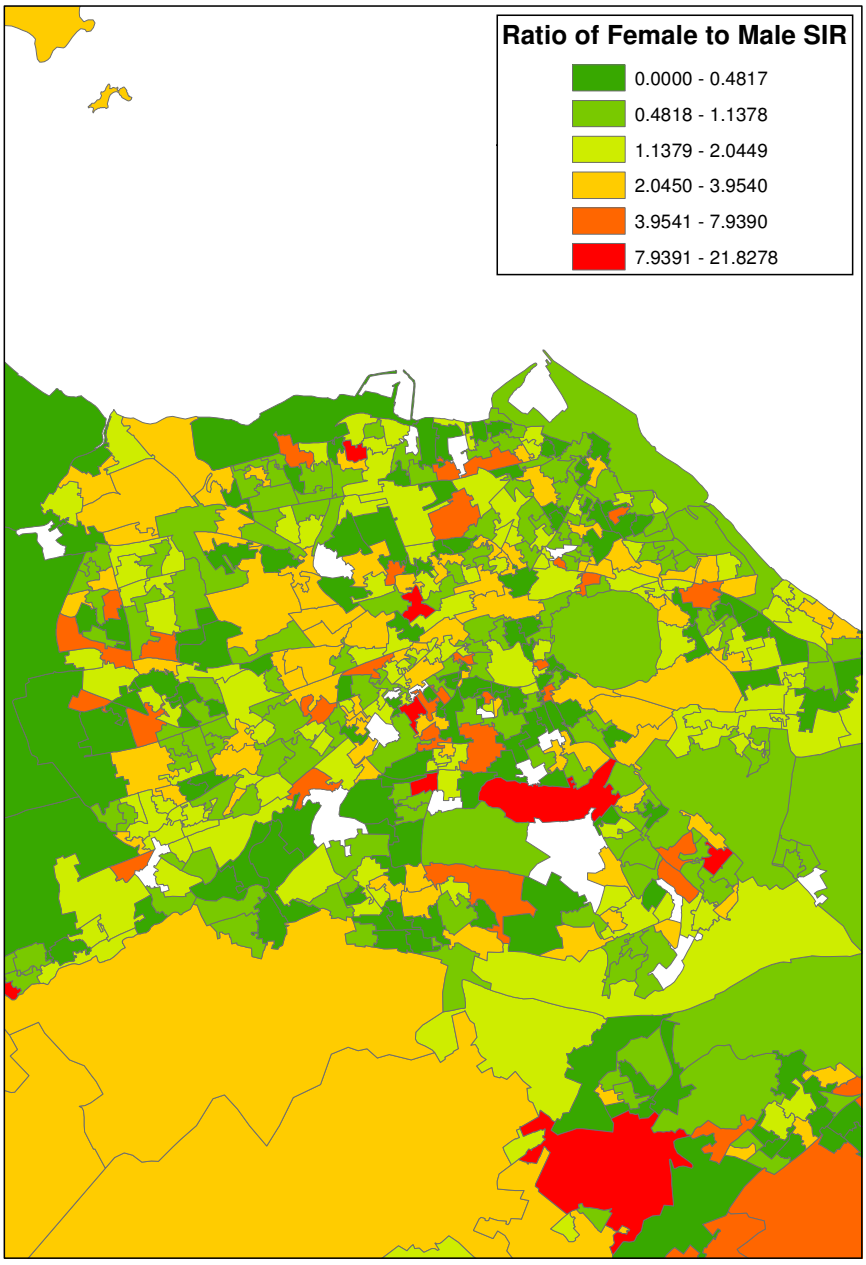


Figure 4.39: Data Zone Map of the Ratio of Female to Male SIR in the Edinburgh Area

sections. Viewing the SIR pattern in such a way makes it harder to compare several areas of the maps simultaneously than if they were shown in a single figure. In an attempt to highlight any differences within the SIR maps or

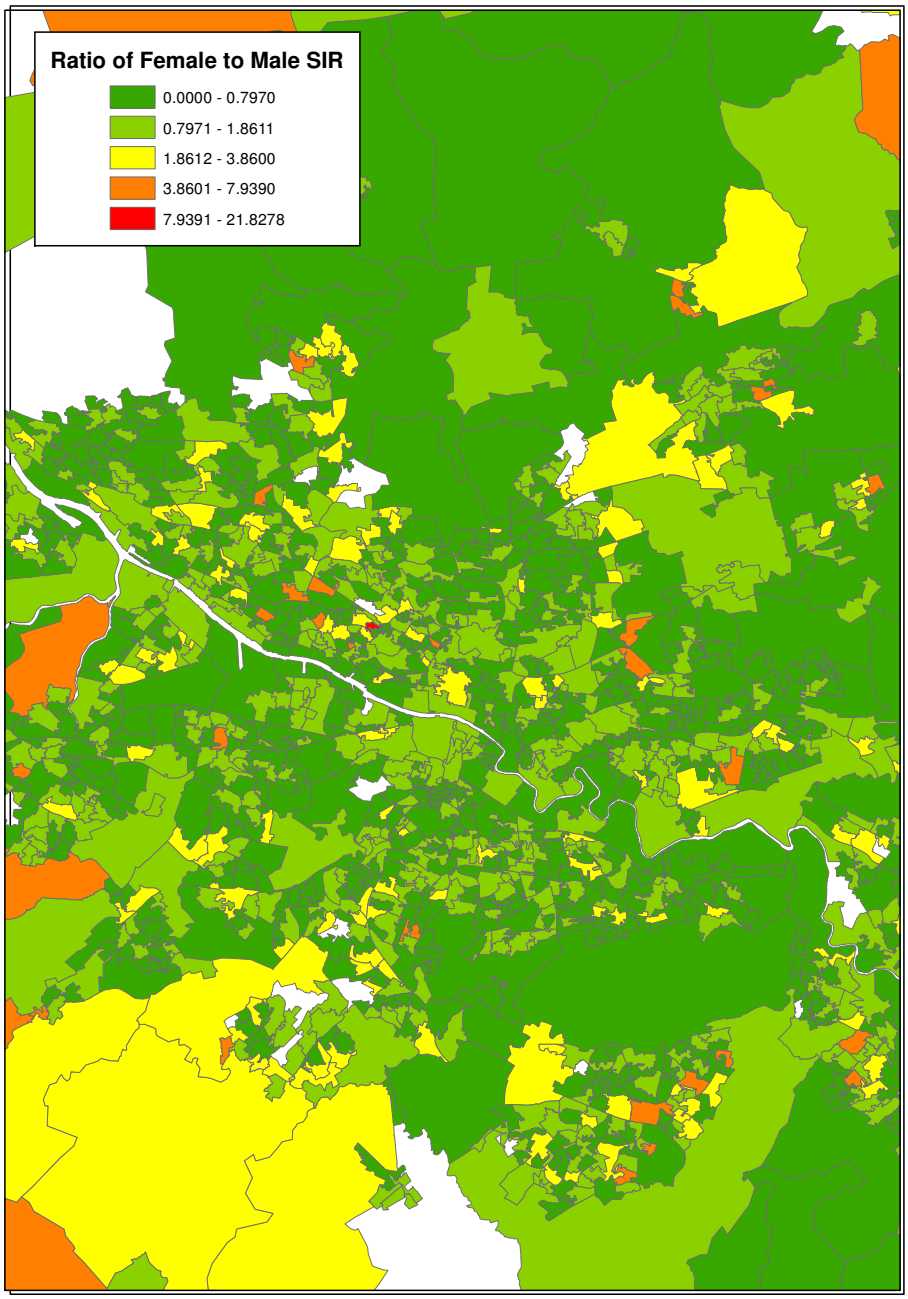


Figure 4.40: Data Zone Map of the Ratio of Female to Male SIR in the Glasgow Area

between male and female SIR values which were missed using sectioned maps, two further boxplots have been produced. Boxplots of SIR value by local authority are shown in Figure 4.43 for males and Figure 4.44 for females.

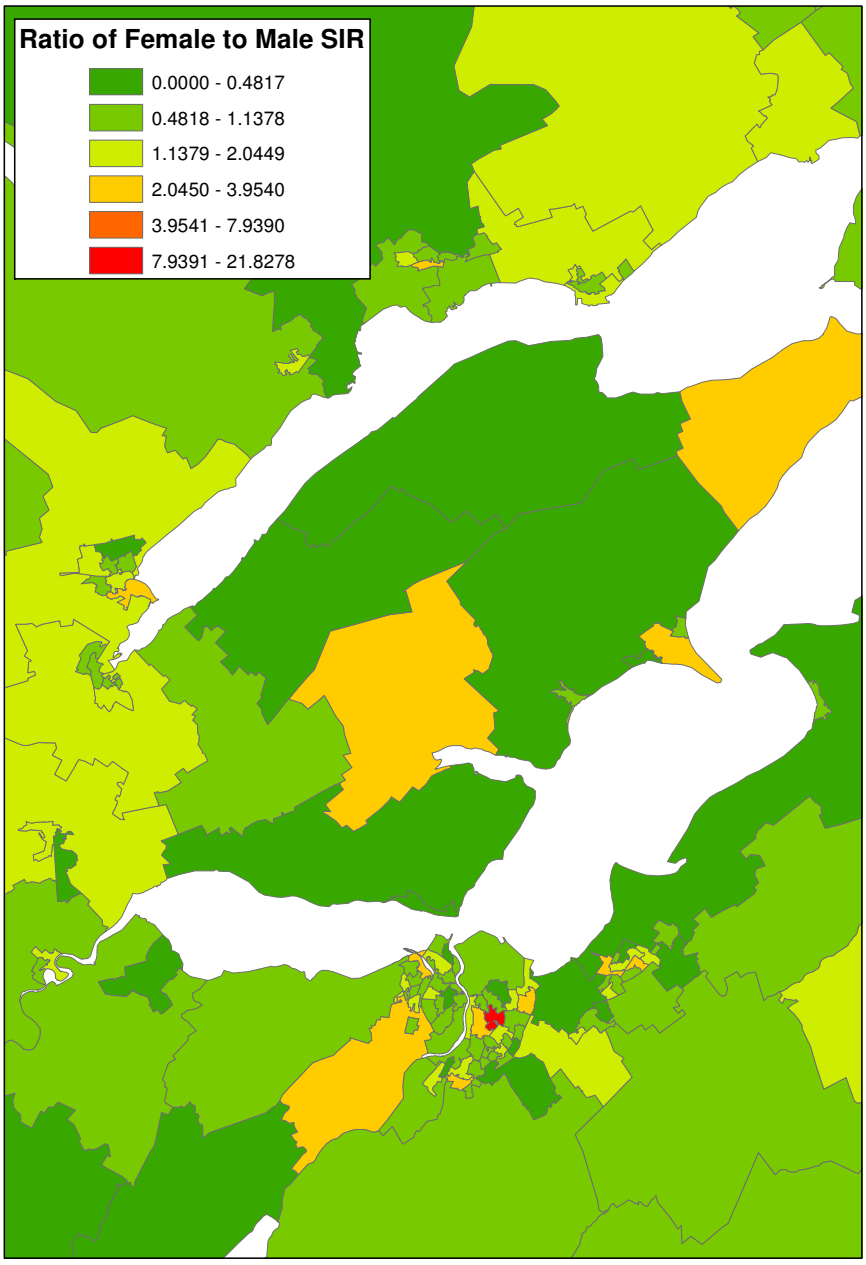


Figure 4.41: Data Zone Map of the Ratio of Female to Male SIR in the Inverness Area

Figure 4.43 suggests that the local authorities with the highest median male SIR value for the period in question are Eilean Siar (the Outer Hebrides), Glasgow City and Inverclyde. This result is no surprise in terms of

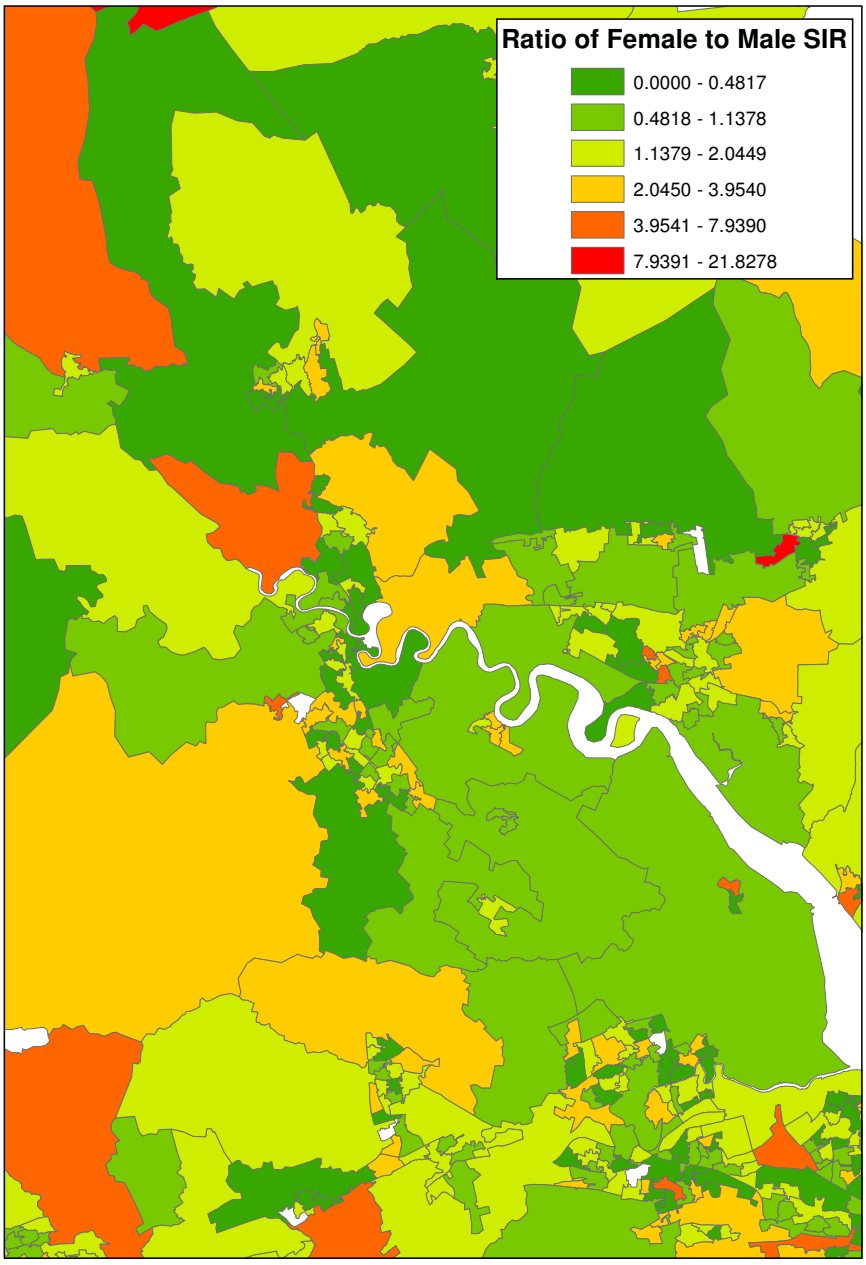


Figure 4.42: Data Zone Map of the Ratio of Female to Male SIR in the Stirling Area

Glasgow City, but Eilean Siar is an area which may not have been expected to have such a high average male SIR. The Scottish data zone map of male SIR (Figure 4.13) does suggest that there were high values experienced in

Eilean Siar, but none appears to be in the most deprived category of 3.14 to 7.95. The similarity in average male SIR value in Eilean Siar compared with that of Glasgow City must be due to the fact that, although there are many extremely high values experienced within Glasgow, these are cushioned by the appearance of several low male SIR values between 0 to 0.47.

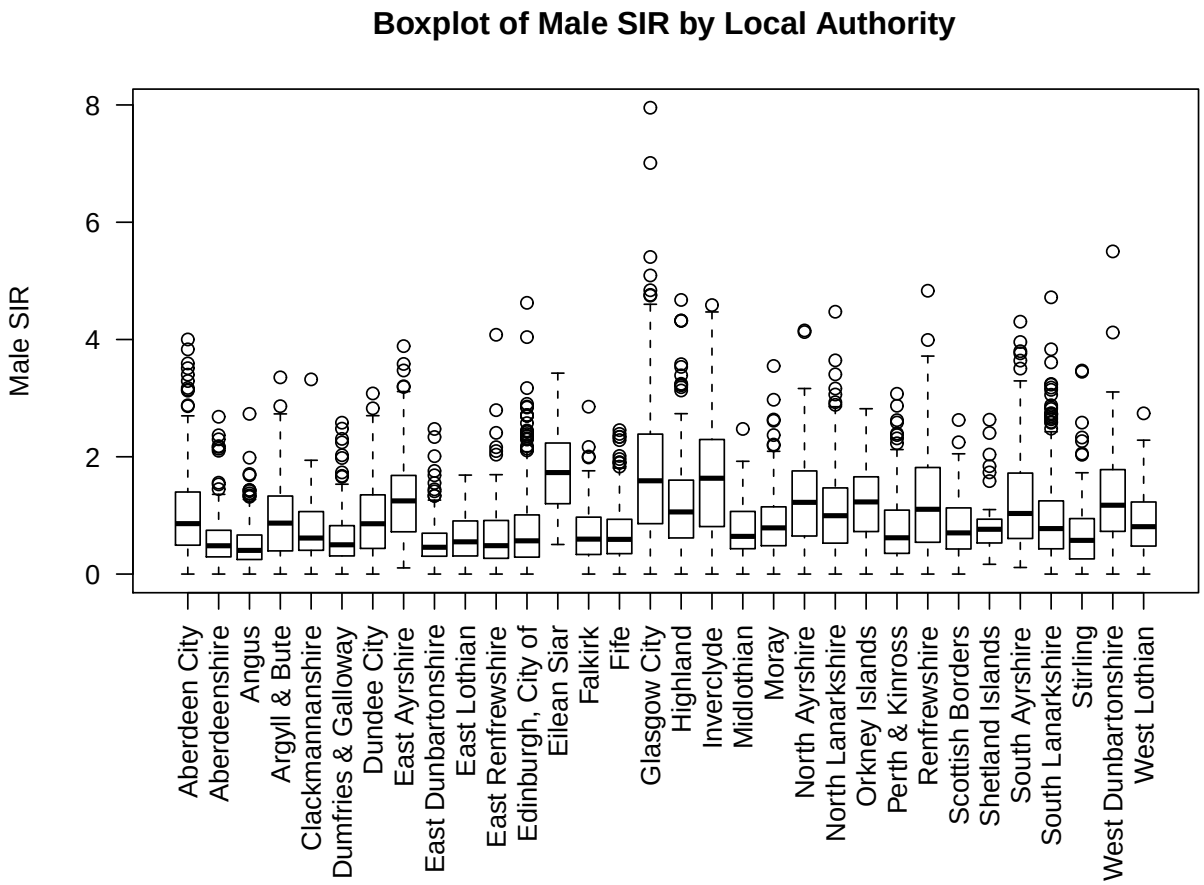


Figure 4.43: Boxplot of Male SIR by Local Authority

The boxplots of female SIR by local authority (Figure 4.44) are now considered. For females the three local authorities with the highest median SIR value are Eilean Siar, East Ayrshire and South Ayrshire. Eilean Siar appears to stand out more for the females than for the males as having the highest median value. The patterns then do not seem to be the same for male and female SIR across Scotland. The plots for males and females also suggest

that alcohol-related risk in Scotland occurs to differing degrees throughout the country, with high risks being experienced in both rural and inner city locations.

However due to the sparseness of the female data compared to that of the and the discussed drawbacks of the SIR method, there should not be too much weight placed upon these plots.

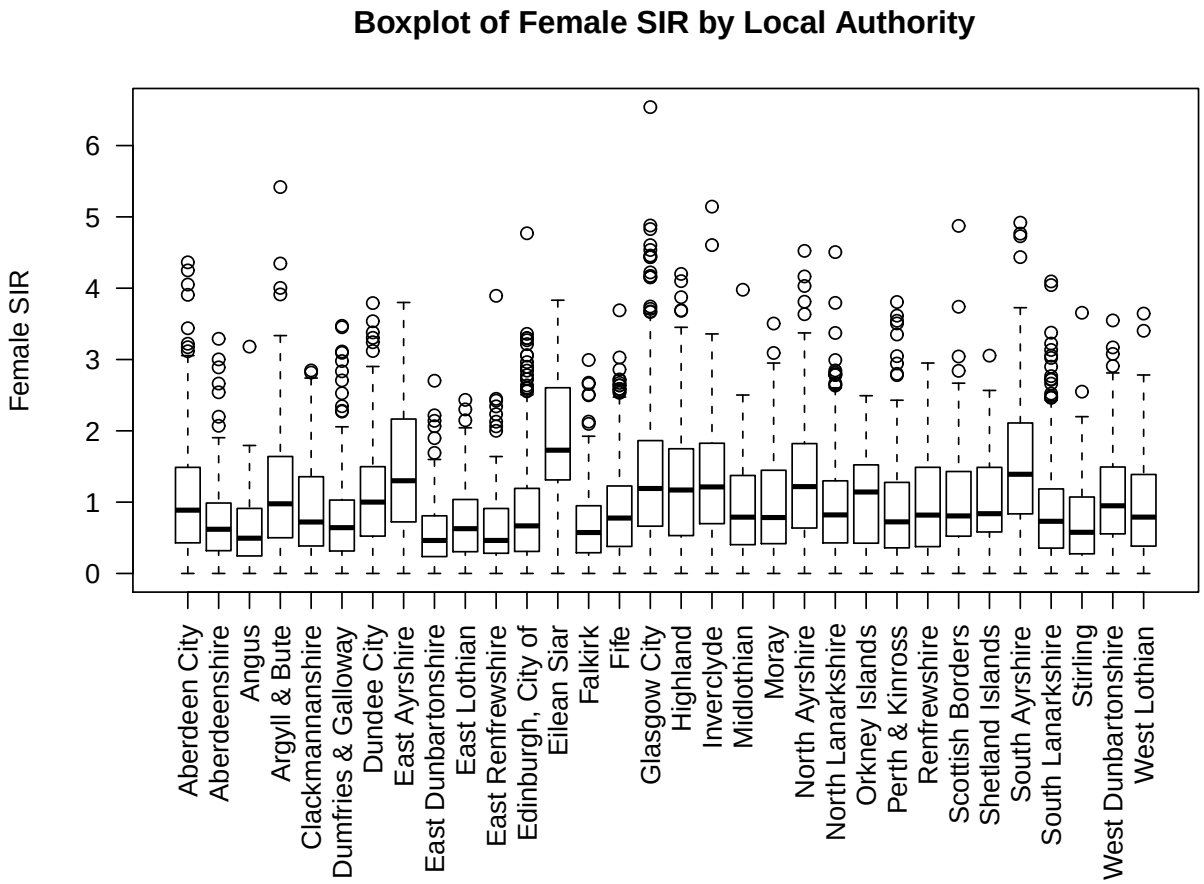


Figure 4.44: Boxplot of Female SIR by Local Authority

Chapter 5

BYM Models for Combined Data

Although the SIR is a quick and easy risk estimate, it has some serious drawbacks which have been discussed earlier in this thesis. With the aim of overcoming these problems model-based estimates for the relative risk of alcohol-related death or hospitalisation in each data zone across Scotland will now be considered.

Firstly, models will be fitted to the combined male and female data. The basis for the model structure used is that of the Besag, York and Mollié model (Besag et al. (1991)) discussed in Chapter 3. This is a spatial Bayesian model which considers both correlated and uncorrelated heterogeneity. The models are based on the expected and observed counts of alcohol-related deaths and hospitalisations in each data zone.

5.1 Models Considered

This chapter will investigate nine different models for the combined alcohol-related relative risk across the data zones. These models will differ in terms of both fixed effects and random effects. As random effects each model will include either uncorrelated heterogeneity (v), correlated heterogeneity (u) or

a convolution prior $(u + v)$. Area deprivation score is the only fixed effect which is explicitly fitted in any of the models. The expected count data used to fit these models has already been standardised for age and sex, so they should not be included at the model-building stage. Deprivation score has been modelled as a fixed effect in two different ways; firstly in a linear manner and secondly by assigning a separate parameter to each of the 10 deprivation scores.

Table 5.1 gives a summary of the nine different models compared in this section; it indicates how the deprivation score has been incorporated (if at all) and what random effects have been included.

Model Name	**Fixed Effects**	**Random Effects**
Model A-v	none	v
Model A-u	none	u
Model A	none	$u + v$
Model B-v	linear deprivation	v
Model B-u	linear deprivation	u
Model B	linear deprivation	$u + v$
Model C-v	non-linear deprivation	v
Model C-u	non-linear deprivation	u
Model C	non-linear deprivation	$u + v$

Table 5.1: Models for Combined Alcohol-Related Relative Risks

As explained in section 3.6, the Besag, York and Mollié model assumes that the relative risk in area i, θ_i, is given by

$$\theta_i = exp\,(\alpha + u_i + v_i)$$

where $exp(\alpha)$ is the baseline or overall level of relative risk. Model A fits exactly the BYM model.

Model B incorporates a linear term for area deprivation score, giving

$$\theta_i = exp\,(\alpha + \beta d_i + u_i + v_i)$$

where β is a parameter and d_i is the deprivation score in area i. The deprivation parameter β has been assigned a vague normal prior with mean 0 and a precision (inverse variance) of e_{-5}.

Model C goes one step further and adds a non-linear deprivation term to the basic model A. This gives

$$\theta_i = exp\left(\alpha + \beta_{d_i} + u_i + v_i\right)$$

where there is a separate parameter, β_1 to β_{10}, for each of the 10 deprivation scores. The parameter for the worst deprivation score of 1, β_1, has been arbitrarily set to zero and the remaining 9 parameters are given vague normal prior distributions with mean zero and precision e_{-5}. So, for Model C we have

$$\begin{aligned} \beta_1 &= 0 \text{ and} \\ \beta_j &\sim N(0, exp(-5)), \end{aligned}$$

for j in 2:10.

The background relative risk α is said to follow an improper flat prior in all 9 models. This is the most vague form of prior; it is effectively a uniform distribution across the entire real line, which means α has an equal prior probability of being any real value.

The code for all of the models specifies a normal prior distribution with mean zero for the uncorrelated heterogeneity and a conditional autoregressive prior for the correlated heterogeneity, so

$$\begin{aligned} v_i &\sim N\left(0, \tau_v^2\right) \text{ and} \\ \left[u_i | u_j, i \neq j, \tau_u^2\right] &\sim N\left(\overline{u_i}, \tau_i^2\right) \end{aligned}$$

where τ_v^2, $\overline{u_i}$ and τ_i^2 are as described in section 3.6.

Vague gamma hyperprior distributions have been assigned to the inverse variance hyperparameters of both random effects. In particular,

$$\tau_v^2 \sim gamma(0.5, 0.0005) \text{ and}$$

$$\tau_u^2 \sim gamma(0.5, 0.0005).$$

This hyperprior distribution has been chosen since it is sufficiently vague and commonly used in disease mapping studies where there is no strong prior knowledge.

All nine models have been run using OpenBUGS and the code for Model A, Model B and Model C is shown below in appendix section 10.1, 10.2 and 10.3 respectively. The code for the other variations of these models can be easily obtained by omitting the redundant parts of the code; for example, Model A-v can be obtained by deleting all parts of the Model A code which relate to the correlated heterogeneity random effect u.

5.2 Convergence

The aim of using any of the sampling methods discussed in Chapter 3 section is to simulate a Markov chain whose equilibrium distribution is the desired distribution (Gilks et al. (1996)). It is hoped that the joint distribution of the simulated values will converge, or stabilise, to the joint posterior distribution. Often such simulations will take a number of iterations to converge, but the length of this so called 'burn-in' period varies greatly between different studies and different models. It is necessary to carry out a number of convergence checks in order to determine a suitable number of burn-in iterations. All parameter estimates are based only on iterations after the burn-in period, so the values obtained during this period are effectively forgotten. There must be enough post-burn-in iterations to allow accurate posterior estimates to be calculated from the samples.

The models investigated simulate a separate relative risk parameter and in some cases two random effects for every single area. Given that there are 6505 data zones in the study it proved impractical to record these parameter values at every iteration. For all nine combined models the relative risk has been fully monitored for a subset of the data zones and a summary monitor has been set for the remaining areas. All other parameters in the models

have been fully monitored. A summary monitor gives exact estimates of the mean and standard deviation of the simulated parameter sample along with approximate 95% credible intervals.

The relative risk estimate was recorded at every iteration for the data zones given in Table 8.2 below.

Data zone Code	**Relative Risk Parameter**	**Reason Chosen**
S01006393	θ_{115}	poor deprivation score
S01006438	θ_{14}	poor deprivation score
S01006490	θ_{1}	good deprivation score
S01006505	θ_{2}	good deprivation score
S01003744	θ_{2521}	rural area
S01003915	θ_{2692}	rural area
S01003380	θ_{3044}	urban/city area
S01002325	θ_{4687}	urban/city area
S01005521	θ_{985}	island / no neighbouring areas
S01000447	θ_{6238}	island / no neighbouring areas

Table 5.2: Data Zones with Fully Monitored Relative Risk Estimates

It was found that adequate convergence was achieved by all of the combined data models after a burn-in of 10,000 iterations, after which each model was run for a further 150,000 iterations. Two identical sampling algorithms were run simultaneously from different starting points, in order to allow more robust checks for convergence. A variety of methods were used to ascertain convergence and these are discussed below.

Convergence will only be discussed in detail for Model C-u, since the same checking methods were used and satisfied for all nine models.

Firstly, the history plots for a selection of the relative risk and other parameters from Model C-u are shown in Figures 5.5, 5.6, 5.7 and 5.8. These plots show the parameter value at each iteration post burn-in period against the iteration number, for both chains on the same plot. Every one of these history plots indicates that Model C-u has converged well. They exhibit no obvious patterns or trends and the lines for each chain form consistently overlapping horizontal bands across the plots. This is strong evidence that both chains have converged, or settled, to a stable posterior distribution.

However, it must be remembered that it is still possible that the simulation has just become 'stuck' in a certain area of the parameter space.

For the same subset of Model C-u parameters the Gelman-Rubin diagnostic plots, as discussed in section 3.4.3, are shown in Figure 5.3 and Figure 5.4. Again, all of these plots suggest that both chains have achieved adequate convergence. This is because the green line, which shows the width of the central 80% interval of the pooled chains, and the blue line, which shows the average width of the 80% intervals within the individual chains, are both stable and the red line which represents their ratio is stable at a value of 1. In fact, the intervals are so similar for the individual chains and the pooled chains that the blue line almost completely obscures the green line.

A further indication of the satisfactory convergence of Model C-u is that, for the same subset of parameters, the posterior density plots shown in Figure 5.1 and Figure 5.2 all appear to be smooth. A lack of convergence often results in such parameter posterior density plots appearing more uneven and 'spikey'.

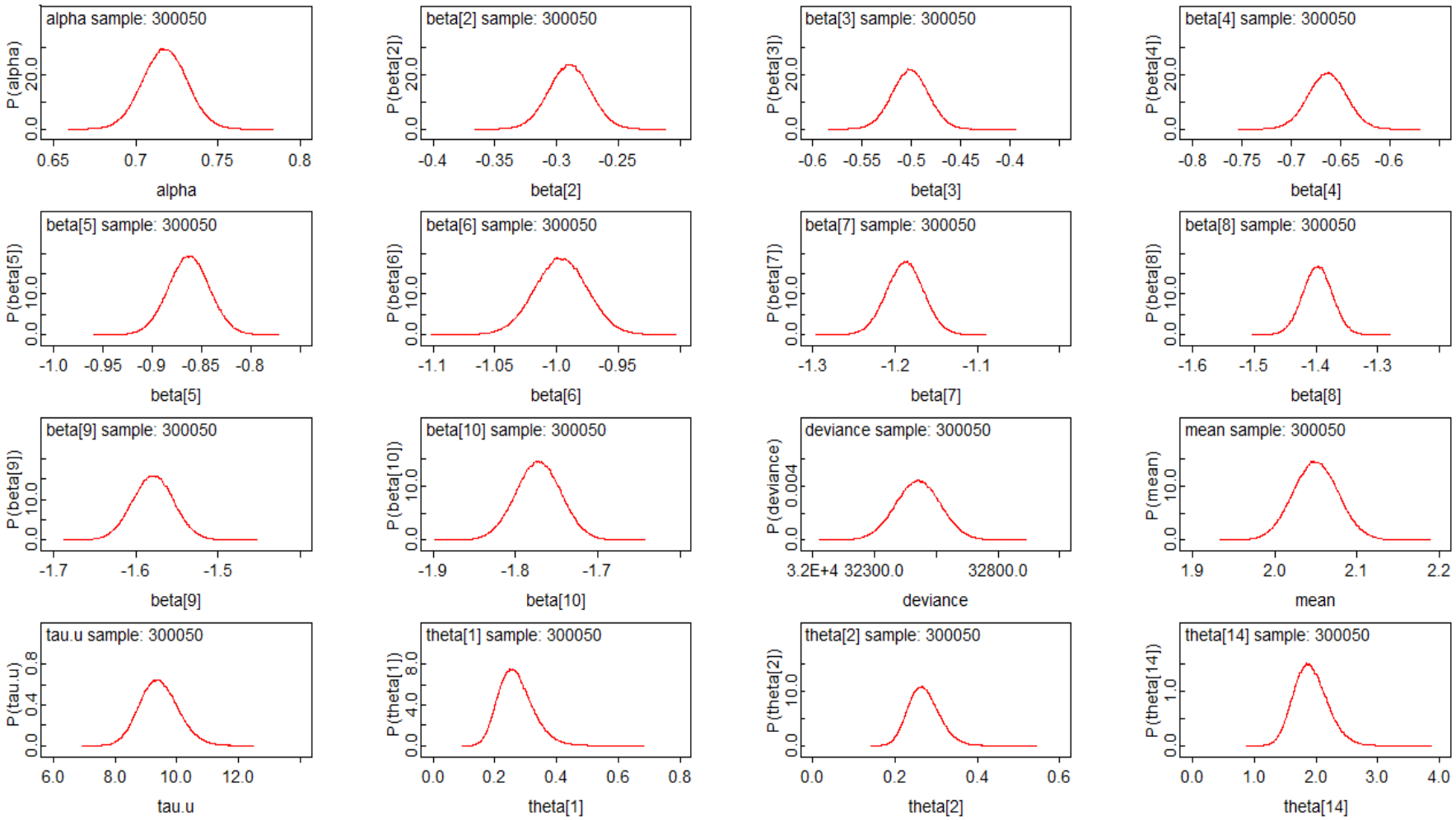


Figure 5.1: Posterior Density Plots for a Subset of Model C-u Parameters (part 1)

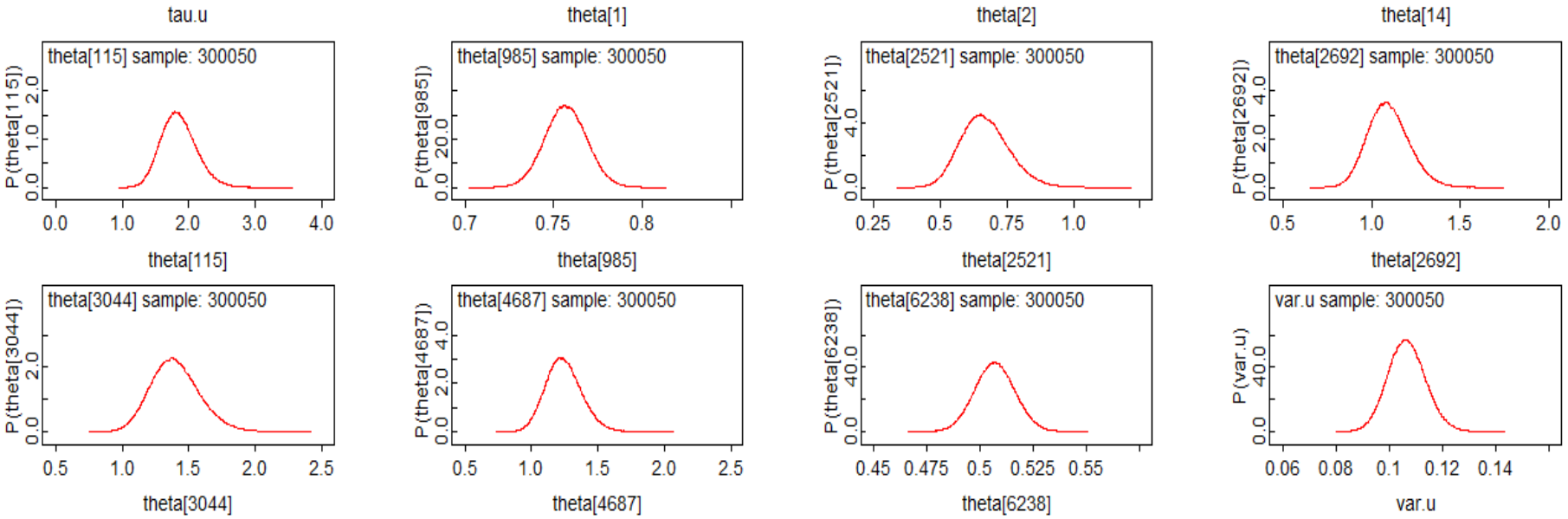


Figure 5.2: Posterior Density Plots for a Subset of Model C-u Parameters (part 2)

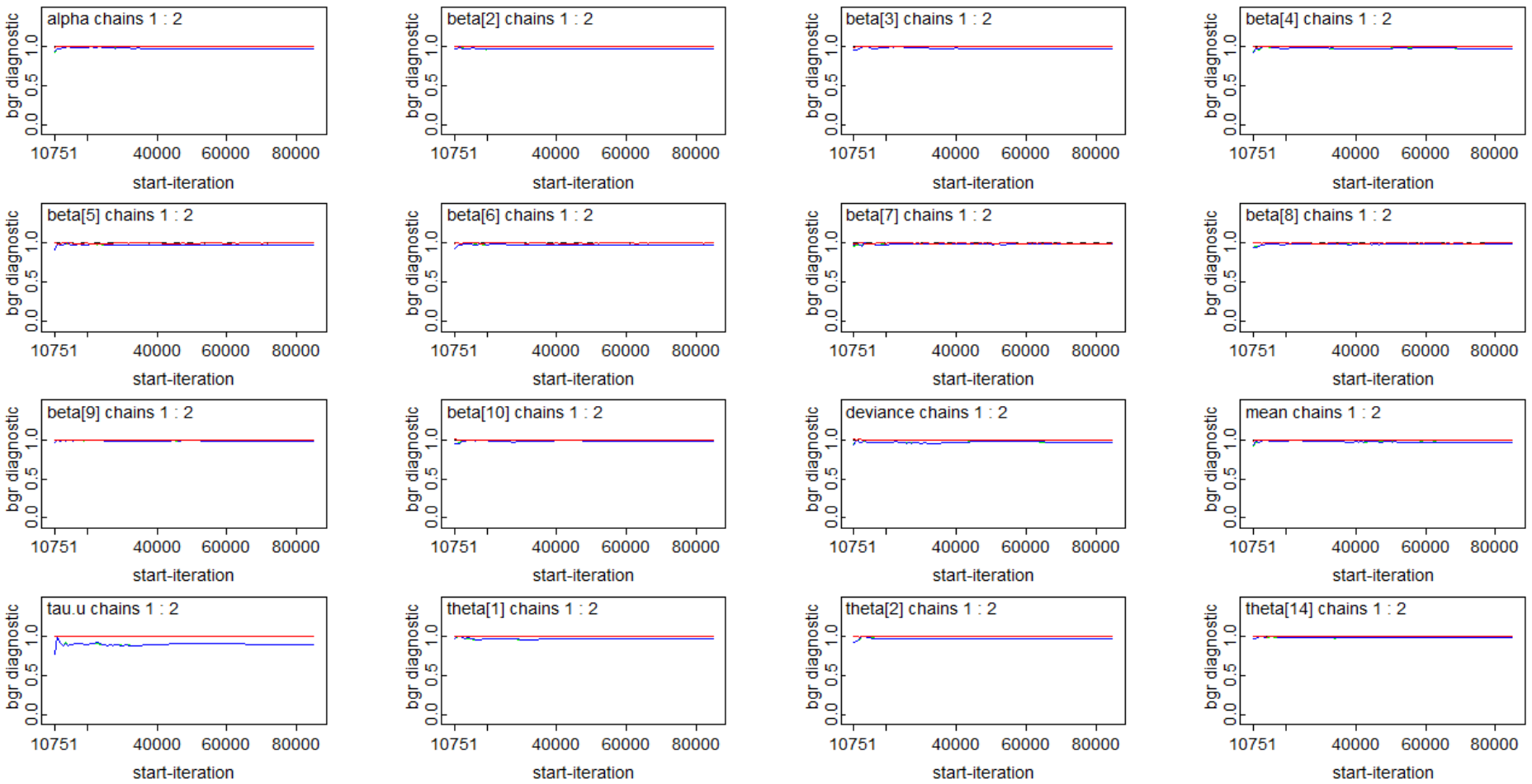


Figure 5.3: BGR Diagnostic Plots for a Subset of Model C-u Parameters (part 1)

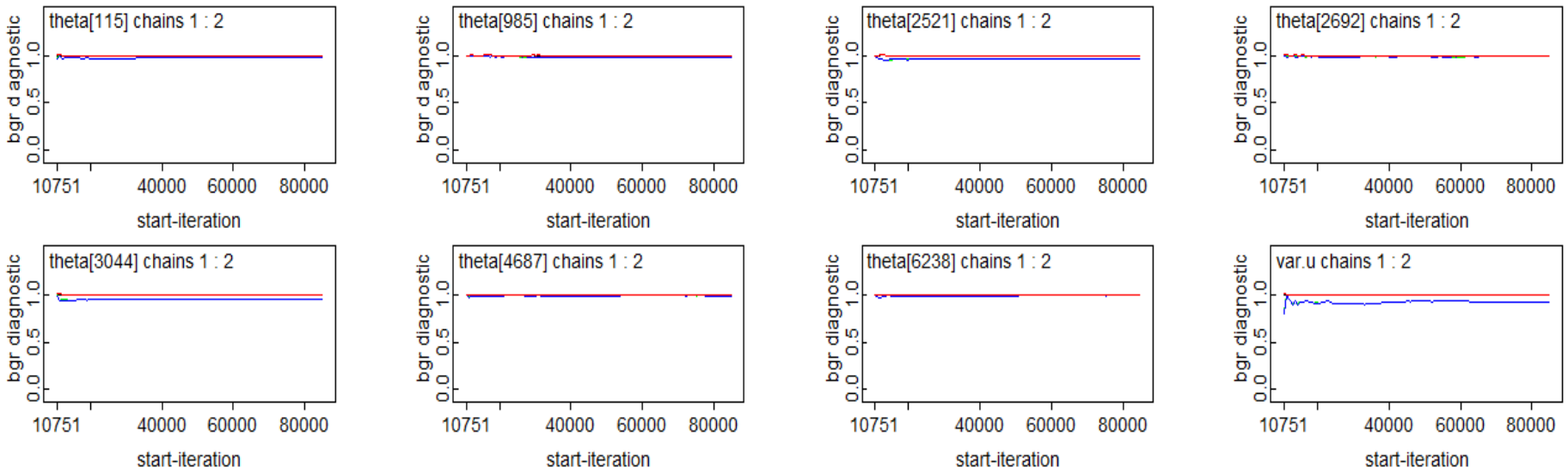


Figure 5.4: BGR Diagnostic Plots for a Subset of Model C-u Parameters (part 2)

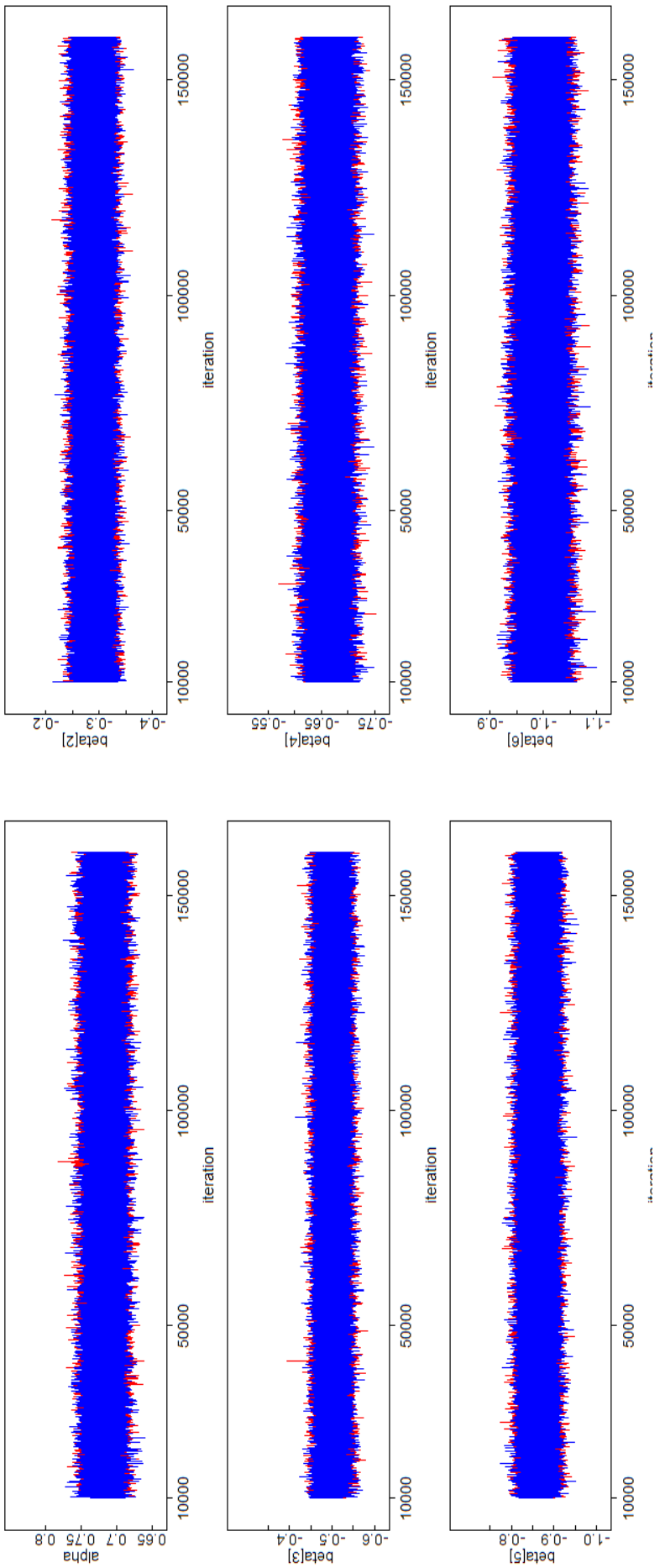


Figure 5.5: Simulation History Plots for a Subset of Model C-u Parameters (part 1)

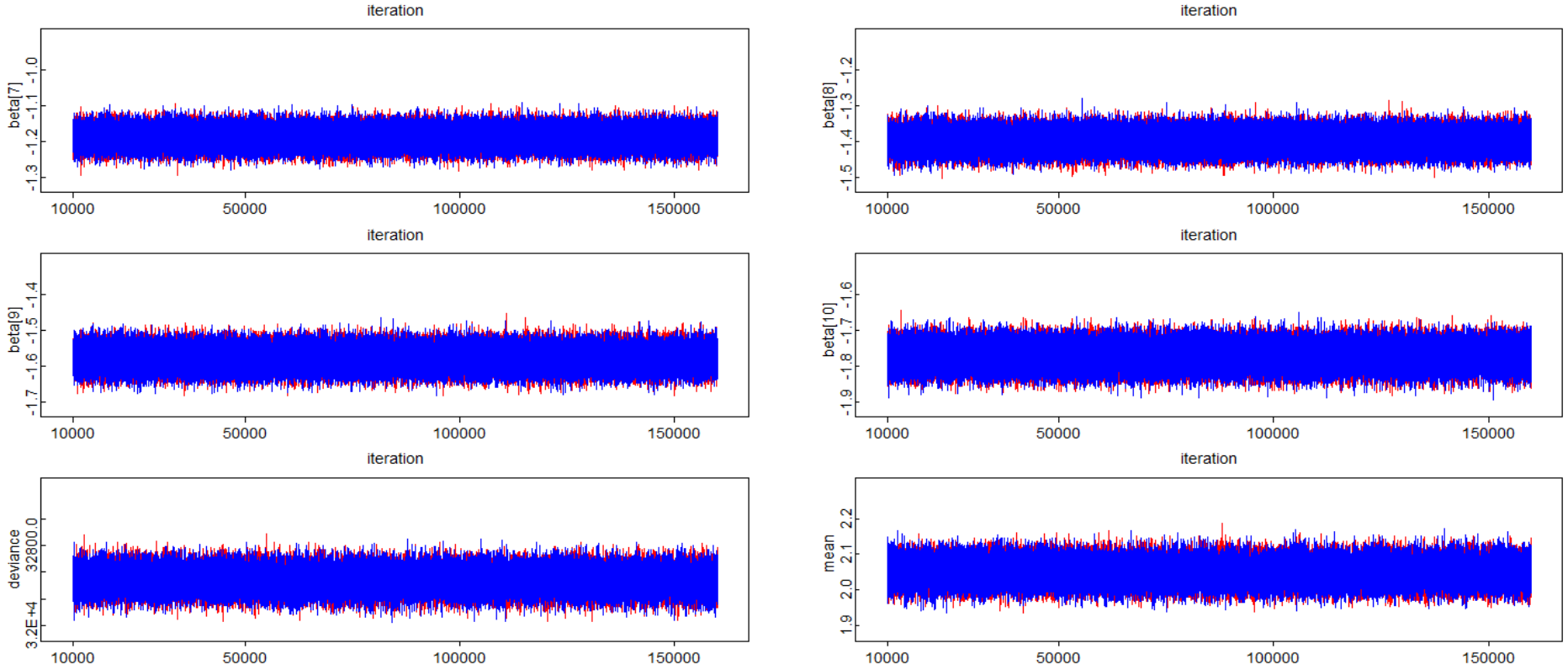


Figure 5.6: Simulation History Plots for a Subset of Model C-u Parameters (part 2)

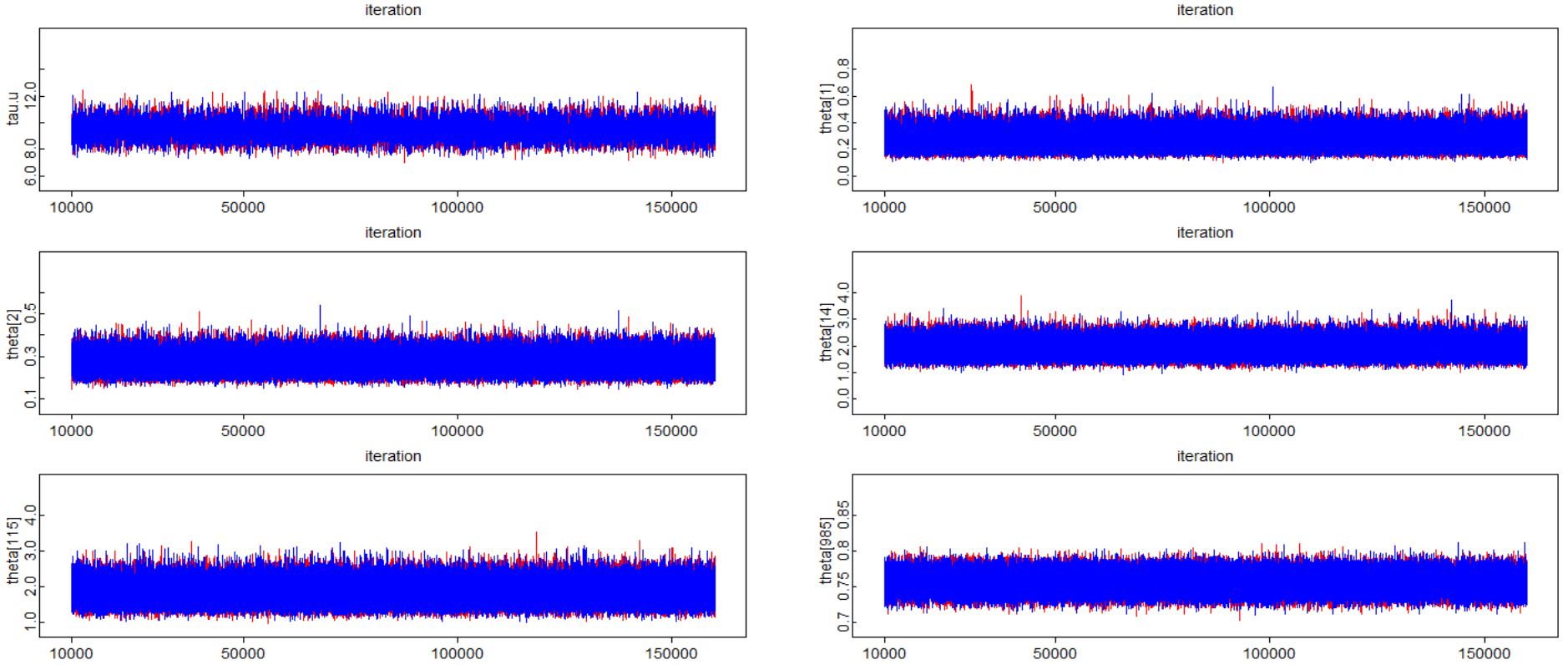


Figure 5.7: Simulation History Plots for a Subset of Model C-u Parameters (part 3)

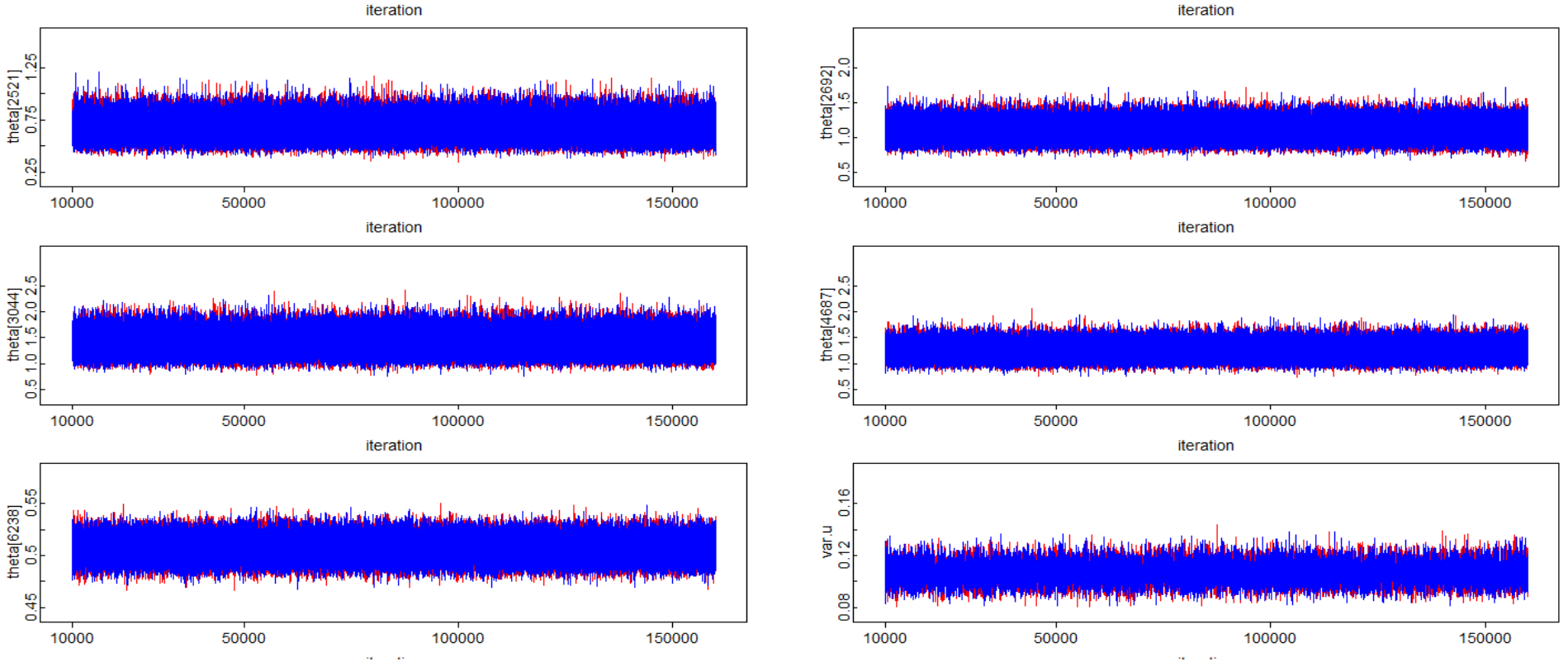


Figure 5.8: Simulation History Plots for a Subset of Model C-u Parameters (part 4)

5.3 DIC

As discussed in Chapter 3 the Deviance Information Criterion, DIC, is a commonly used measure of goodness of fit for spatial Bayesian models such as the BYM model. Table 5.3 gives the deviance, DIC and related values calculated using the pD method for all nine models fitted to the combined data in this chapter. In this table the model results are sectioned in two ways: firstly, split by the fixed effects they contain, either no deprivation, linear deprivation or non-linear deprivation, and secondly, the results are also split into models which were fitted using only correlated/spatial heterogeneity (u), uncorrelated heterogeneity (v) or both ($u + v$).

Table 5.3 shows that the DIC is lowest for the model which does not incorporate deprivation but includes both uncorrelated and correlated heterogeneity effects (Model A). This is definitely not what one would expect, since when comparing the data zone deprivation score maps and data zone SIR maps the patterns shown are very similar. It is therefore highly unlikely that deprivation score does not explain a significant amount of the variation in relative risks. However, when fitting a spatially smooth model which includes an equally smooth covariate, it is not uncommon that the model selection process suggests to remove the covariate, even though it seems to be highly relevant. The issue is that both the covariate and the smooth random effects compete with each other as they have similar explanatory power.

Both the posterior mean of the deviance ($\overline{D}$) and the point estimate of the deviance obtained by substituting in the posterior means of the other model parameters ($\widehat{D}$) are given. The DIC section on the WinBUGS website indicates that $\widehat{D}$ is a better measure of fit than $\overline{D}$ which can be considered more of a measure of adequacy. So Table 5.3 suggests that Model A-v fits the data best, which is again unexpected.

The most obvious problem presented by Table 5.3 is the negative pD values. The pD value should represent the effective number of parameters in

the model, but for two of the nine models investigated (Model A and Model B-u) this value is negative. This is possible but highly undesirable. Such negative pD values can occur when there is a non-log-concave likelihood, when the posterior for a parameter is especially asymmetric or bimodal or when there is another situation that causes the posterior mean to be a poor summary statistic causing large deviance values. Upon investigation it does not appear that any of the posterior distributions for the fully monitored parameters are particulary multi-modal or skewed. Unfortunately, since it was not possible to fully monitor the majority of the model parameters it is not possible to rule out posterior multi-modality or asymmetry as a cause of the negative pD values.

Due to the negative pD results achieved using the DIC values in Table 5.3 it has been decided to focus on an alternative method for calculating the DIC. This method uses an alternative to pD known as p*D which was developed by Andrew Gelman (as discussed in Chapter 3). Since each p*D value is calculated as a proportion of the corresponding parameter sample variance they cannot be negative. Table 5.4 gives the p*D, DIC calculated using p*D and deviance values. In this table the lowest DIC of 36591 corresponds to Model C-u which includes non-linear deprivation and correlated heterogeneity. The p*D method of calculating DIC therefore results in much more intuitive model selection.

5.4 Model Selection

It would be normal practice to choose Model C-u as the 'best' model since it has the lowest DIC value. The selection of Model C-u indicates that the relationship between alcohol-related relative risk and deprivation score in Scotland is not linear. This is consistent with previous discussions in Chapter 4, which noted that there was a larger increase in average SIR value between the deprivation scores of 1 and 2 than between any other

pair of consecutive scores. The difference in average combined SIR between sequential deprivation scores increases as deprivation worsens. So when using the p*D method of DIC, it seems to lead to a sensible model choice in this case.

It must, however, be remembered that these models are based on assumptions which are set by the modeller by way of the prior and hyperprior distributions for the parameters and hyperparameters. Although a model which incorporates non-linear deprivation and only correlated heterogeneity has been chosen under the current priors, it needs to be investigated whether this would normally be the case. With limited time to complete this project a comprehensive analysis of how sensitive the model results are to the priors is not possible. However, sensitivity to the gamma(0.5, 0.0005) hyperpriors assigned to τ_u^2 and τ_v^2 will be examined.

5.5 Hyperprior Sensitivity Analysis

The nine models given in Table 5.1 will be re-run with different hyperprior distributions assigned to τ_u^2 and τ_v^2. These sensitivity models will use the same names as those given in Table 5.1 but with 'Sens' appended at the end, for example 'Model A-Sens'. The reason for running the models with different hyperpriors is firstly to see whether the choice of prior will affect the model selection, and secondly, to see by how much the estimated alcohol-related relative risk estimates are affected by the alternative priors.

The alternative hyperpriors considered in each model (where necessary) are

$$\tau_u^2 \sim \text{Gamma}(1, 1) \text{ and}$$
$$\tau_v^2 \sim \text{Gamma}(1, 1).$$

These distributions are much less vague and very different from the previous hyperpriors used. They are not ideal as a first choice. However, if the models

can be fitted using such different hyperpriors and still give similar results, this will suggest that the models are not too sensitive to hyperprior choice.

Convergence of all sensitivity models was also monitored and checked and these models were also found to converge adequately after 10,000 iterations. As with the original models, each sensitivity model was then run for a further 150,000 simulations.

Once all nine sensitivity models had been run the DIC values were calculated. Table 5.5 gives the deviance statistics, pD and DIC values calculated using the pD method for the sensitivity models. Negative pD values also arise for the sensitivity models, again only for models which contain the correlated heterogeneity term u_i.

Due to the occurrence of these negative pD values, the p*D method of calculating DIC will also be used for the sensitivity models. Table 5.6 gives the deviance statistics, p*D value and DIC value calculated using p*D for all nine sensitivity models. Using this method the DIC value was lowest for the sensitivity model which incorporates fixed effects for non-linear area deprivation and both correlated and uncorrelated heterogeneity random effects (Model C-Sens).

If model selection is carried out for both the original and the sensitivity models using the p*D method of calculating DIC, the lowest values are observed for different models. However, both Model C-u and Model C-Sens include the spatial random effect u_i and a non-linear area deprivation score fixed parameter. The need for the additional non-spatial random effect is probably because the priors assigned to the precision terms of both random effects in the sensitivity models are much more restrictive. The original precision priors specify a mean of 1000 and a variance of 2,000,000 compared to both a mean and variance of 1 for the sensitivity models.

The similarity in the models chosen using very different priors suggests that the model structures are not overly sensitive to the hyperprior choice. It is also of interest to compare the actual parameter estimates of both chosen

models.

Some of the results from Model C-u and Model C-Sens are given in Table 5.7. The point estimates for all fully monitored parameters from these models are given along with corresponding credible intervals. The point estimates are calculated by treating the simulated chain values as a sample from the true posterior distributions and taking the mean of each parameter sample. The credible intervals used are equivalent to frequentist confidence intervals. The 95% central Bayesian credible intervals given comprise the 2.5% and the 97.5% quantiles of each parameter sample for the fully monitored parameters.

With the exception of the precision and variance parameters, every Model C-u parameter estimated in Table 5.7 lies within the corresponding Model C-Sens confidence interval and vice versa. This indicates that the choice of hyperprior has not dramatically affected the alcohol-related relative risks in each area. The results for the variance parameters in Table 5.7 show that, although Model C-Sens contains both correlated and uncorrelated heterogeneity and Model C-u contains only correlated heterogeneity, Model C-Sens attributes over 72% of the total variance to spatial effects.

Given the strong similarities between the Model C-u and Model C-Sens results for combined male and female alcohol-related relative risk across Scotland, Model C-u will be considered the final model since it has a much more appropriate hyperprior distribution and it has been shown not to be very sensitive to hyperprior choice.

5.6 Model Results

A selection of the parameters fitted in Model C-u are shown in Table 5.7. This table shows that none of the 95% credible intervals for the Model C-u deprivation parameters overlap or contain zero. This is strong evidence that all deprivation scores have a significant effect on combined alcohol-related relative risks in Scotland and hence should all be included in the model.

This information further supports the model choice; if the DIC had suggested this model but the individual deprivation score parameters proved not to be significant, this would potentially lead to an alternative model choice.

Box plots of the simulated area deprivation score parameters in Model C-u, β_2 to β_{10} (β_1 was arbitrarily set to zero so was not simulated), are shown in Figure 5.9. Although a non-linear fixed effect was found to be the most appropriate way to include the area deprivation score, the β parameter estimates themselves appear to be fairly linear. All of the deprivation score parameter estimates are negative and the value of β_{d_i} gets progressively smaller as d_i increases from a score of 2 to a score of 10. This is as one would expect; there is a larger decrease in the relative risk estimate for less deprived areas. The chosen model suggests that it is highly likely that, on average, the least deprived data zones with a deprivation score of 10 have an alcohol-related relative risk which is between only 0.161 and 0.179 times that of the most deprived data zones.

The structure of Model C-u means that the chosen model assumes that the alcohol-related relative risk in each data zone depends on the risk estimates in the neighbouring areas.

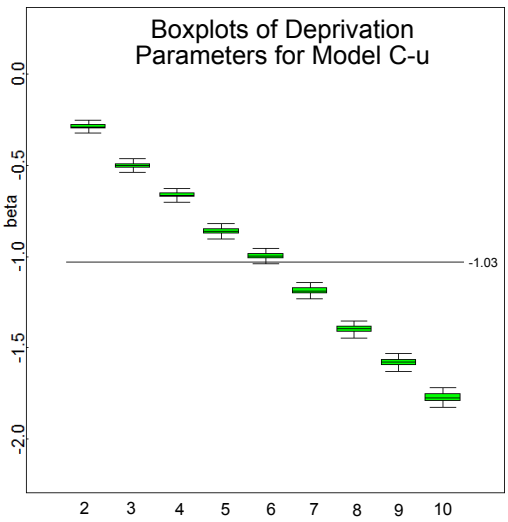


Figure 5.9: Boxplots of Deprivation Parameters for Model C-u

The ten hightest and ten lowest alcohol-related relative risks calculated using Model C-u are given in Table 5.8. It must be remembered that the credible intervals given for the model parameters which have been assigned a summary monitor are only approximate.When we compare this to the equivalent for the combined SIR values (Table 4.2) it is immediately apparent that there are no longer any zero relative risk estimates. Other than this however, the results do seem to be similar to those obtained using the SIR method. All of the ten lowest values correspond to data zones with an area deprivation score of at least 8, with 6 having the least deprived score of 10. For the highest combined relative risks estimated using this model, all ten correspond to data zones with the worst deprivation score of 1.

It should also be noted the most extreme high-risk estimates from Model C-u are lower than those obtained via the the combined SIR method. The modelling process has therefore reduced the problem of extremely low and extremely high risk estimates experienced with the SIR method due to the rarity of the disease and the extremely small study regions.

5.7 Alcohol-Related Relative Risk Maps

In this section the alcohol-related relative risk estimates calculated for each data zone in Model C-u have been mapped. A data zone map of Scotland depicting the combined relative risk estimates is shown in Figure 5.10, along with magnified sections of this map for Aberdeen (Figure 5.11), Ayrshire (Figure 5.12), the Dundee area (Figure 5.13), Edinburgh (Figure 5.14), Glasgow (Figure 5.15), the Inverness area (Figure 5.16) and Stirling (Figure 5.17).

Before any comparisons between the maps in this chapter and the combined SIR maps given in Chapter 4 can be made, it must be noted that the legend cut points used are not the same. However, bearing this in mind, it can be seen that the modelled risk estimates appear to give a very similar

general pattern to that of the combined SIR method. The modelled SIR values are, however, much smoother due to the incorporation of correlated heterogeneity. Large blocks of colour, which represent clusters of certain risk categories, can be observed in the maps for modelled risks.

There is a possibility that some would say the risk estimates for Model C-u have been forced to be over smooth by only allowing for spatial random effects in the model. The appropriateness of this model would depend on the intended use. If specific individual data zones are of interest then it may be decided that uncorrelated heterogeneity should also be included. However, if dealing with very small areas and/or rare diseases, it may be desired to include only correlated heterogeneity, where this seems reasonable, in order to encourage smoothing over differences between areas which only occur due to chance rather than to any real differences. For example, if two neighbouring areas are expected to experience 0.5 deaths in any given period but one experiences none and the other experiences 1, without including spatial random effects these would areas would have very different risk estimates.

	No Covariates				Non-linear Deprivation				Linear Deprivation			
Random Effects	$\overline{\mathbf{D}}$	$\hat{\mathbf{D}}$	**pD**	**DIC**	$\overline{\mathbf{D}}$	$\hat{\mathbf{D}}$	**pD**	**DIC**	$\overline{\mathbf{D}}$	$\hat{\mathbf{D}}$	**pD**	**DIC**
$\mathbf{v_i}$	32080	27200	4887	36970	32080	29300	2785	34870	32070	29250	2823	34890
$\mathbf{u_i}$	32330	32230	96.54	32420	32430	32310	119.5	32550	32420	32460	-43.3	32370
$\mathbf{u_i + v_i}$	32150	33010	-861.3	31290	32070	31540	531.3	32600	32060	31670	390.1	32450

Table 5.3: Deviance Statistics and DIC using the pD Method

	No Covariates				Non-linear Deprivation				Linear Deprivation			
Random Effects	$\overline{D}$	$\widehat{var}\{D\}$	p*D	DIC	$\overline{D}$	$\widehat{var}\{D\}$	p*D	DIC	$\overline{D}$	$\widehat{var}\{D\}$	p*D	DIC
v_i	32080	12996.0	6498.0	38578.0	32080	11794.0	5897.0	37977.0	32070	11794.0	5897.0	37967.0
u_i	32330	12122.0	6061.0	38391.0	32430	8322.9	4161.5	36591.5	32420	8416.2	4208.1	36628.1
$u_i + v_i$	32150	12746.4	6373.2	38523.2	32070	10691.6	5345.8	37415.8	32060	10753.7	5376.8	37436.8

Table 5.4: Deviance and DIC using the p*D Method

	No Covariates				Non-linear Deprivation				Linear Deprivation			
Random Effects	$\overline{D}$	$\widehat{D}$	pD	DIC	$\overline{D}$	$\widehat{D}$	pD	DIC	$\overline{D}$	$\widehat{D}$	pD	DIC
v_i	32080	27190	4888	36970	32060	29250	2808	34870	32040	29200	2846	34890
u_i	32320	32320	7.244	32330	32390	32590	-195.4	32200	32390	32440	-49.88	32340
$u_i + v_i$	32070	32320	-252.8	31820	31820	31050	772.4	32590	31820	31040	780.1	32600

Table 5.5: Deviance and DIC using pD Method (Sensitivity Models)

	No Covariates				Non-linear Deprivation				Linear Deprivation			
Random Effects	$\overline{D}$	$var\{D\}$	p*D	DIC	$\overline{D}$	$var\{D\}$	p*D	DIC	$\overline{D}$	$var\{D\}$	p*D	DIC
v_i	32080	12950.4	6475.2	38555.2	32060	11620.8	5810.4	37870.4	32040	11642.4	5821.2	37861.2
u_i	32320	12100.0	6050.0	38370	32390	8134.2	4067.1	36457.1	32390	8250.1	4125.0	36515.0
$u_i + v_i$	32070	12056.0	6028.0	38098.0	31820	8545.2	4272.6	36092.6	31820	8560.0	4280.0	36100.0

Table 5.6: Deviance and DIC using p*D Method (Sensitivity Models)

	Model 6-u		Model 6-Sens	
Parameter	**Estimate**	**95% Credible Interval**	**Estimate**	**95% Credible Interval**
α	0.7177	(0.6913, 0.7441)	0.7236	(0.6952, 0.7518)
β_1	0	NA	0	NA
β_2	-0.2894	(-0.3226, -0.2563)	-0.2971	(-0.3336, -0.2607)
β_3	-0.5014	(-0.5375, -0.4655)	-0.5116	(-0.5504, -0.4727)
β_4	-0.6636	(-0.7018, -0.6258)	-0.6728	(-0.7135, -0.6322)
β_5	-0.8627	(-0.9034, -0.8223)	-0.8747	(-0.918, -0.8316)
β_6	-0.9963	(-1.038, -0.9548)	-1.009	(-1.053, -0.965)
β_7	-1.188	(-1.232, -1.144)	-1.205	(-1.252, -1.159)
β_8	-1.398	(-1.444, -1.351)	-1.415	(-1.464, -1.366)
β_9	-1.579	(-1.628, -1.53)	-1.595	(-1.647, -1.544)
β_{10}	-1.772	(-1.826, -1.718)	-1.788	(-1.844, -1.732)
τ_u^2	9.43	(8.281, 10.72)	14.45	(12.34, 16.87)
τ_v^2	NA	NA	38.44	(32.77, 45.05)
var(u)	0.1065	(0.09324, 0.1208)	0.06964	(0.05927, 0.08102)
var(v)	NA	NA	0.02619	(0.0222, 0.03052)
θ_1	0.269	(0.1756, 0.3924)	0.2651	(0.1655, 0.4014)
θ_2	0.2704	(0.2038, 0.351)	0.2714	(0.1822, 0.3886)
θ_{14}	1.917	(1.434, 2.497)	1.901	(1.354, 2.574)
θ_{115}	1.849	(1.381, 2.409)	1.88	(1.336, 2.545)
θ_{985}	0.7569	(0.7339, 0.7802)	0.7561	(0.5585, 0.9971)
θ_{2521}	0.6679	(0.5071, 0.8622)	0.6835	(0.4698, 0.9562)
θ_{2692}	1.095	(0.8849, 1.34)	1.068	(0.7626, 1.446)
θ_{3044}	1.399	(1.075, 1.782)	1.306	(0.9342, 1.762)
θ_{4687}	1.243	(1.002, 1.521)	1.337	(0.9593, 1.804)
θ_{6238}	0.5068	(0.4887, 0.5253)	0.4889	(0.3552, 0.6555)

Table 5.7: Selection of Parameters from Model C-u and Model C-Sens

Rank	Data Zone	Intermediate Geography	Local Authority	Deprivation	Mean RR	95% Credible Interval
Lowest 10 RRs	S01004888	Westfield	North Lanarkshire	10	0.1687	(0.03261, 0.4241)
	S01005516	North & East Isles	Shetland Islands	8	0.1941	(0.03828, 0.487)
	S01000972	Gretna & Eastriggs	Dumfries & Galloway	9	0.1997	(0.1257, 0.2981)
	S01005350	Bishopton	Renfrewshire	10	0.2027	(0.1073, 0.3368)
	S01005515	North & East Isles	Shetland Islands	8	0.2033	(0.04037, 0.5088)
	S01005354	Bishopton	Renfrewshire	10	0.2037	(0.1075, 0.3386)
	S01005296	Houston South	Renfrewshire	10	0.2077	(0.1083, 0.3499)
	S01005356	Bishopton	Renfrewshire	10	0.2092	(0.1161, 0.3352)
	S01005301	Houston South	Renfrewshire	10	0.2098	(0.1087, 0.3546)
	S01000981	Gretna and Eastriggs	Dumfries & Galloway	9	0.2111	(0.1512, 0.2849)
Highest 10 RRs	S01006260	IZ Thirteen	West Dunbartonshire	1	3.526	(2.698, 4.499)
	S01003158	Toryglen & Oatlands	Glasgow City	1	3.536	(2.642, 4.584)
	S01006061	Shawfield & Clincarthill	South Lanarkshire	1	3.596	(2.785, 4.542)
	S01003860	Inverness Merkinch	Highland	1	3.681	(2.756, 4.772)
	S01005592	Ayr North Harbour, Wallacetown & Newton South	South Ayrshire	1	3.682	(2.889, 4.61)
	S01003862	Inverness Merkinch	Highland	1	3.754	(2.726, 4.98)
	S01003855	Inverness Merkinch	Highland	1	3.823	(2.998, 4.767)
	S01005594	Ayr North Harbour, Wallacetown & Newton South	South Ayrshire	1	3.865	(3.052, 4.789)
	S01003849	Inverness Merkinch	Highland	1	4.023	(3.092, 5.103)
	S01003043	Glenwood North	Glasgow City	1	4.108	(2.955, 5.495)
	S01003313	Parkhead West & Barrowfield	Glasgow City	1	4.293	(3.546, 5.132)

Table 5.8: Table of Fitted Combined Alcohol-Related Relative Risks

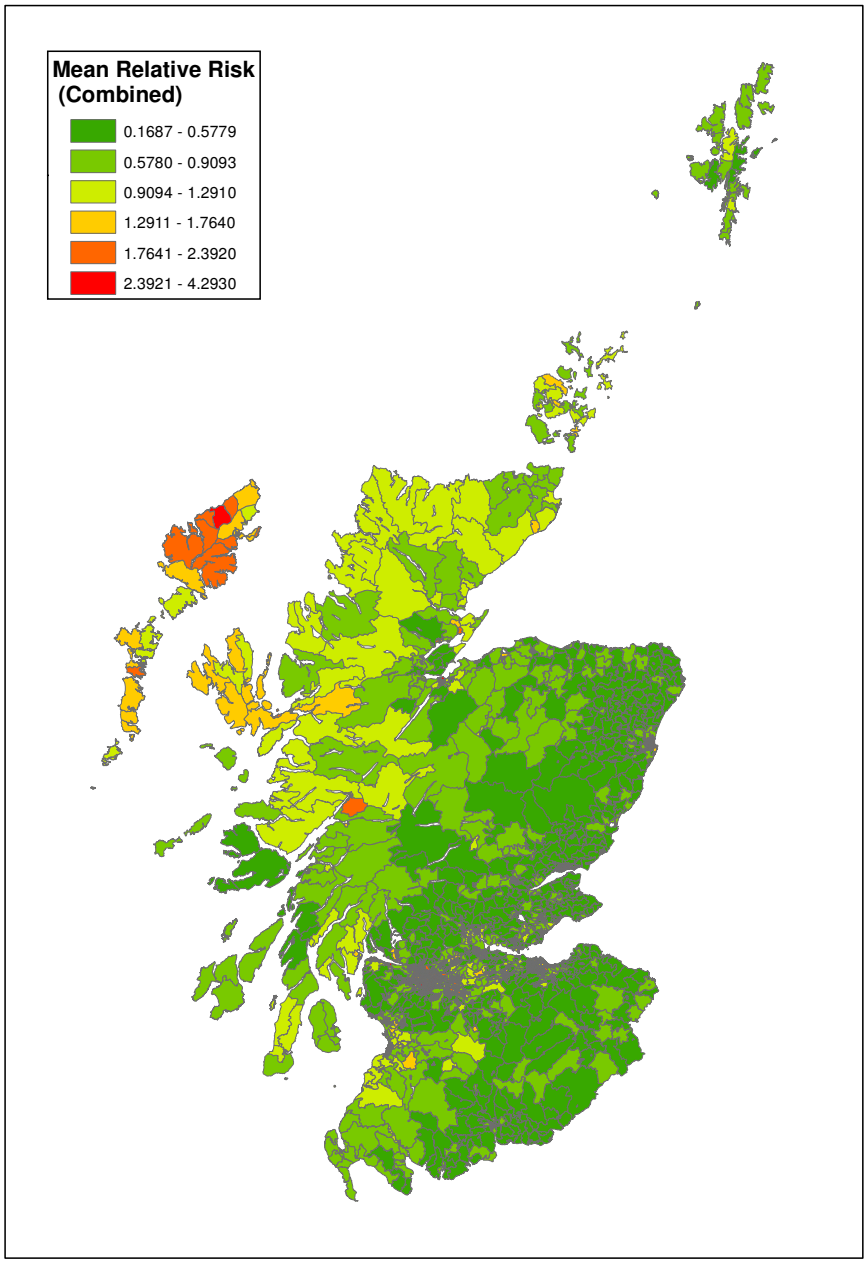


Figure 5.10: Data Zone Map of Mean Alcohol-Related Relative Risk

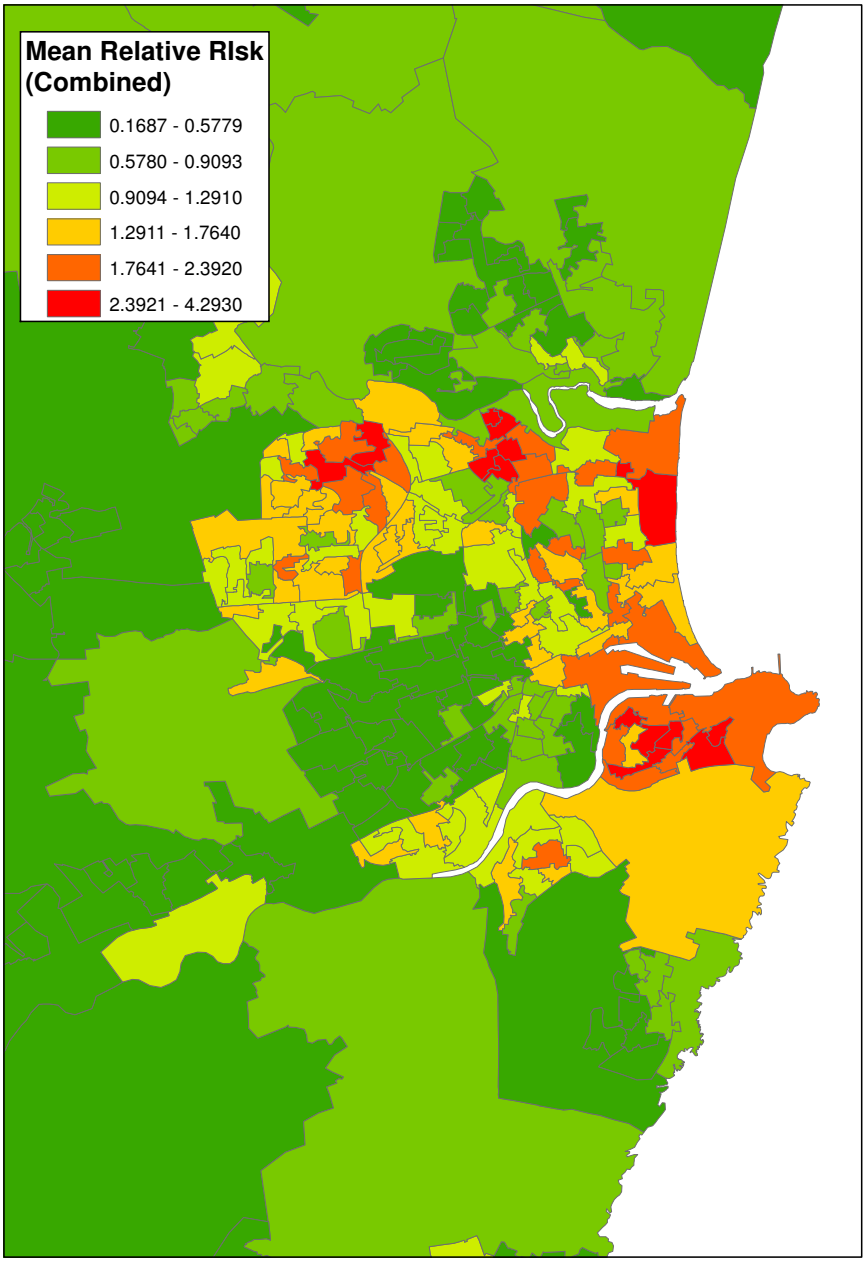


Figure 5.11: Aberdeen Area Data Zone Map of Mean Alcohol-Related Relative Risk

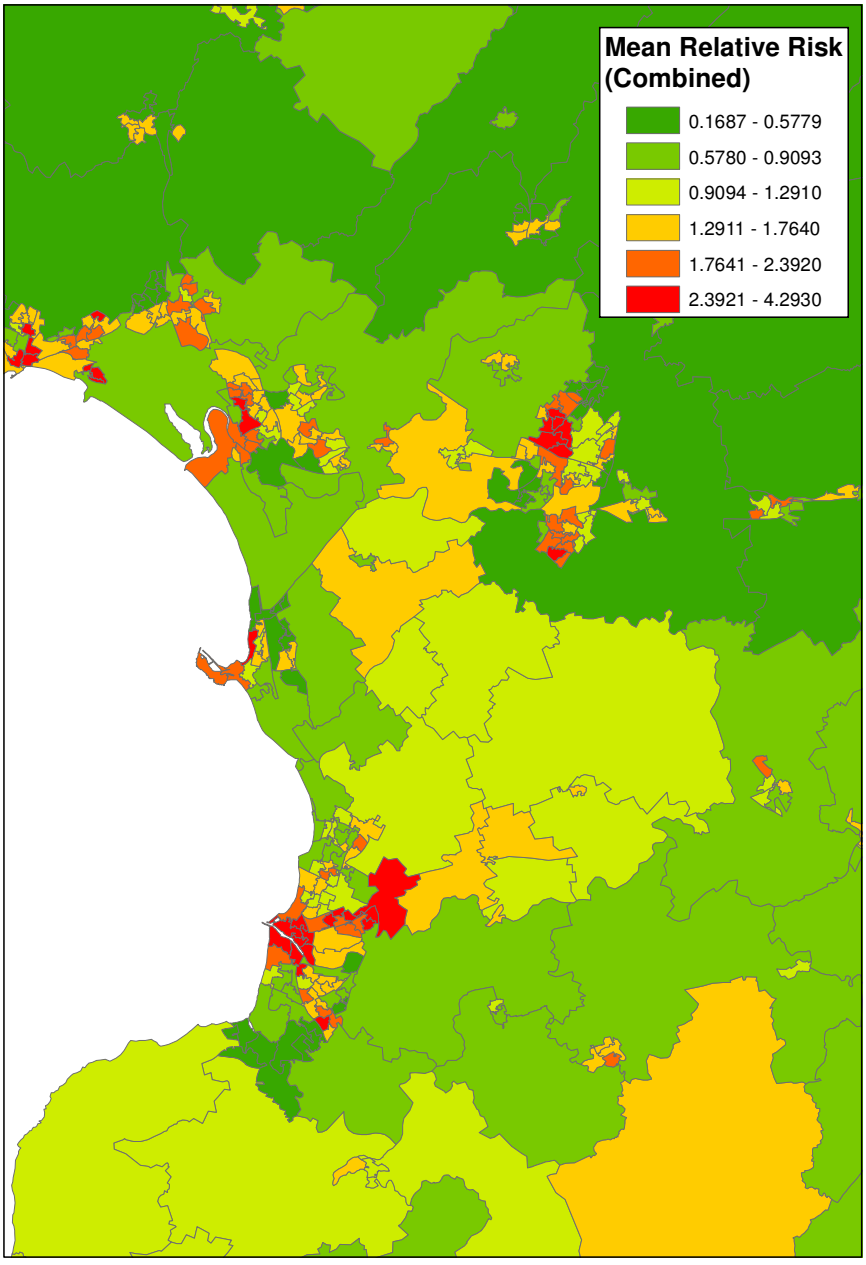


Figure 5.12: Ayrshire Area Data Zone Map of Mean Alcohol-Related Relative Risk

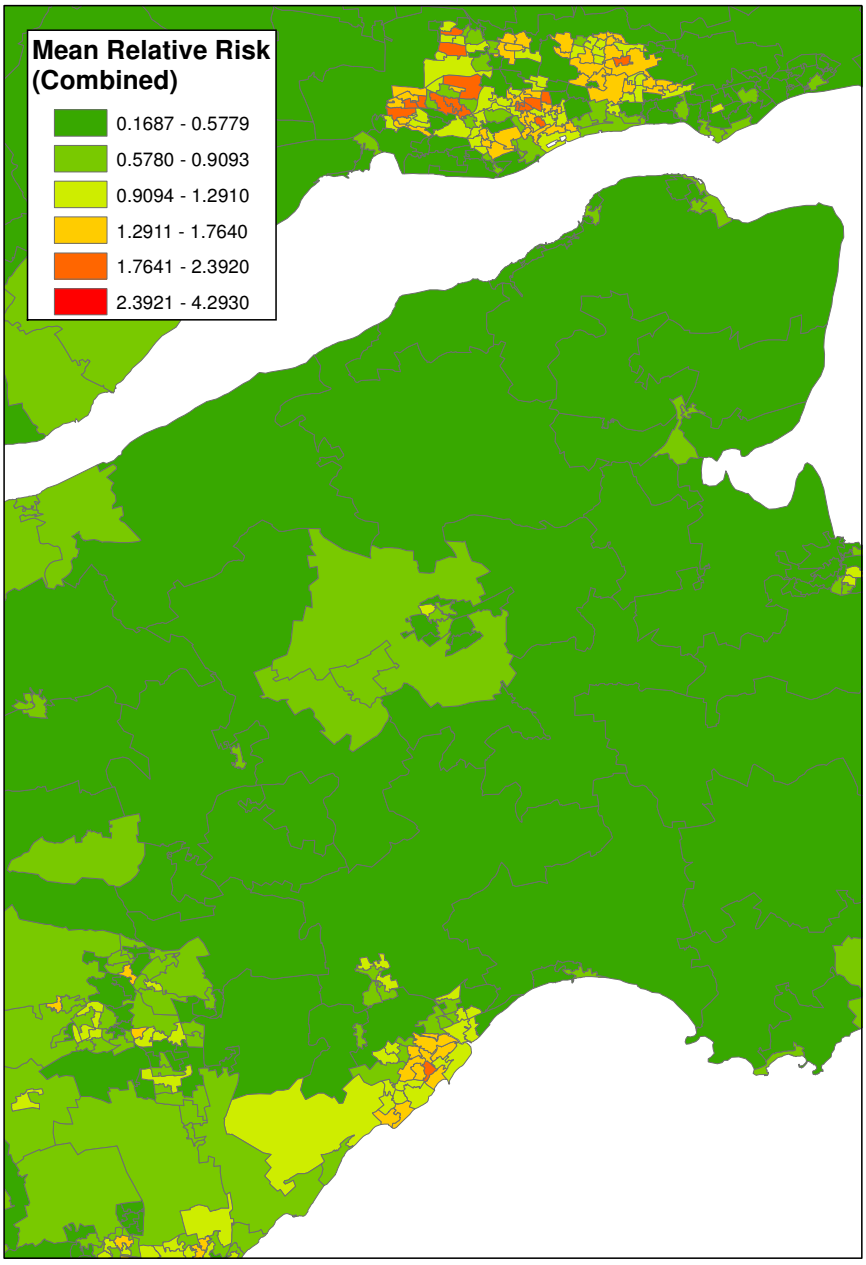


Figure 5.13: Dundee Area Data Zone Map of Mean Alcohol-Related Relative Risk

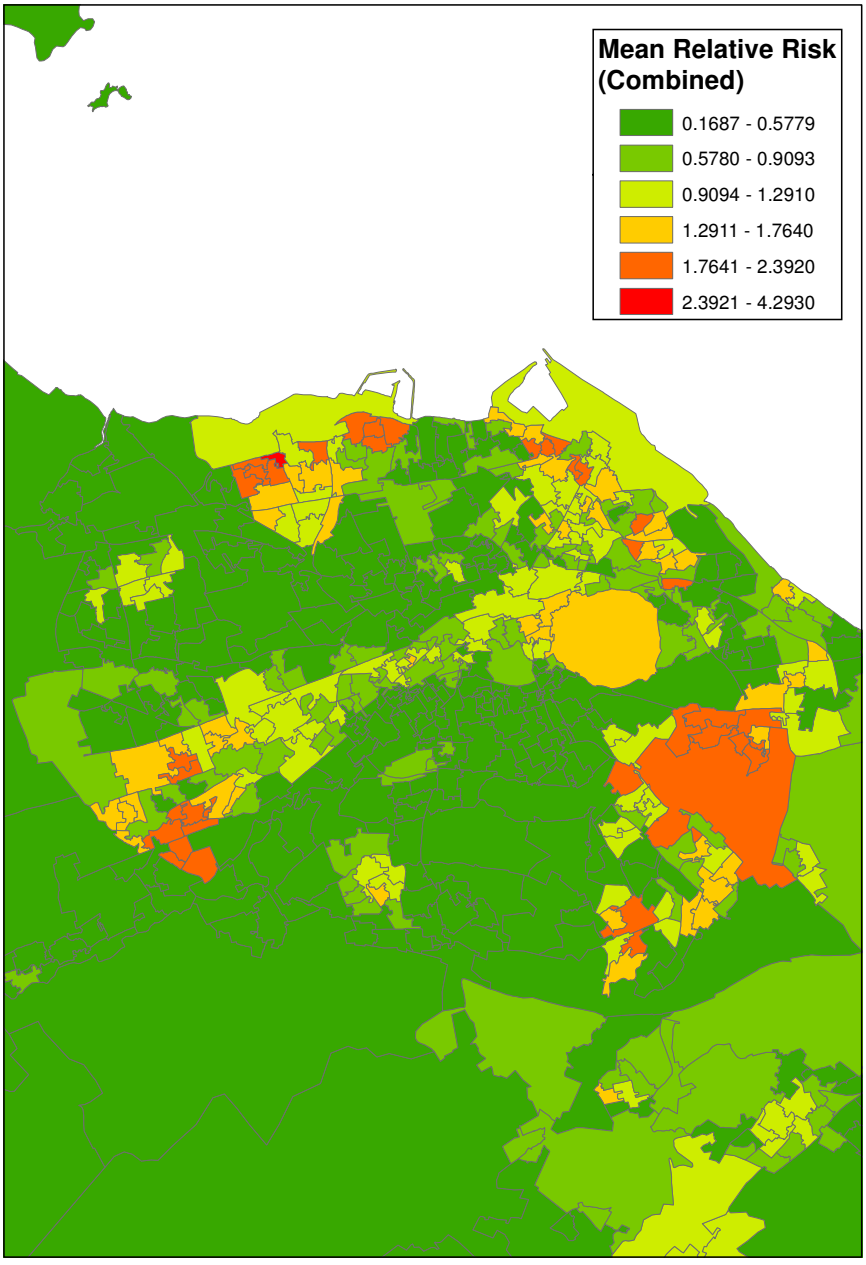


Figure 5.14: Edinburgh Area Data Zone Map of Mean Alcohol-Related Relative Risk

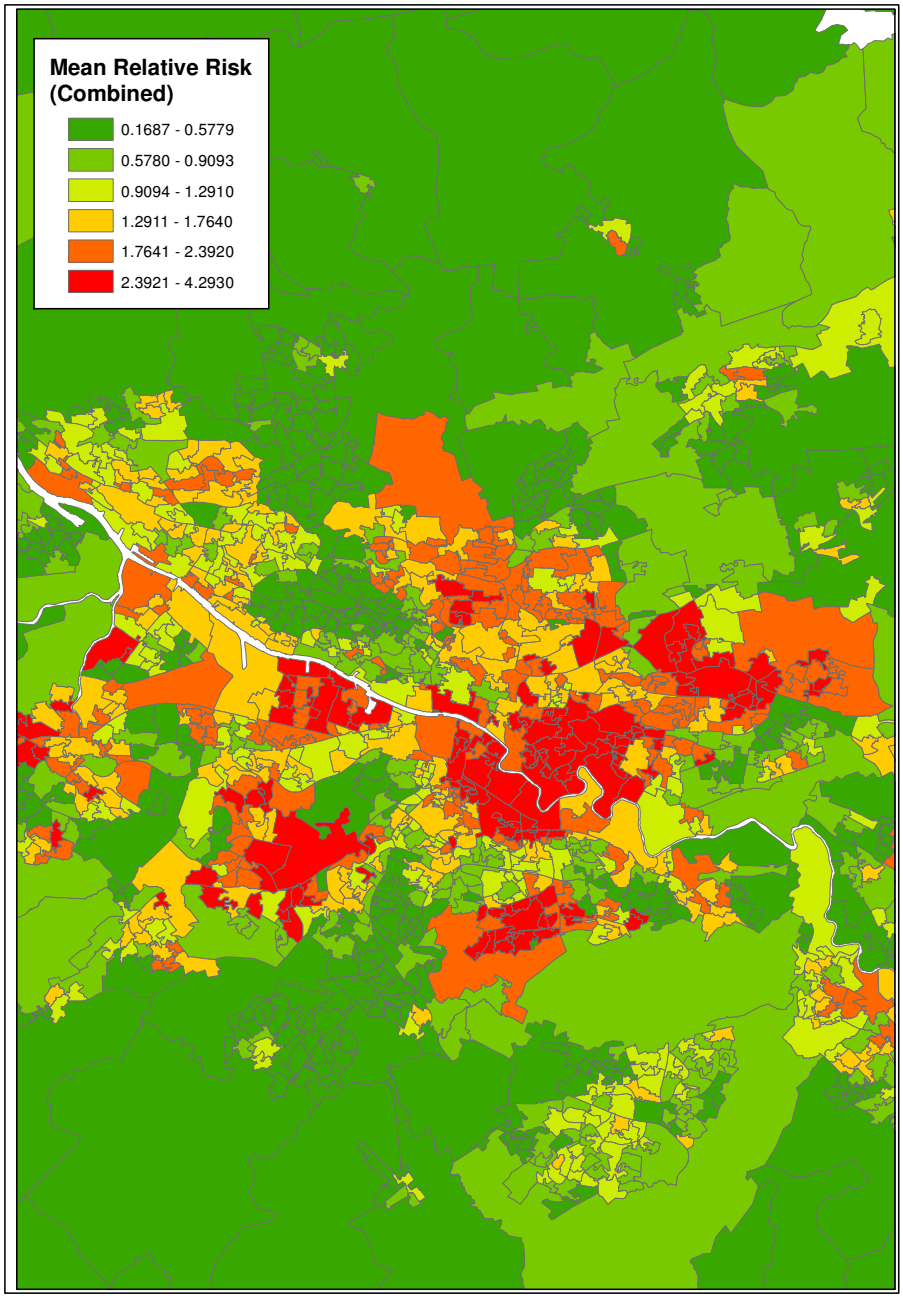


Figure 5.15: Glasgow Area Data Zone Map of Mean Alcohol-Related Relative Risk

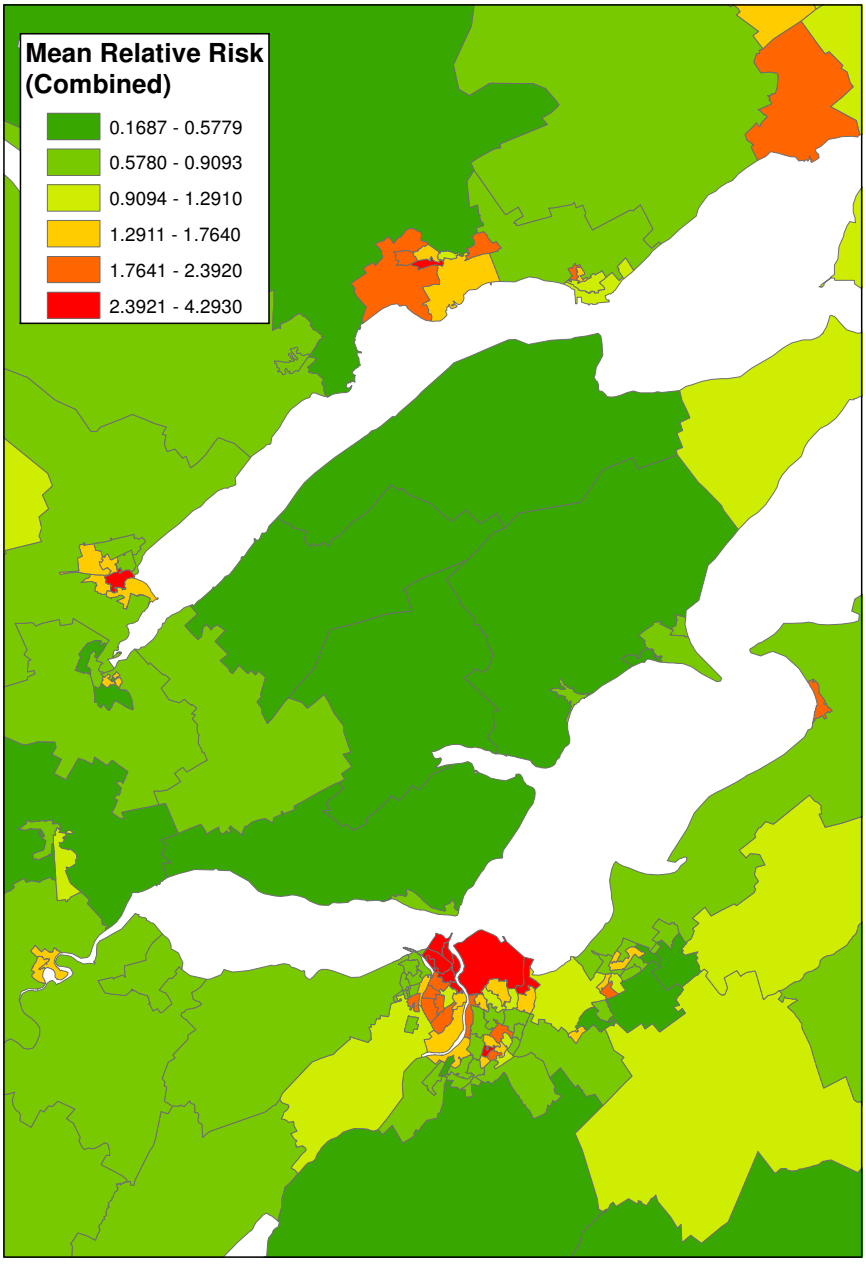


Figure 5.16: Inverness Area Data Zone Map of Mean Alcohol-Related Relative Risk

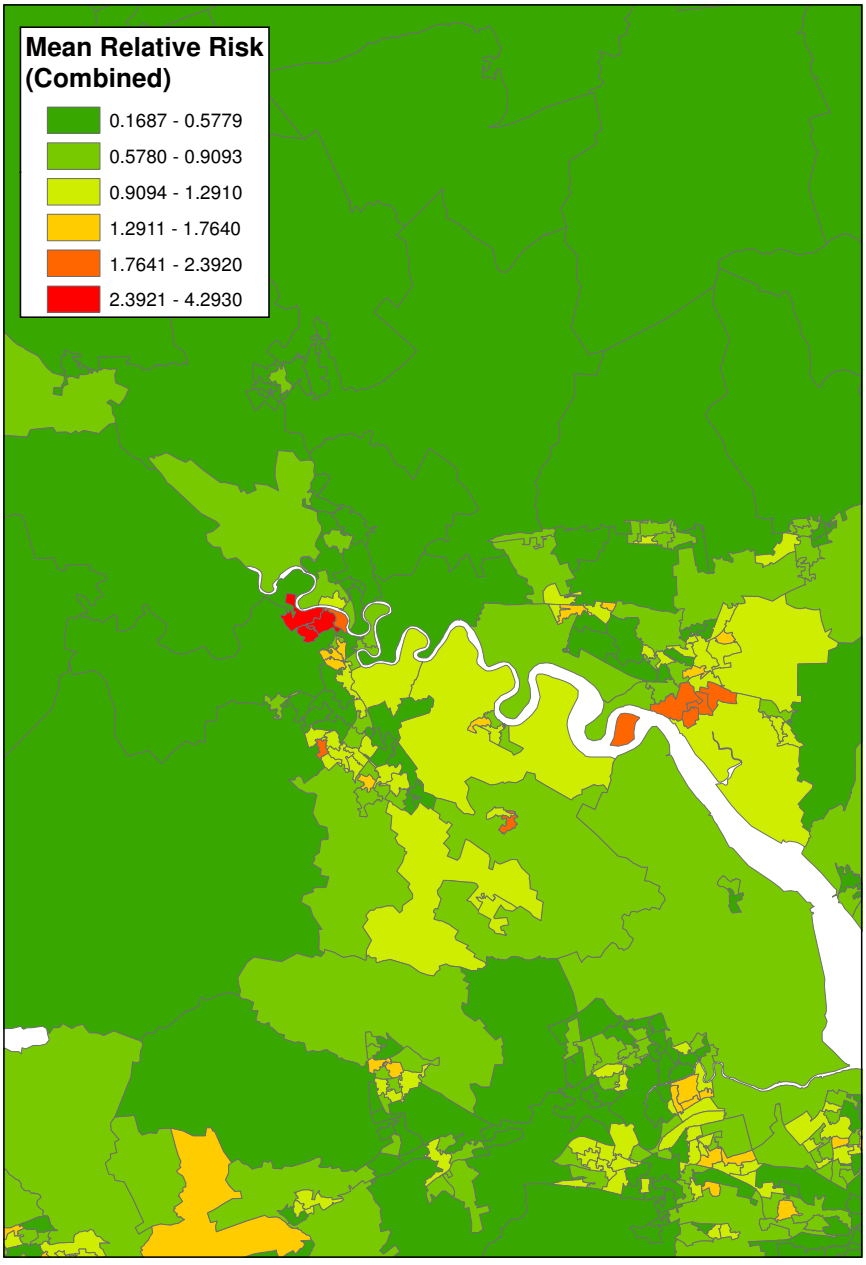


Figure 5.17: Stirling Area Data Zone Map of Mean Alcohol-Related Relative Risk

Chapter 6

BYM Models for Male Data

The previous Chapter investigates various spatial Bayesian models for the combined male and female relative alcohol-related risk across Scotland. As well as looking at the combined data, it is also of interest to analyse the male and female deaths and hospitalisations separately. Doing so will allow comparisons to be drawn between the results for each gender, as well as potentially providing stronger evidence of any relationships suggested. For example, if the chosen models for male, female and combined alcohol-related risk all suggest that area deprivation score is a significant factor, it allows one to be more confident in the models produced due to the consistency in their results.

This Chapter will consider several Bayesian models for male relative alcohol-related risk at the data zone level across Scotland.

6.1 Models Considered

The models considered for the male alcohol-related risks are exactly the same as those fitted to the combined data in Chapter 5, but now fitted to the male only death and hospitalisation data. The data used consists of every male alcohol-related death and hospitalisation in Scotland during years 2002 to 2006 inclusive. The expected number of occurrences in each data zone has

been calculated using age-standardisation as described in Chapter 3.

As was the case for the combined models, the male models discussed are inspired by the Besag, York and Mollié model (Besag et al. (1991)). The male models vary in terms of both fixed effects and random effects. The only fixed effect considered is a bona fide area deprivation score. This score is included in two different ways; either in a linear manner or by assigning a separate parameter to each of the ten deprivation scores. Again there are two separate random effects considered; correlated heterogeneity (u) and uncorrelated heterogeneity (v).

The nine male models considered are summarised with respect to the fixed and random effects included in Table 6.1 below.

Model Name	**Fixed Effects**	**Random Effects**
Male Model A-v	none	v
Male Model A-u	none	u
Male Model A	none	$u + v$
Male Model B-v	linear deprivation	v
Male Model B-u	linear deprivation	u
Male Model B	linear deprivation	$u + v$
Male Model C-v	non-linear deprivation	v
Male Model C-u	non-linear deprivation	u
Male Model C	non-linear deprivation	$u + v$

Table 6.1: Models for Male Alcohol-Related Relative Risks

Since these models are exactly the same as those considered for the combined data, the discussion in Section 5.1 regarding model structures, parameters and prior distributions still hold. The most important points will, however, be summarised here.

Vague normal priors with mean zero and precision e^{-5} have been assigned to all deprivation parameters with the exception of β_1 in the non-linear case, which has been arbitrarily set to zero.

The codes for all of the models specify a normal prior distribution with

mean zero for the uncorrelated heterogeneity and a conditional autoregressive prior for the correlated heterogeneity, so that

$$\begin{aligned} v_i &\sim N\left(0, \tau_v^2\right) \text{ and} \\ \left[u_i | u_j, i \neq j, \tau_u^2\right] &\sim N\left(\overline{u_i}, \tau_i^2\right) \end{aligned}$$

where τ_v^2, $\overline{u_i}$ and τ_i^2 are as described in section 3.6 of Chapter 3.

Vague Gamma hyperprior distributions have been assigned to the inverse variance hyperparameters of both random effects. In particular

$$\tau_v^2 \sim \text{Gamma}(0.5, 0.0005) \text{ and}$$
$$\tau_u^2 \sim \text{Gamma}(0.5, 0.0005).$$

This hyperprior distribution has been chosen since it is sufficiently vague and commonly used in disease mapping studies where there is no strong prior knowledge.

These male models were run using OpenBUGS and the code for Male Model A, Male Model B and Male Model C is shown in Appendices section 1.1, 1.2 and 1.3 respectively. Note that this is exactly the same code as was used for the equivalent combined models; the only difference lies in the data to which they were fitted. Again, the code for all other male models can be derived from this code by deleting any redundant sections; for example delete all code in Male Model A which relates to uncorrelated heterogeneity, v, in order to obtain Male Model A-u.

6.2 Convergence

As discussed in Section 5.1.2 when using any of the sampling methods discussed in Chapter 3, it is hoped that the joint distribution of the simulated Markov Chains will converge, or stabilise, to the joint posterior distribution.

Due to the large number of data zones, and hence parameters in each model, it was not practicable to store the simulated value of every parameter

at each iteration. Instead the central model parameters, along with a chosen subset of the data zone relative risk parameters, have been fully monitored. The data zone relative risk parameters chosen to be monitored, along with reasons why they were chosen, are shown below in Table 8.2. The remaining parameters had a summary monitor set. A summary monitor means that at each iteration the summary statistics for that variable are updated, but that the simulated parameter value itself is not stored. This results in exact estimates of the mean and standard deviation of the simulated parameter sample, but only approximate 95% credible intervals.

Data zone Code	**Relative Risk Parameter**	**Reason Chosen**
S01006393	θ_{115}	poor deprivation score
S01006438	θ_{14}	poor deprivation score
S01006490	θ_{1}	good deprivation score
S01006505	θ_{2}	good deprivation score
S01003744	θ_{2521}	rural area
S0100391	θ_{2692}	rural area
S01003380	θ_{3044}	urban/city area
S01002325	θ_{4687}	urban/city area
S01005521	θ_{985}	island / no neighbouring areas
S01000447	θ_{6238}	island / no neighbouring areas
S01003031	θ_{2885}	lowest total population
S01000799	θ_{5792}	highest male population
S01002622	θ_{3557}	highest female population
S01003313	θ_{3046}	highest male SIR
S01006473	θ_{89}	zero male SIR value
S01006341	θ_{172}	zero male SIR value
S01003043	θ_{2889}	very high male SIR

Table 6.2: Data Zones with Fully Monitored Male Relative Risk Estimates

Convergence was monitored for all male alcohol-related risk models and it was found that all appeared to converge well after a burn-in period of 50,000 iterations. For each model two chains were run for a further 150,000 post-burn-in simulations. The resulting diagnostic plots from all nine models

suggest that this is a long enough chain length. Since the same convergence checks were carried out and satisfied for all male models, they will only be discussed in detail for Male Model C-u.

Firstly, the sample history plots for each of the fully monitored parameters in Male Model C-u are shown in Figures 6.1 to 6.6. For each of the chosen parameters, a line plot of simulated value against iteration number is shown. The simulation path for one chain is shown in blue while the second chain is shown in red. These plots all show that for the duration of the 150,000 post-burn-in iterations the simulated parameter values for each chain consistently overlap. Moreover, none of the plots exhibits any obvious patterns or trends and the simulated values form horizontal bands across each plot. Although these plots can only be looked at for a subset of the total model parameters, they give strong evidence that the model has converged. Since two different chains with different starting points are consistently giving similar values, it suggests that the chains have in fact settled to the appropriate posterior distribution.

For the same subset of Male Model C-u parameters the Gelman-Rubin diagnostic plots, as discussed in section 3.4.3, are shown in Figure 6.7 and Figure 6.8. All of these plots also suggest that more than adequate convergence has been achieved. This is because the green line, which represents the width of the central 80% interval of the pooled chains, and the blue line, which shows the average width of the 80% intervals within the individual chains, are both stable and the red line which shows their ratio has settled to a value of 1. As was the case for the combined models, the intervals are so similar for the individual chains and the pooled chains that the blue line almost completely obscures the green line.

A further visual check of model convergence is to look at the sample probability density plots. These are shown for the same selection of Male Model C-u parameters in Figure 6.9 and Figure 6.10. Each of these plots shows a smooth transition in probability between different parameter values.

These plots also suggest that the simulated chains have been run for a long enough post-burn-in period. Even in cases where models have been shown to converge, if enough iterations are not run, these density plots can appear uneven and 'lumpy' in appearance.

The evidence presented above gives strong evidence that Male Model C-u has achieved adequate convergence after an initial 50,000 iterations and a further 150,000 iterations for two simulation chains. This was also found to be the case for the other 8 alcohol-related male relative risk models.

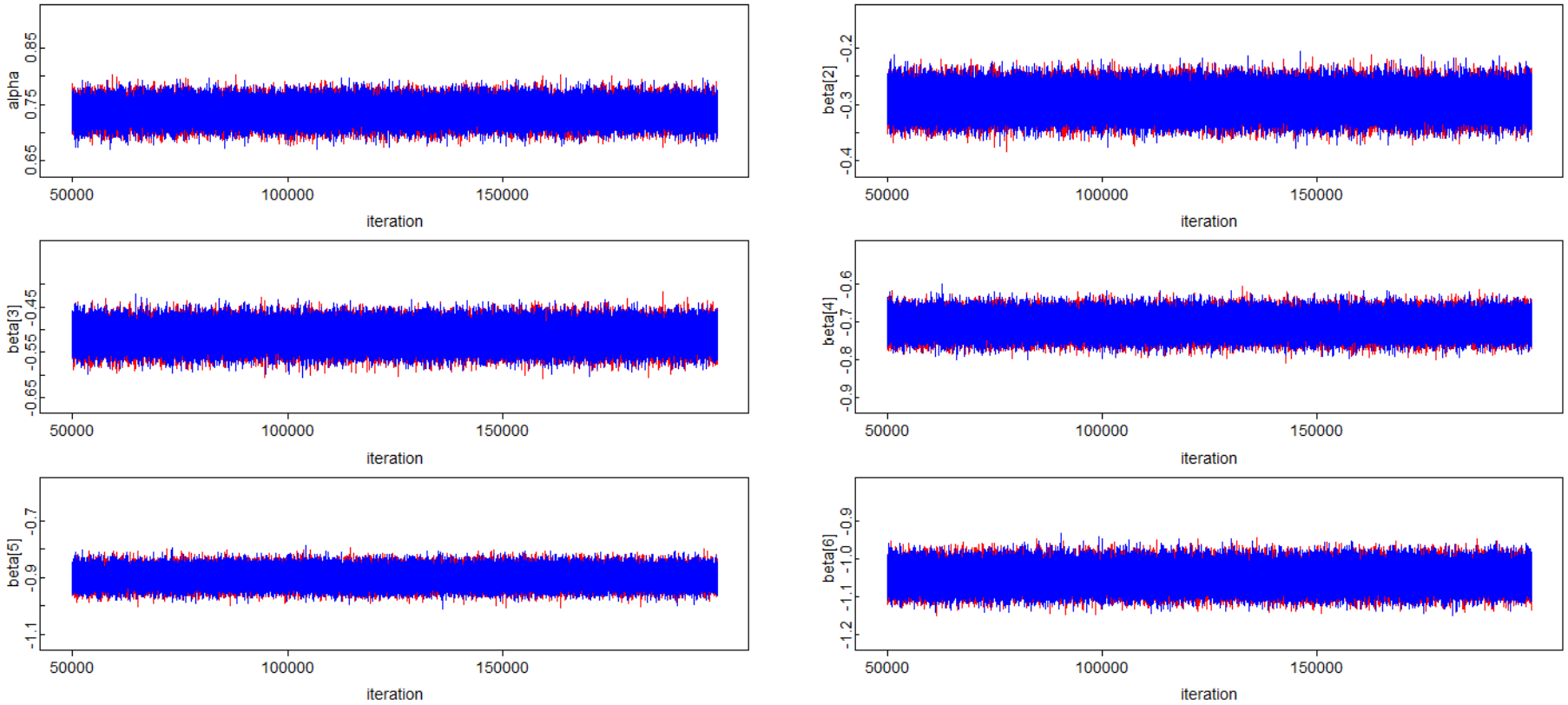


Figure 6.1: Simulation History Plots for a Subset of Male Model C-u Parameters (part 1)

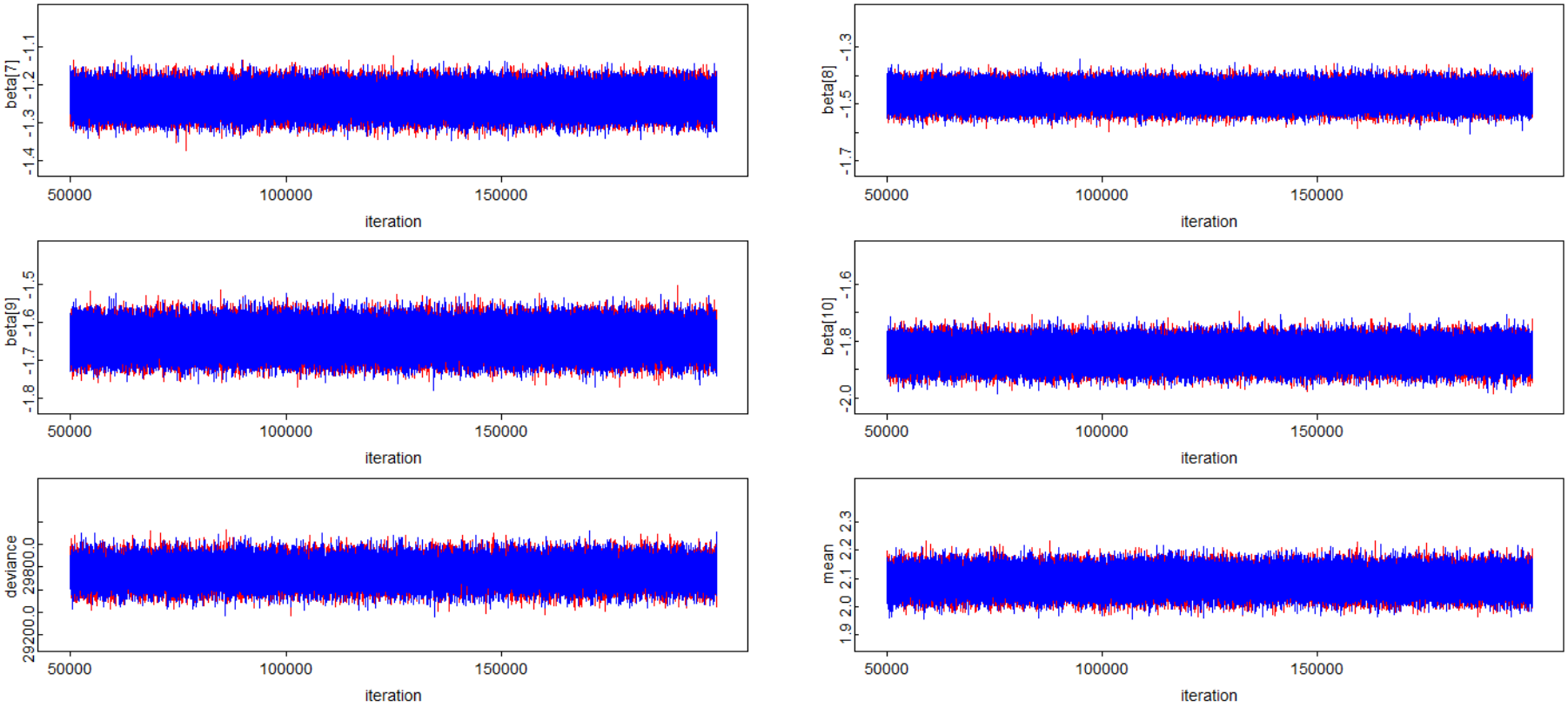


Figure 6.2: Simulation History Plots for a Subset of Male Model C-u Parameters (part 2)

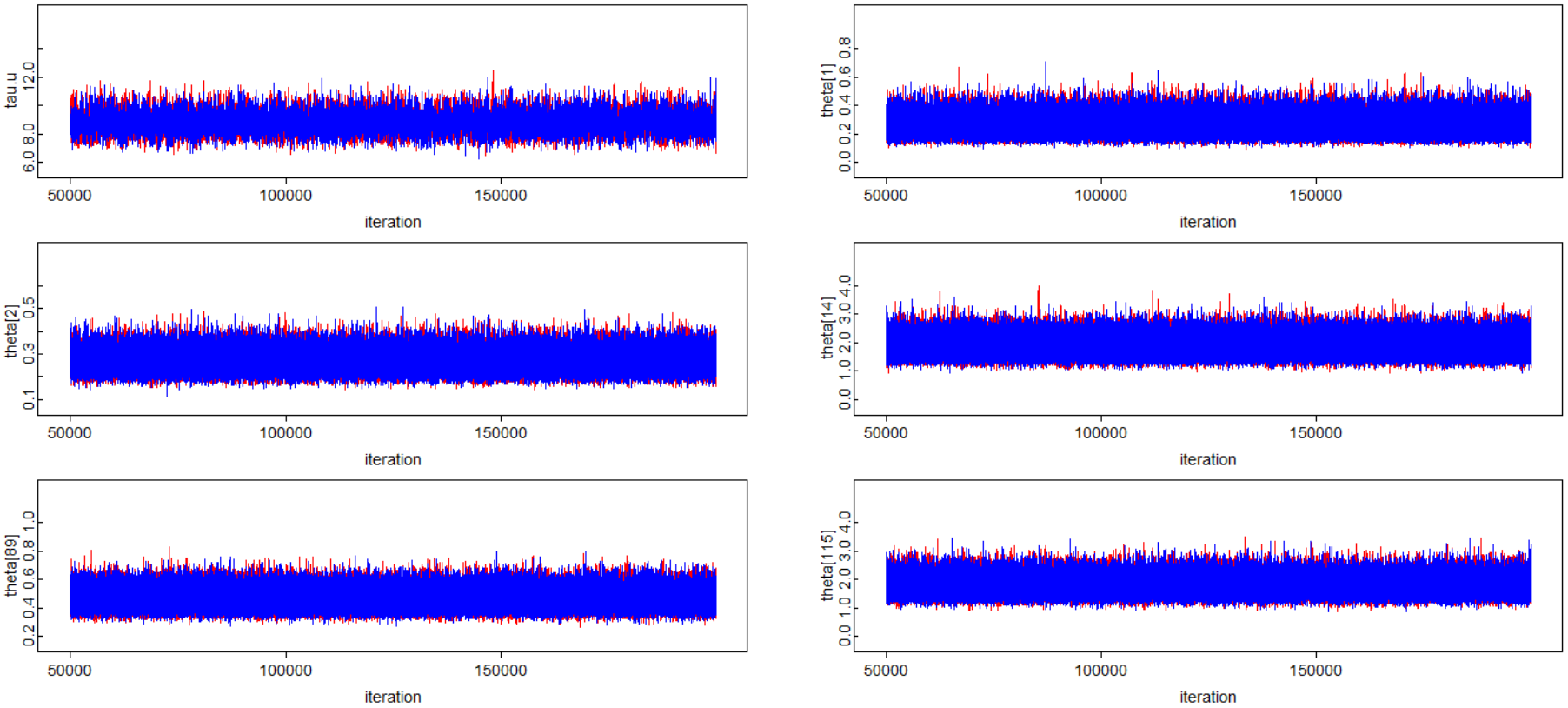


Figure 6.3: Simulation History Plots for a Subset of Male Model C-u Parameters (part 3)

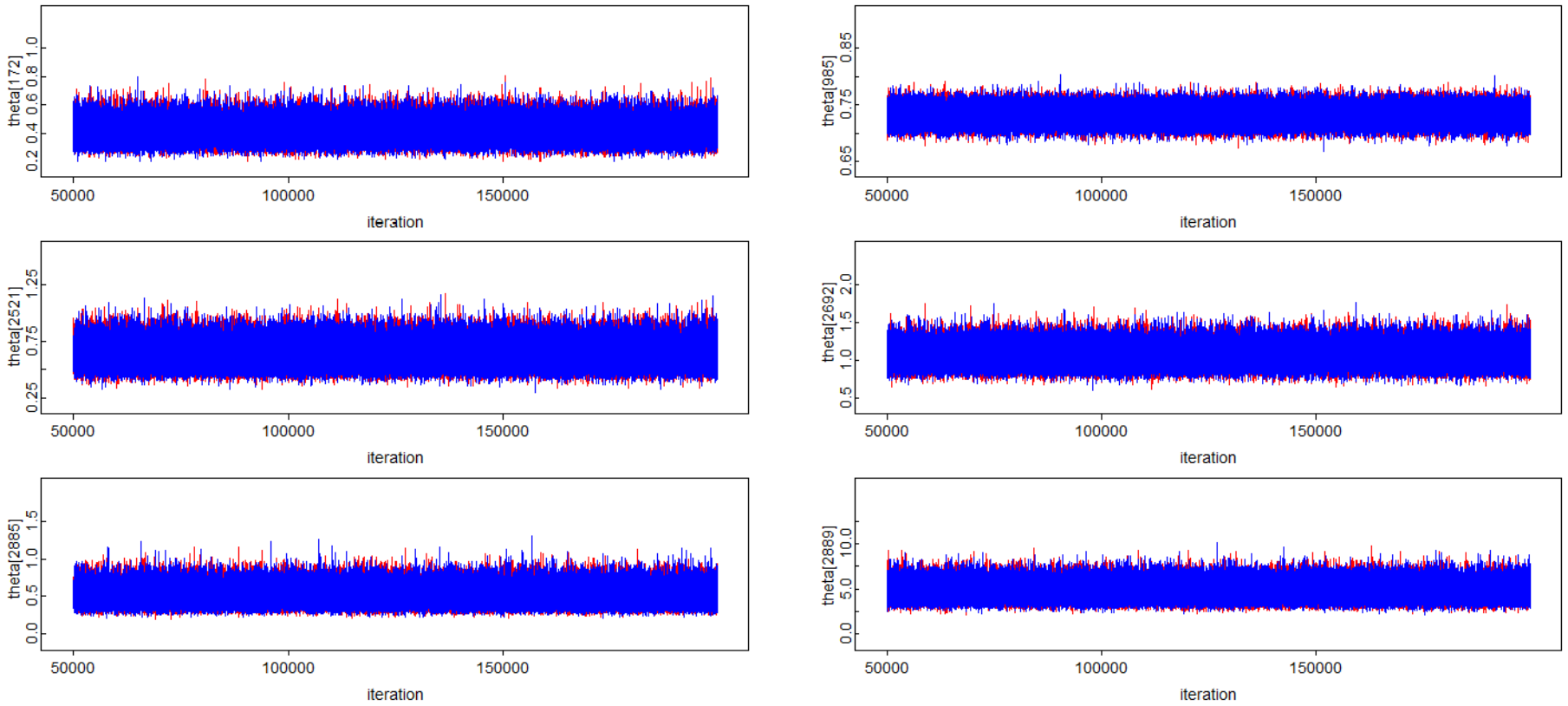


Figure 6.4: Simulation History Plots for a Subset of Male Model C-u Parameters (part 4)

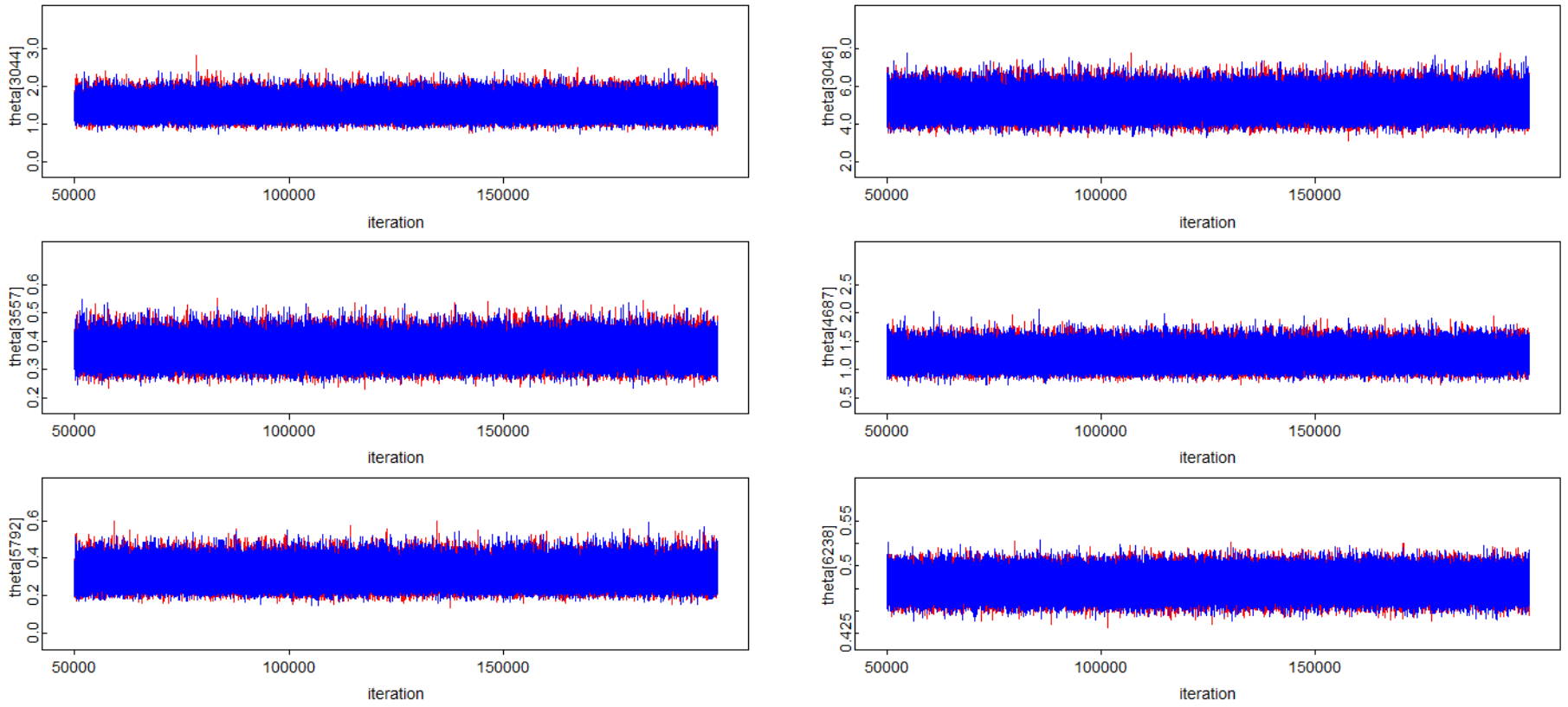


Figure 6.5: Simulation History Plots for a Subset of Male Model C-u Parameters (part 5)

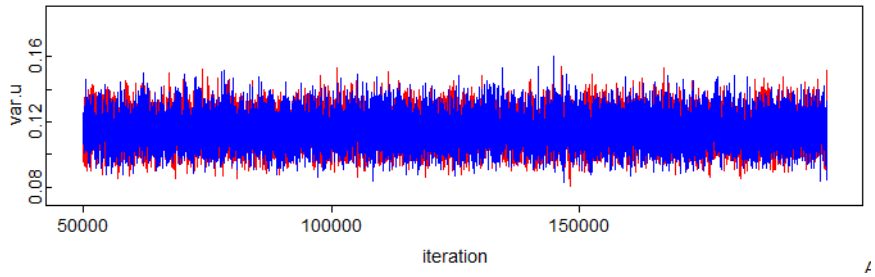


Figure 6.6: Simulation History Plots for a Subset of Male Model C-u Parameters (part 6)

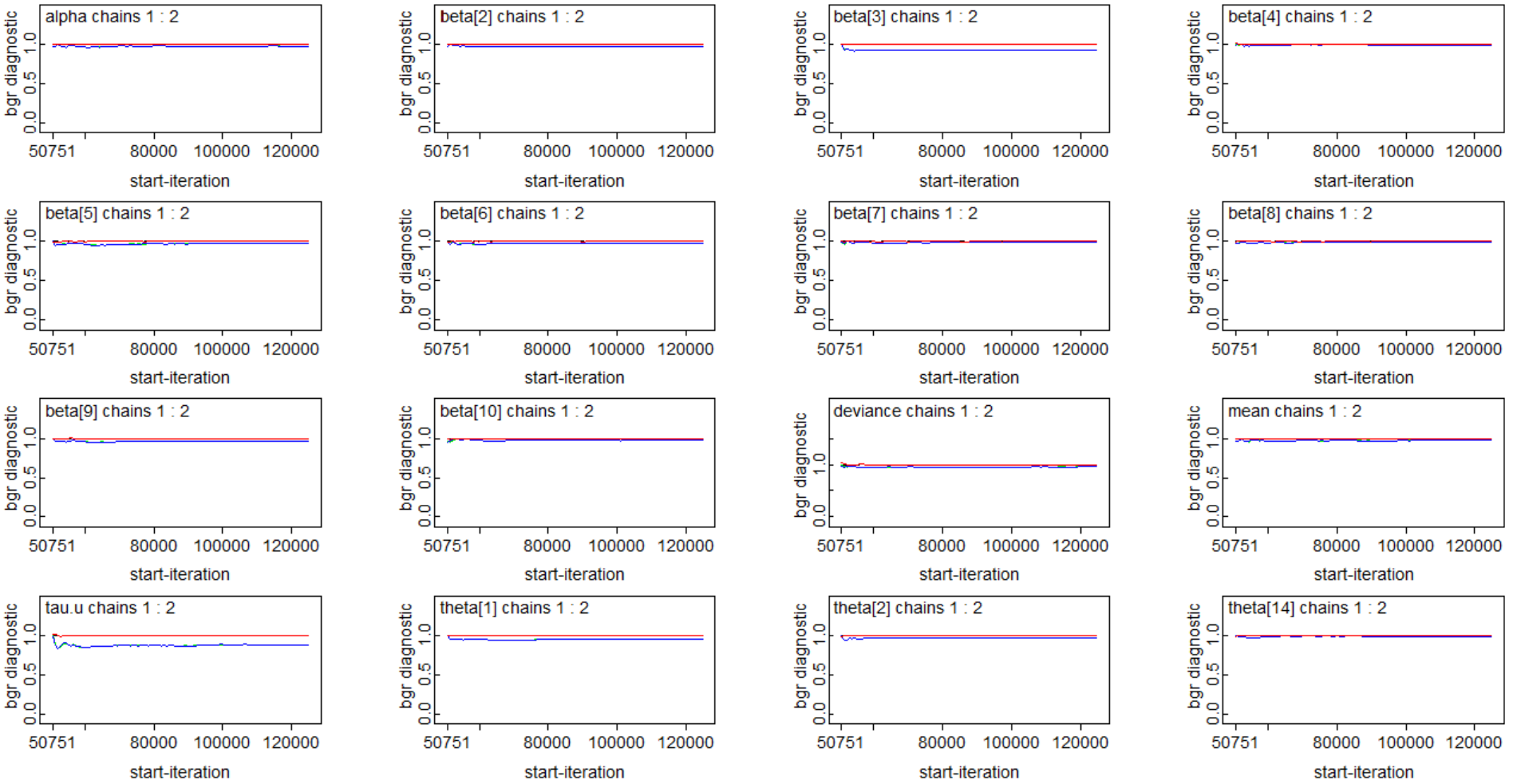


Figure 6.7: BGR Diagnostic Plots for a Subset of Male Model C-u Parameters (part 1)

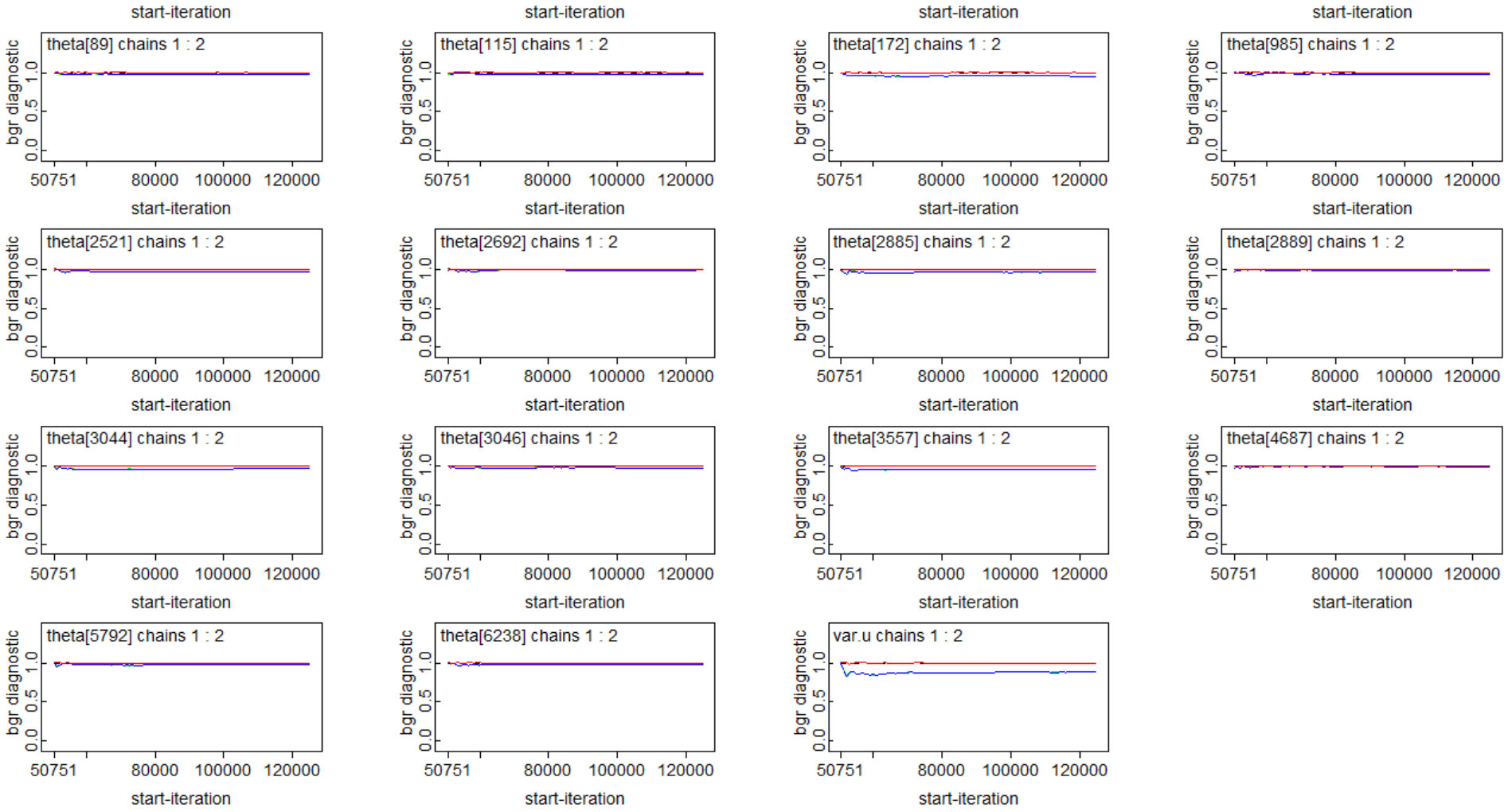


Figure 6.8: BGR Diagnostic Plots for a Subset of Male Model C-u Parameters (part 2)

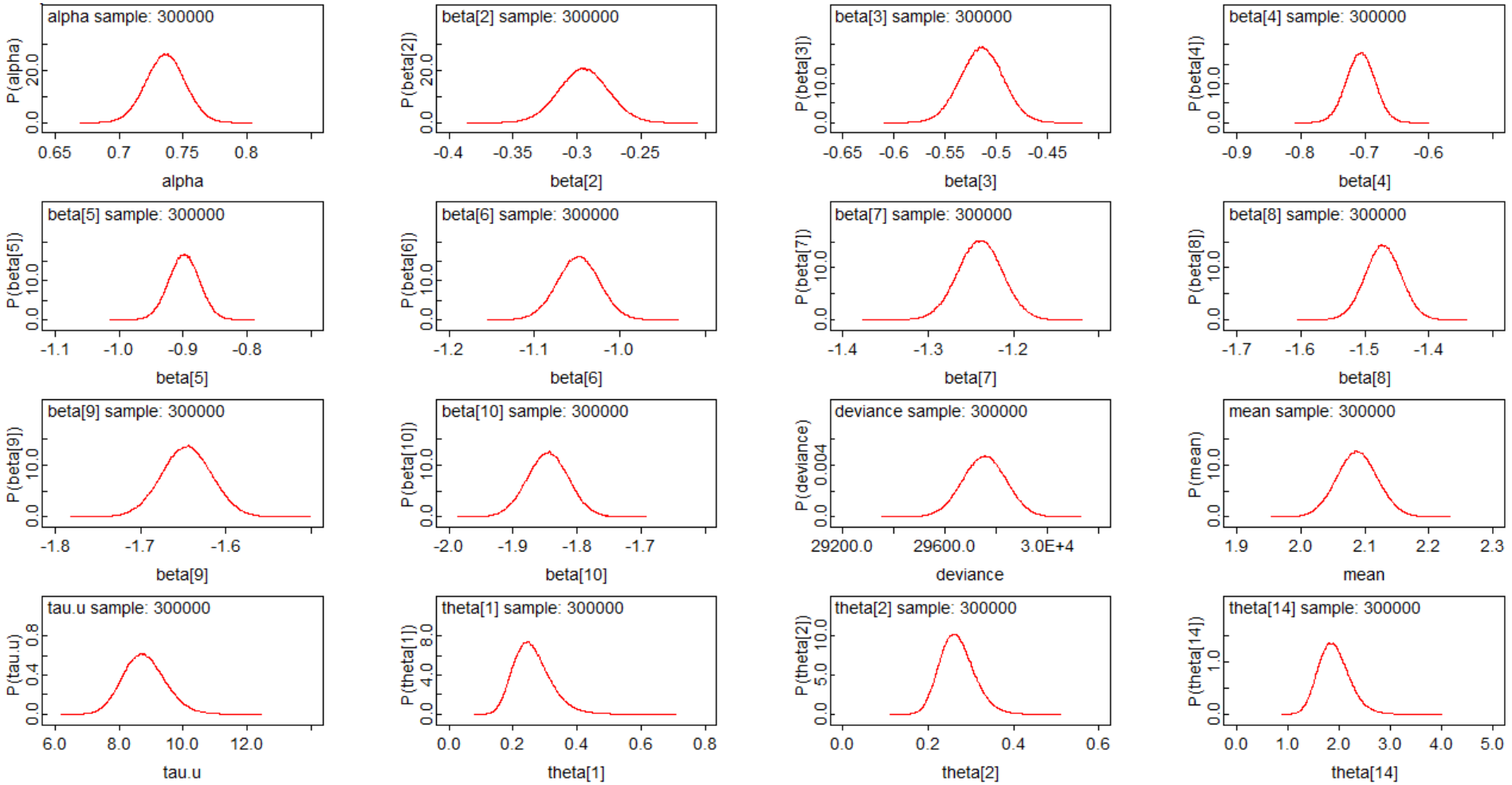


Figure 6.9: Posterior Density Plots for a Subset of Male Model C-u Parameters (part 1)

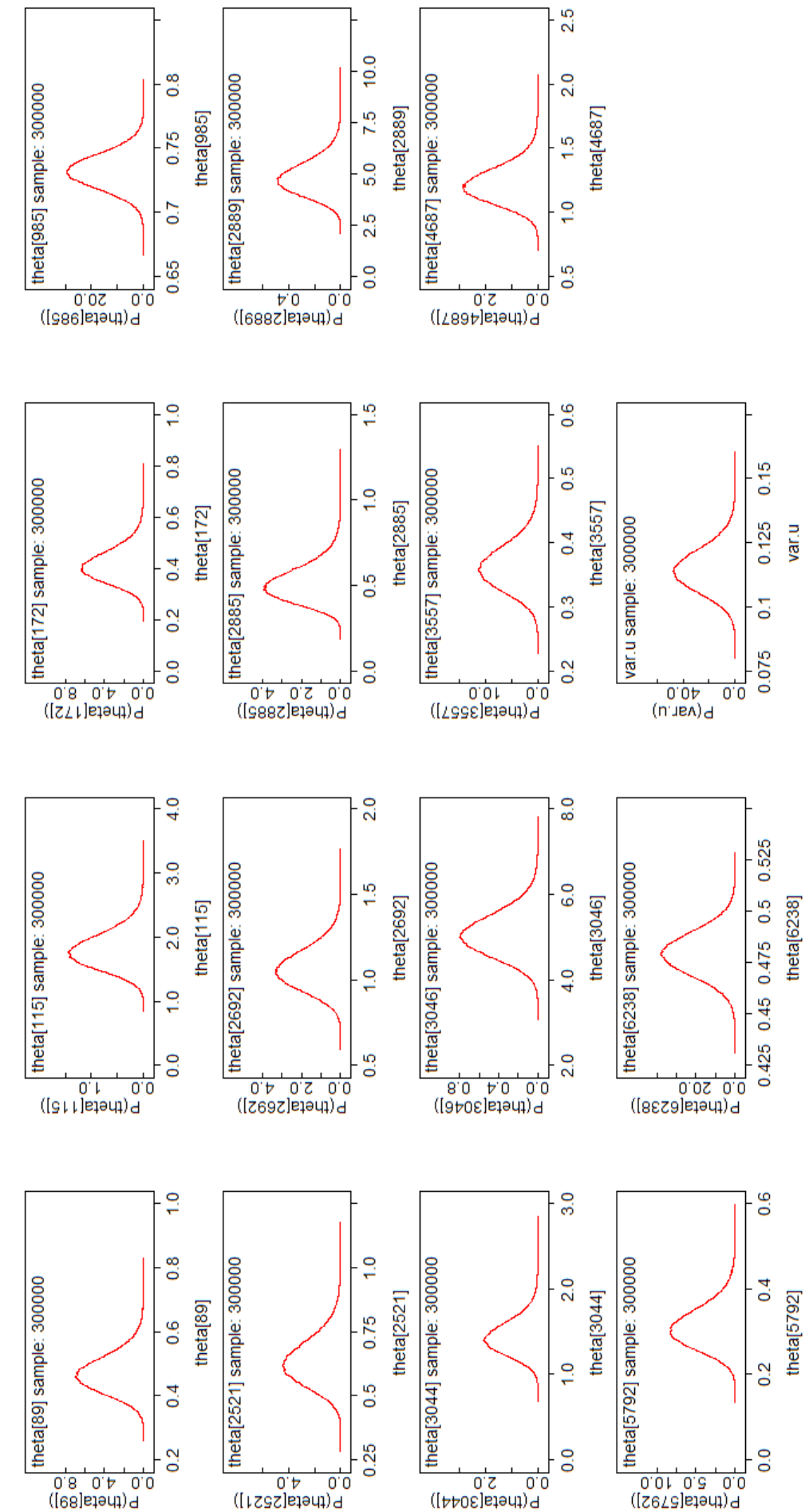


Figure 6.10: Posterior Density Plots for a Subset of Male Model C-u Parameters (part 2)

6.3 DIC

Now that all nine male models have been run and are deemed to have achieved satisfactory convergence it is necessary to choose between them. As discussed in Chapter 3, the Deviance Information Criterion, DIC, is a commonly used measure of goodness of fit for spatial Bayesian models such as those discussed in this chapter.

Table 6.3 gives the DIC for each male model calculated using the pD method, along with the corresponding pD and deviance values. In this table the model results are are sectioned in two ways: firstly, split by the fixed effects they contain, either no deprivation, linear deprivation or non-linear deprivation, and secondly, the split by whether the models contain only correlated heterogeneity (u), uncorrelated heterogeneity (v) or both ($u + v$).

By looking at this table it is immediately obvious that the problem of negative pD values experienced for the combined models is also an issue for the male models. There are three instances of negative pD values here, all in models with contain an element of spatial correlation between data zones. As was discussed in the previous chapter, negative pD values are a possible but very undesirable phenomenon. The value pD is supposed to represent the effective number of parameters in the model, and so obviously this value should not be negative. Such negative values can occur when there is a non-log-concave likelihood, when the posterior for a parameter is especially asymmetric or bimodal or when there is another situation that causes the posterior mean to be a poor summary statistic causing large deviance. The diagnostic plots discussed above in the Convergence section, along with the equivalent plots for the other male models, showed no evidence of particularly asymmetric or non-unimodal distributions, and there is no other obvious reason why the posterior mean should be a poor parameter estimate.

The lowest DIC value observed in Table 6.3 is for Male Model A which includes both correlated and uncorrelated random effects and no fixed effects.

This model choice is definitely not what one would expect given the strong similarities exhibited between the deprivation maps and the male SIR maps discussed in Chapter 4. From previous subjective impressions it is highly likely that area deprivation score will account for some of the variation in data zone male alcohol-related risk across Scotland. Also, as the WinBUGS website suggests, $\widehat{D}$ is a better measure of fit than $\overline{D}$ which can be considered more a measure of adequacy. This means that Table 6.3 suggests that Male Model A-v, which contains only uncorrelated heterogeneity and no covariates, fits the male data best. This too is a very unlikely outcome given previous strong indications of a link between male alcohol-related risk and deprivation levels.

Due to the negative pD values and dubious model choice in which they result, it has been decided to instead calculate DIC using the p*D method. This is the same method which was used in the previous combined model chapter. The p*D method is discussed in Chapter 3, and bases the estimated number of effective parameters, p*D, on half the variance of the model deviance.

The DIC values for the male models of alcohol-related relative risk calculated using the p*D method are shown in Table 6.4 below. Obviously, since the estimates of p*D are just half the variance of the model deviance, they are all positive. The lowest DIC using this method is for the model which includes non-linear deprivation and only spatial heterogeneity, Male Model C-u. This is a much more believable and reasonable sounding model choice. Furthermore, this is the same model structure as was chosen for the combined male and female data, which strengthens confidence in any inferred relationships.

6.4 Male Model Selection

In normal circumstances the model with the lowest DIC would be chosen as the 'best' model. For the models of male alcohol-related relative risk in Scotland, this would be Male Model C-u. The selection of this model indicates that the relationship between alcohol-related relative risk and area deprivation is not linear. This is consistent with subjective impressions given in Chapter 4, which suggested that there was a greater increase in average male SIR value between the deprivation scores of 1 and 2 than between any other pair of consecutive scores. Male Model C-u also includes spatial or correlated heterogeneity, which seems reasonable since in the male SIR maps shown in Chapter 4 similar values do tend to cluster together, even though the maps were not overly smooth.

It seems then, that using DIC calculated via the p*D method leads to a reasonable model selection for the male risks. However, as was the case for the combined models, it must be remembered that the male models are based on assumptions which are set by the modeller and incorporated via the prior and hyperprior distributions for the parameters and hyperparameters. Although Male Model C-u seems a reasonable choice, it is possible that the same model structure would not be chosen if different prior and hyperprior distributions were considered. Since there has been limited time to complete this project, a full and comprehensive sensitivity analysis was not possible. However, model sensitivity to the Gamma (0.5,0.0005) hyperprior distributions specified for the inverse variance parameters τ_u^2 and τ_v^2 will be examined.

6.5 Male Hyperprior Sensitivity Analysis

All male models listed in Table 6.1 will be re-run with different hyperprior distributions assigned to τ_u^2 and τ_v^2. The names used for these sensitivity models will be the same as those given in Table 6.1 but with 'Sens' appended at the end, e.g. 'Male Model A-Sens'. The main reason for running the

models with different hyperpriors is to see whether the same model structure would be selected. It is also of interest to see how parameter estimates and credible intervals are affected even if the same model structure is chosen.

The alternative priors used are the same as those used for the combined models. In each model the following hyperpriors will be used where necessary

$$\tau_u^2 \sim \text{Gamma}(1,1) \text{ and}$$
$$\tau_v^2 \sim \text{Gamma}(1,1).$$

Again, these distributions are much less vague and very different from the original hyperpriors used. This would not be the ideal distributions to use as first choice for these models. However, if the male models can be fitted using two very different hyperprior distributions and still give similar results, this would provide strong evidence that the models are not too sensitive to hyperprior choice.

Convergence of all male sensitivity models was monitored and checked, with all converging well after 50,000 iterations. As with the original male models, two chains were then simulated for a further 150,000 iterations. This resulted in samples of 300,000 simulated values for each fully monitored parameter.

Once all of the male sensitivity models had been run the DIC values were calculated. Table 6.5 gives the deviance, pD and DIC values calculated using the pD method for each male sensitivity model. This table shows that the negative pD issue discussed earlier in the DIC section is still present in the male sensitivity models. Again, it only affects those models which incorporate correlated heterogeneity, u, in some way. The lowest DIC value corresponds to a model which includes no fixed effects, which due to strong suggestions of a link between deprivation score and male alcohol-related risk seems dubious.

Due to the negative pD values and questionable model selection obtained, it has been decided to instead calculate DIC using the p*D method seen

before. The DIC values obtained via this method are given in Table 6.6 along with the deviance statistics and p*D values required to calculate these figures.

Since the p*D estimate of the effective number parameters is a positive proportion of the deviance variance, all p*D values are positive. Using this method the lowest DIC value corresponds to a Male Model C-Sens, which fits non-linear deprivation and both correlated and uncorrelated heterogeneity to the male data.

Changing the hyperpriors used has therefore led to a slight difference in model selection. From the original male models, Male Model C-u was chosen, which includes non-linear deprivation and only correlated heterogeneity. Under both the original and sensitivity methods then, both chosen models suggest that there is a non-linear relationship between male alcohol-related risk in Scotland and deprivation score, and both incorporate a spatial random effect. Given how different the hyperprior distributions assigned were, this slight difference does not pose a large problem.

It is also of interest to compare the parameter estimates between the chosen original and sensitivity models. The estimates and 95% credible intervals for all fully monitored parameters in Male Model C-u and Male Model C-Sens are given in Table 6.7. Ignoring the variance and precision parameters, for all but one of the remaining parameters, the estimate from each model lies within the corresponding credible interval from the other. The one parameter for which this does not hold is the relative risk parameter θ_{3046}. For this parameter the estimate from the sensitivity model does not lie within the credible interval from the original model. The two estimates for this parameter do seem a little different. However, the estimate from the sensitivity model does not lie too far away from the upper bound of the credible interval from the original model. I do not think that this difference should cause too much alarm. The parameter in question corresponds to the data zone which experienced the highest male SIR value, but the risk parameter

which corresponds to the data zone with the second-highest male SIR, θ_{2889}, did not experience such a problem.

Obviously, since the Male Model C-u includes only correlated heterogeneity when Male Model C-Sens includes both correlated and uncorrelated heterogeneity the variance parameter estimates will not be the same. However, even though Male Model C-Sens contains both spatial and non-spatial random effects, it assigns over 71% of the variation to correlated heterogeneity. The need for the additional non-spatial random effect will most likely be because the priors assigned to the precision terms of both random effects in the sensitivity models are much more restrictive. The original precision priors specify a mean of 1000 and a variance of 2,000,000 compared to both a mean and variance of 1 for the sensitivity models.

Given the strong similarities between the Male Model C-u and Male Model C-Sens results for male alcohol-related relative risk across Scotland, Male Model C-u will be chosen as the final model since it has a much more appropriate and vague hyperprior distribution and it has been shown not to be very sensitive to hyperprior choice.

6.6 Male Model Results

A selection of the parameters fitted in the final male relative risk model, Male Model C-u, are shown in Table 6.7. None of the 95% credible intervals for the nine deprivation scores overlaps or contains zero. This strongly suggests that all deprivation scores have a significant effect on male alcohol-related relative risks in Scotland and hence should be included in the model. This evidence further supports the model choice; if the DIC suggested this model but it was shown that the individual deprivation score parameters were not significant, this could potentially lead to an alternative model selection. Boxplots of the simulated samples of area deprivation parameter values from Male Model C-u are shown in Figure 6.11. Although the chosen

model does not contain linear deprivation, the individual area deprivation score parameters themselves appear to be fairly linear, with a larger jump in value between deprivation scores 1 and 2. These boxplots also show that the modelled relationship between deprivation and risk is monotonic.

All of the deprivation score parameter estimates are negative and the value of β_{d_i} gets progressively smaller as d_i increases from a score of 2 to 10. This was also observed for the combined models and is what one would expect; there is a larger decrease in the relative risk estimates for less deprived areas. The chosen model suggests it is highly likely that, on average, the least deprived data zones with an area deprivation score of 10 have a male alcohol-related relative risk which is between only 0.148 and 0.168 times that of the most deprived areas.

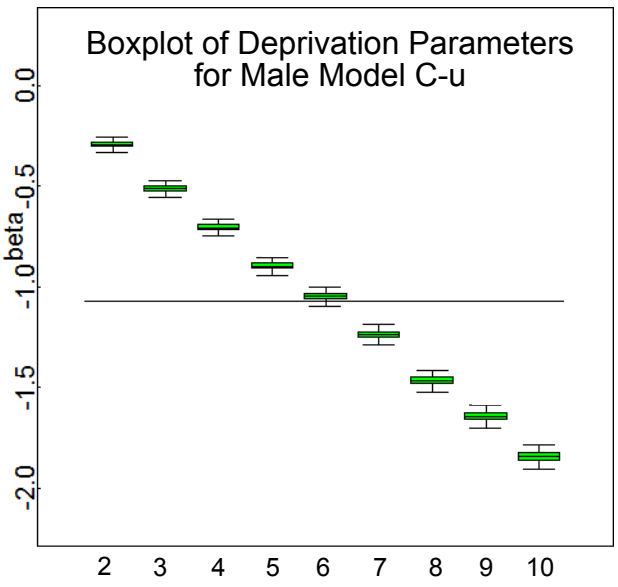


Figure 6.11: Boxplots of Deprivation Parameters for Male Model C-u

Since Male Model C-u includes only correlated heterogeneity, it ensures that 100% of the variance in alcohol-related risk which remains after fitting the fixed effect of deprivation is ascribed to spatial effects, or correlated heterogeneity. This means that the model assumes that male alcohol-related risk in each data zone depends on the risk estimates in its neighbouring areas.

The ten highest and lowest male alcohol-related relative risk estimates calculated using Male Model C-u are given in Table 6.8. This table also gives the data zone which these risk estimates correspond to, along with the appropriate deprivation score and intermediate geography name. It should be remembered that, for all of the parameters which have not been fully monitored, the 95% credible intervals given are only approximations produced by OpenBUGS. The ten highest male risk estimates correspond to data zones with an area deprivation score of 1, as was the case for the highest 10 male SIR values. It is also noticeable that 4 of the 10 highest risk estimates fall within Glasgow City. This agrees both with previous research and the male SIR results which suggest that there are high levels of male alcohol-related health problems in the Glasgow area. However, the results in Table 6.8 also suggest that there may be clusters of data zones in the Highland region that experience relatively high numbers of male alcohol-related deaths and hospitalisations.

Another feature highlighted in Table 6.8 is that the ten lowest relative risk estimates are non-zero, which was not the case for the 10 lowest male SIR values. The ten lowest model-based male risk estimates all correspond to data zones with an area deprivation score of 8 or greater, with five having the 'best' deprivation score of 10.

The male relative risks estimated using Bayesian spatial models range from 0.1496 to 5.078 whereas the male SIR values range from 0 to 7.952. This smaller range of risk estimates shows that the use of spatial Bayesian modelling techniques has successfully reduced the problem of extremely low and extremely high risk estimates experienced with the SIR methods. These

extreme male SIR values are caused by the rartiy of the disease and the extremely small study regions.

6.7 Male Alcohol-Related Relative Risk Maps

The main aim of this research is to map the alcohol-related health risk across Scotland. In this section the estimates of male alcohol-related relative risk, calculated using Male Model C-u, are plotted for the whole of Scotland at the data zone level of geography.

A data zone map of Scotland depicting the male relative risk estimates is shown in Figure 6.12, with magnified areas of this map shown for Aberdeen (Figure 6.13), Ayrshire (Figure 6.14), the Dundee area (Figure 6.15), Edinburgh (6.16)), Glasgow (Figure 6.17), the Inverness area (Figure 6.18) and Stirling (Figure 6.19).

Again, before any comparisons between these maps and those produced in earlier chapters can be made, it must be noted that the shading cut points used are not the same. Bearing this in mind, the modelled male relative risks appear to give a very similar overall pattern to the male SIR maps shown in Chapter 4. As was the case for the combined risk estimates, the male model-based risk estimates appear much smoother across the country. All of the male relative risk maps show evidence of large blocks or groups of data zones which fall within the same risk category. The removal of the 'noise' experienced in the male SIR maps by using modelling techniques makes it much easier to assess the general pattern of risk in the country and to locate clusters of high-risk areas. For example, the male relative risk estimate map of Glasgow (Figure 6.17) shows, even more strongly than the equivalent male SIR map, that there is a cluster of many data zones towards the East of Glasgow which exhibit very high alcohol-related health risks.

As discussed in section 5.7, using a model which contains only correlated heterogeneity may be regarded by some as forcing any disease maps to be

fairly smooth. The appropriateness of Male Model C-u, as is the case for any model, depends on how the results are to be used.

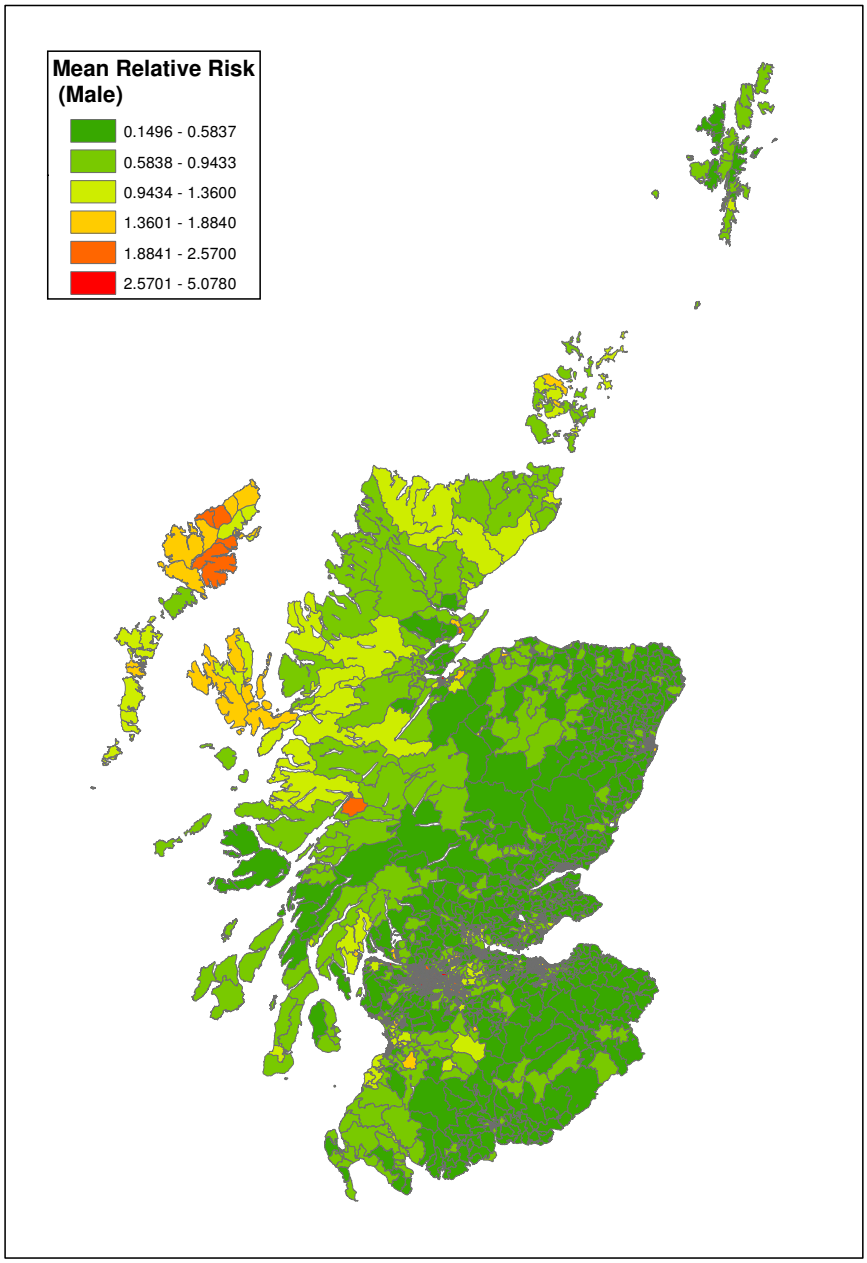


Figure 6.12: Data Zone Map of Mean Male Alcohol-Related Relative Risk

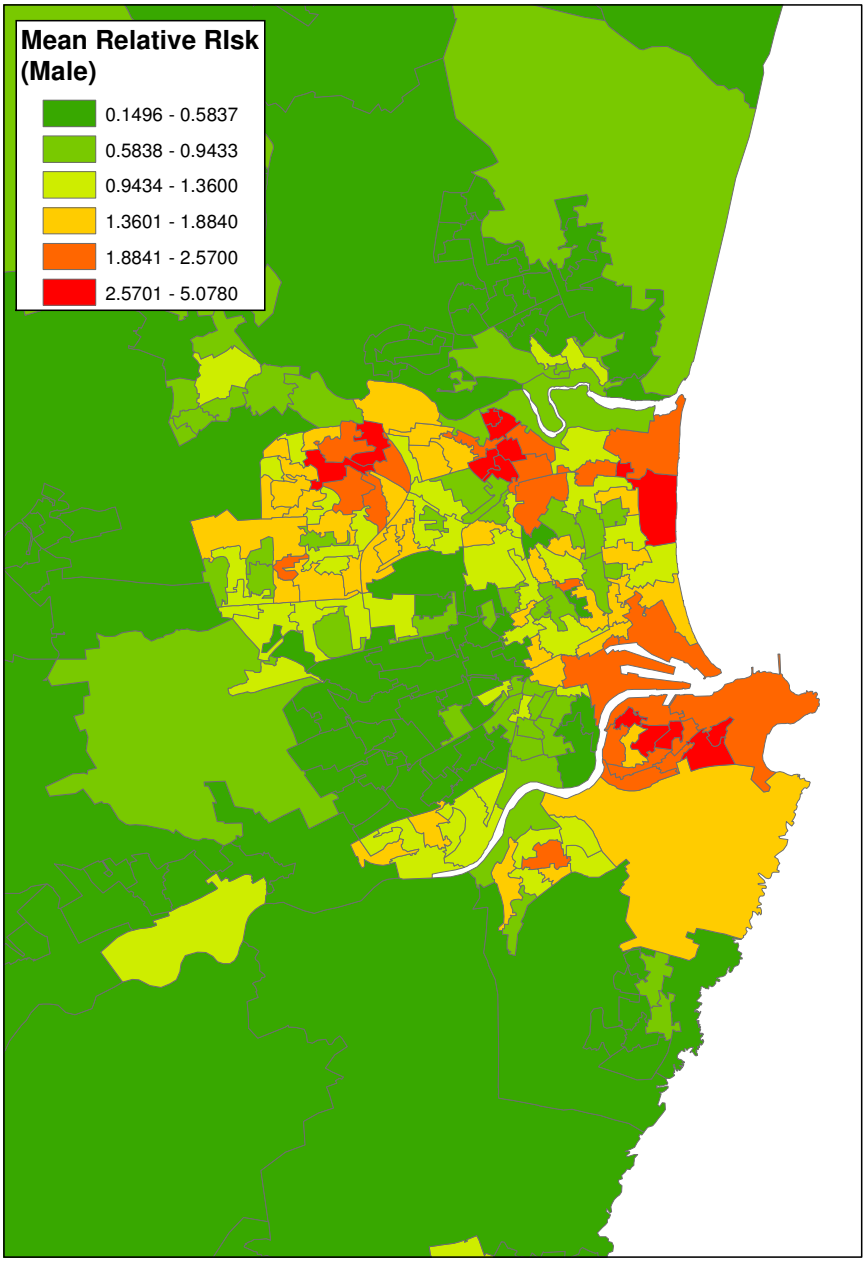


Figure 6.13: Aberdeen Area Data Zone Map of Mean Male Alcohol-Related Relative Risk

	No Covariates				Non-linear Deprivation				Linear Deprivation			
Random Effects	$\overline{\mathbf{D}}$	$\hat{\mathbf{D}}$	**pD**	**DIC**	$\overline{\mathbf{D}}$	$\hat{\mathbf{D}}$	**pD**	**DIC**	$\overline{\mathbf{D}}$	$\hat{\mathbf{D}}$	**pD**	**DIC**
$\mathbf{v_i}$	29430	24870	4554	33980	29460	27050	2404	31860	29440	2.7E+4	2442	31880
$\mathbf{u_i}$	29680	29580	96.43	29780	29750	29880	-127.9	29620	29750	29840	-89.14	29660
$\mathbf{u_i + v_i}$	29530	29960	-430.1	29100	29430	2.9E+4	435.9	29870	29430	28900	529.2	29960

Table 6.3: Male Deviance and DIC using the pD Method

	No Covariates				Non-linear Deprivation				Linear Deprivation			
Random Effects	$\overline{D}$	$\hat{D}$	**p*D**	**DIC**	$\overline{D}$	$\hat{D}$	**p*D**	**DIC**	$\overline{D}$	$\hat{D}$	**p*D**	**DIC**
$\mathbf{v_i}$	29430	12836.9	6418.4	35848.4	29460	11470.4	5735.2	35195.2	29440	11534.8	5767.4	35207.4
$\mathbf{u_i}$	29680	11685.6	5842.8	35522.8	29750	7204.6	3602.3	33352.3	29750	7243.7	3621.9	33371.9
$\mathbf{u_i + v_i}$	29530	12521.6	6260.8	35790.8	29430	10180.8	5090.4	34520.4	29430	10424.4	5212.2	34642.2

Table 6.4: Male Deviance and DIC using the p*D Method

	No Covariates				Non-linear Deprivation				Linear Deprivation			
Random Effects	$\overline{D}$	$\hat{D}$	**pD**	**DIC**	$\overline{D}$	$\hat{D}$	**pD**	**DIC**	$\overline{D}$	$\hat{D}$	**pD**	**DIC**
$\mathbf{v_i}$	29430	24870	4555	33980	29430	2.7E+4	2429	31860	29420	26950	2467	31880
$\mathbf{u_i}$	29680	29640	42.28	29720	29720	29740	-24.81	29690	29720	29650	66.7	29780
$\mathbf{u_i + v_i}$	29440	29640	-204.3	29240	29190	28530	657.8	29850	29190	28500	690.1	29880

Table 6.5: Male Deviance and DIC using the pD Method (Sensitivity Models)

	No Covariates				Non-linear Deprivation				Linear Deprivation			
Random Effects	$\overline{\mathbf{D}}$	$\hat{\mathbf{D}}$	**p*D**	**DIC**	$\overline{\mathbf{D}}$	$\hat{\mathbf{D}}$	**p*D**	**DIC**	$\overline{\mathbf{D}}$	$\hat{\mathbf{D}}$	**p*D**	**DIC**
$\mathbf{v_i}$	29430	12836.9	6418.4	35848.4	29430	11214.8	5607.4	35037.4	29420	11278.4	5639.2	35059.2
$\mathbf{u_i}$	29680	11685.6	5842.8	35522.8	29720	7074.5	3537.2	33257.2	29720	7228.4	3614.2	33334.2
$\mathbf{u_i + v_i}$	29440	11685.6	5842.8	35282.8	29190	7823.4	3911.7	33101.7	29190	7903.2	3951.6	33141.6

Table 6.6: Male Deviance and DIC using the p*D Method (Sensitivity Models)

	Male Model C-u		Male Model C-Sens	
Parameter	**Estimate**	**95% Credible Interval**	**Estimate**	**95% Credible Interval**
α	0.7362	(0.7062, 0.7661)	0.7403	(0.7083, 0.7723)
β_1	0	NA	0	NA
β_2	-0.2949	(-0.3326, -0.2571)	-0.3027	(-0.3442, -0.2615)
β_3	-0.5143	(-0.5553, -0.4735)	-0.5257	(-0.5701, -0.4814)
β_4	-0.7065	(-0.7503, -0.6628)	-0.716	(-0.7628, -0.6693)
β_5	-0.8984	(-0.9451, -0.8518)	-0.912	(-0.9618, -0.8622)
β_6	-1.048	(-1.097, -1.0)	-1.063	(-1.114, -1.012)
β_7	-1.239	(-1.29, -1.188)	-1.257	(-1.311, -1.204)
β_8	-1.471	(-1.526, -1.417)	-1.487	(-1.544, -1.431)
β_9	-1.646	(-1.704, -1.588)	-1.662	(-1.722, -1.602)
β_{10}	-1.846	(-1.909, -1.782)	-1.861	(-1.926, -1.795)
τ_u^2	8.785	(7.581, 10.15)	12.99	(10.99, 15.39)
τ_v^2	NA	NA	32.76	(27.48, 39.1)
var(u)	0.1145	(0.09852, 0.1319)	0.07755	(0.06499, 0.091)
var(v)	NA	NA	0.03078	(0.02558, 0.03638)
θ_1	0.2602	(0.1652, 0.3892)	0.2522	(0.1499, 0.3952)
θ_2	0.2681	(0.198, 0.3544)	0.272	(0.1757, 0.401)
θ_{14}	1.901	(1.376, 2.547)	1.923	(1.308, 2.696)
θ_{89}	0.4708	(0.3651, 0.5964)	0.4392	(0.2935, 0.6284)
θ_{115}	1.803	(1.303, 2.416)	1.814	(1.224, 2.564)
θ_{172}	0.4131	(0.2993, 0.5537)	0.3912	(0.2524, 0.5774)
θ_{985}	0.732	(0.7059, 0.7589)	0.7389	(0.5295, 0.9996)
θ_{2521}	0.6363	(0.4734, 0.837)	0.6385	(0.4214, 0.9235)
θ_{2692}	1.067	(0.8461, 1.324)	1.054	(0.7255, 1.473)
θ_{2885}	0.5145	(0.3363, 0.7547)	0.55	(0.3324, 0.8558)
θ_{2889}	4.832	(3.359, 6.65)	5.043	(3.478, 6.969)
θ_{3044}	1.44	(1.084, 1.867)	1.299	(0.892, 1.807)
θ_{3046}	5.078	(4.139, 6.146)	6.247	(4.992, 7.674)
θ_{3557}	0.3627	(0.2975, 0.4379)	0.3587	(0.2511, 0.4935)
θ_{4687}	1.218	(0.9648, 1.519)	1.372	(0.9506, 1.906)
θ_{5792}	0.3114	(0.2255, 0.4161)	0.2919	(0.1991, 0.4077)
θ_{6238}	0.4795	(0.4588, 0.5004)	0.4758	(0.3345, 0.6546)

Table 6.7: Selection of Parameters from Male Model C-u and Male Model C-Sens

Rank	Data Zone	Intermediate Geography	Local Authority	Deprivation	Mean RR	95% Credible Interval
Lowest 10 RRs	S01005291	Houston South	Renfrewshire	10	0.1496	(0.05005, 0.327)
	S01005296	Houston South	Renfrewshire	10	0.161	(0.06668, 0.3057)
	S01004888	Westfield	North Lanarkshire	10	0.1648	(0.01894, 0.4769)
	S01000972	Gretna & Eastriggs	Dumfries & Galloway	9	0.1713	(0.1026, 0.2655)
	S01005301	Houston South	Renfrewshire	10	0.1726	(0.07153, 0.3279)
	S01005307	Houston North	Renfrewshire	10	0.1728	(0.07354, 0.3203)
	S01000820	Mull, Iona, Coll & Tiree	Argyll & Bute	8	0.1879	(0.06444, 0.393)
	S01000981	Gretna & Eastriggs	Dumfries & Galloway	9	0.1941	(0.1344, 0.2693)
	S01005300	Houston South	Renfrewshire	9	0.1968	(0.08073, 0.3744)
	S01004884	Westfield	North Lanarkshire	9	0.1987	(0.02332, 0.5694)
Highest 10 RRs	S01005594	Ayr North Harbour, Wallacetown & Newton South	South Ayrshire	1	3.709	(2.829, 4.737)
	S01003296	Parkhead West & Barrowfield	Glasgow City	1	3.727	(2.837, 4.781)
	S01006061	Shawfield & Clincarthill	South Lanarkshire	1	3.743	(2.825, 4.832)
	S01003855	Inverness Merkinch	Highland	1	3.94	(2.974, 5.074)
	S01003217	Dalmarnock	Glasgow City	1	4.032	(3.011, 5.255)
	S01003862	Inverness Merkinch	Highland	1	4.221	(2.937, 5.789)
	S01003860	Inverness Merkinch	Highland	1	4.248	(3.075, 5.656)
	S01003849	Inverness Merkinch	Highland	1	4.422	(3.308, 5.723)
	S01003043	Glenwood North	Glasgow City	1	4.832	(3.359, 6.649)
	S01003313	Parkhead West & Barrowfield	Glasgow City	1	5.078	(4.139, 6.145)

Table 6.8: Table of Fitted Male Alcohol-related Relative Risks

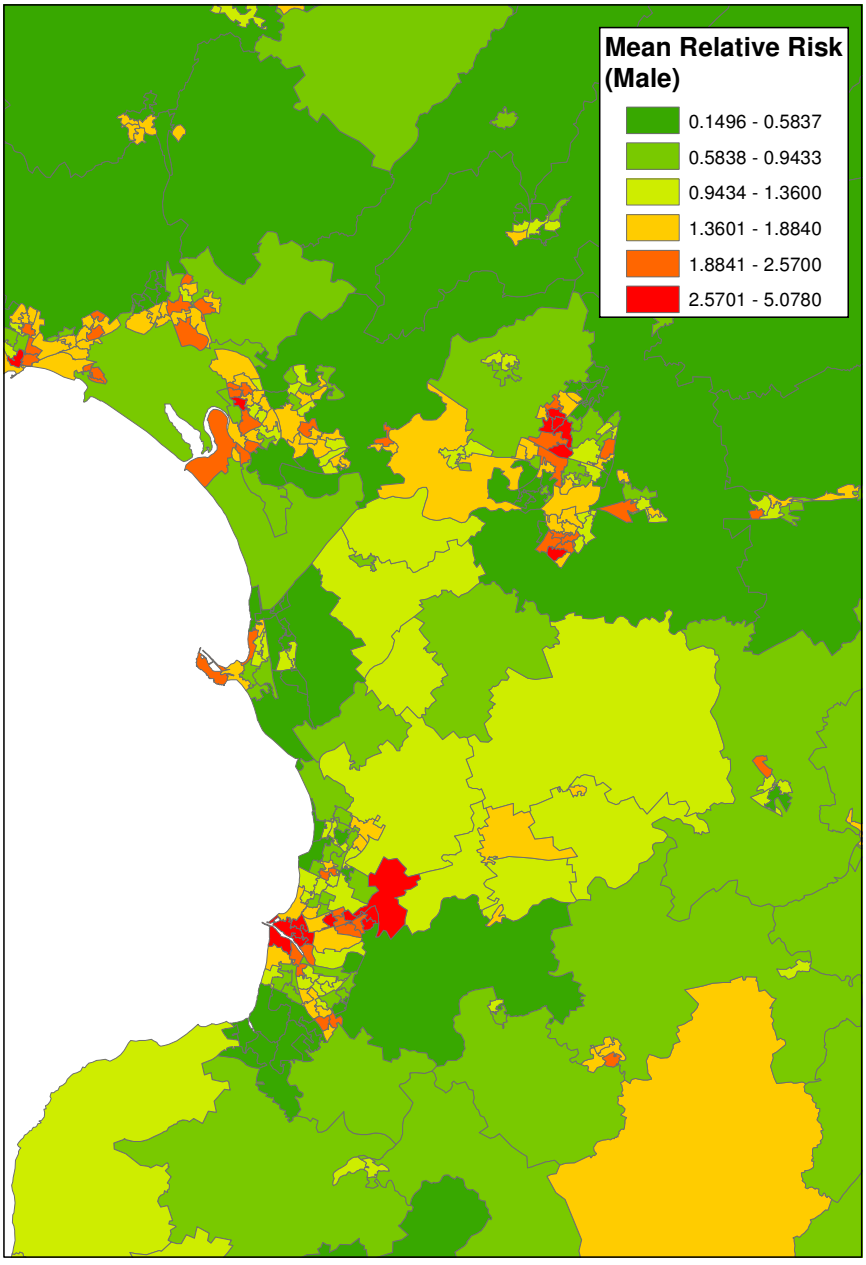


Figure 6.14: Ayrshire Area Data Zone Map of Mean Male Alcohol-Related Relative Risk

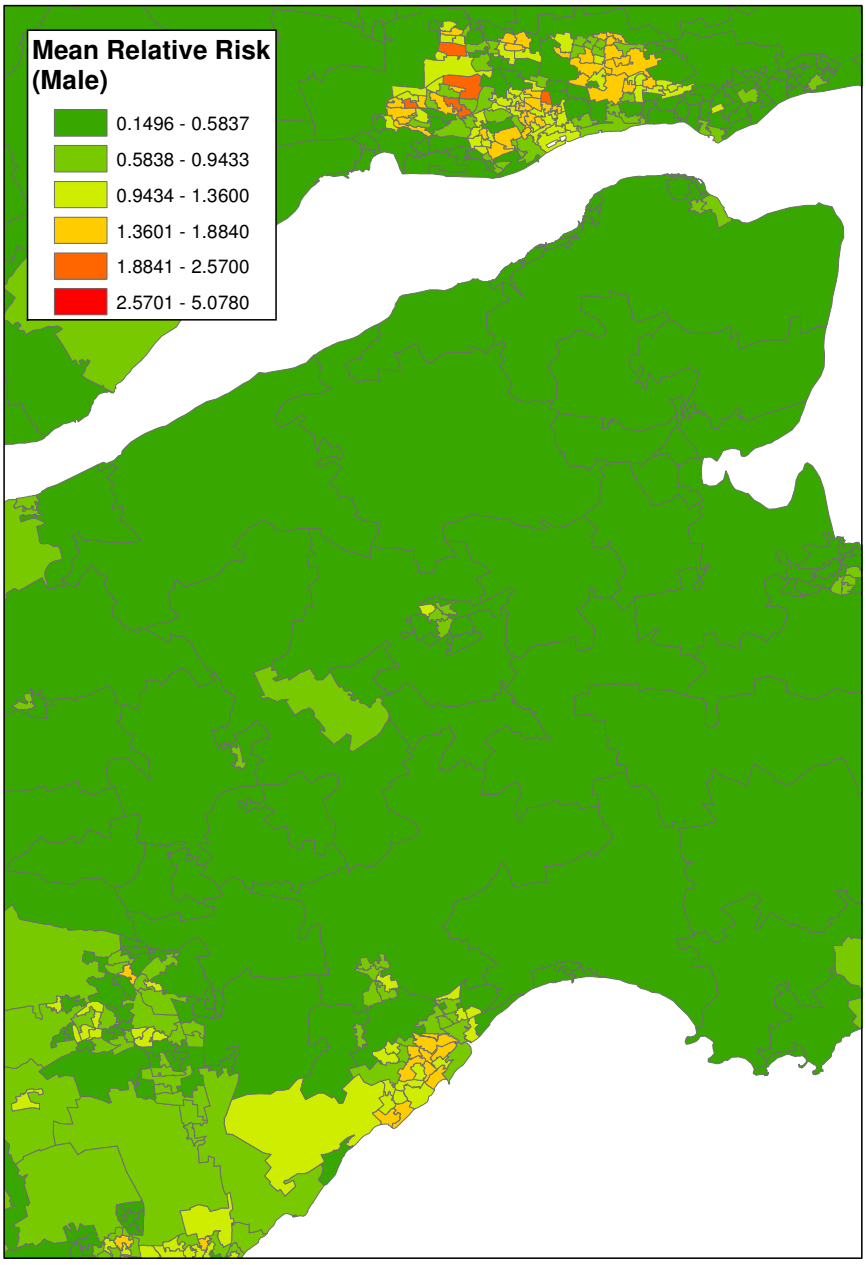


Figure 6.15: Dundee Area Data Zone Map of Mean Male Alcohol-Related Relative Risk

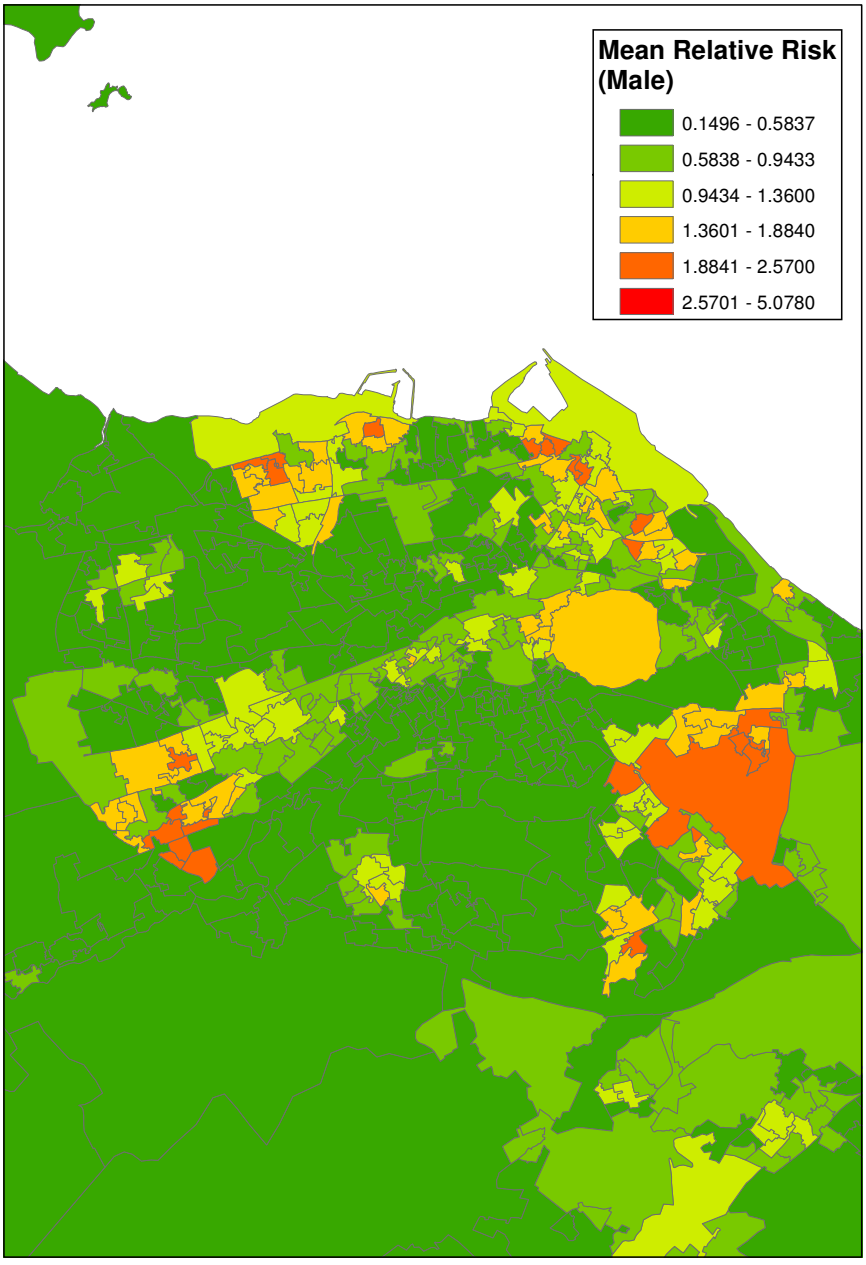


Figure 6.16: Edinburgh Area Data Zone Map of Mean Male Alcohol-Related Relative Risk

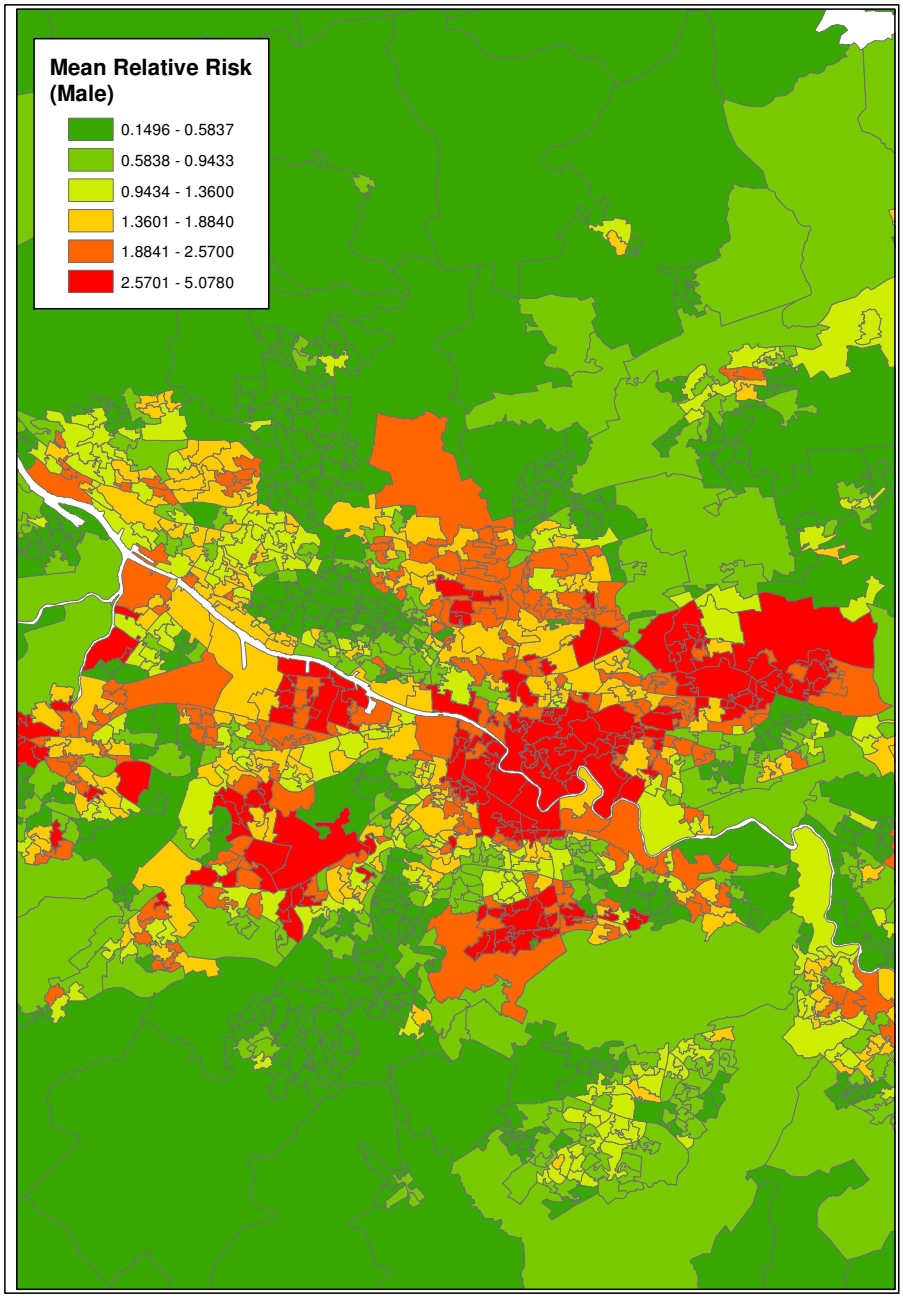


Figure 6.17: Glasgow Area Data Zone Map of Mean Male Alcohol-Related Relative Risk

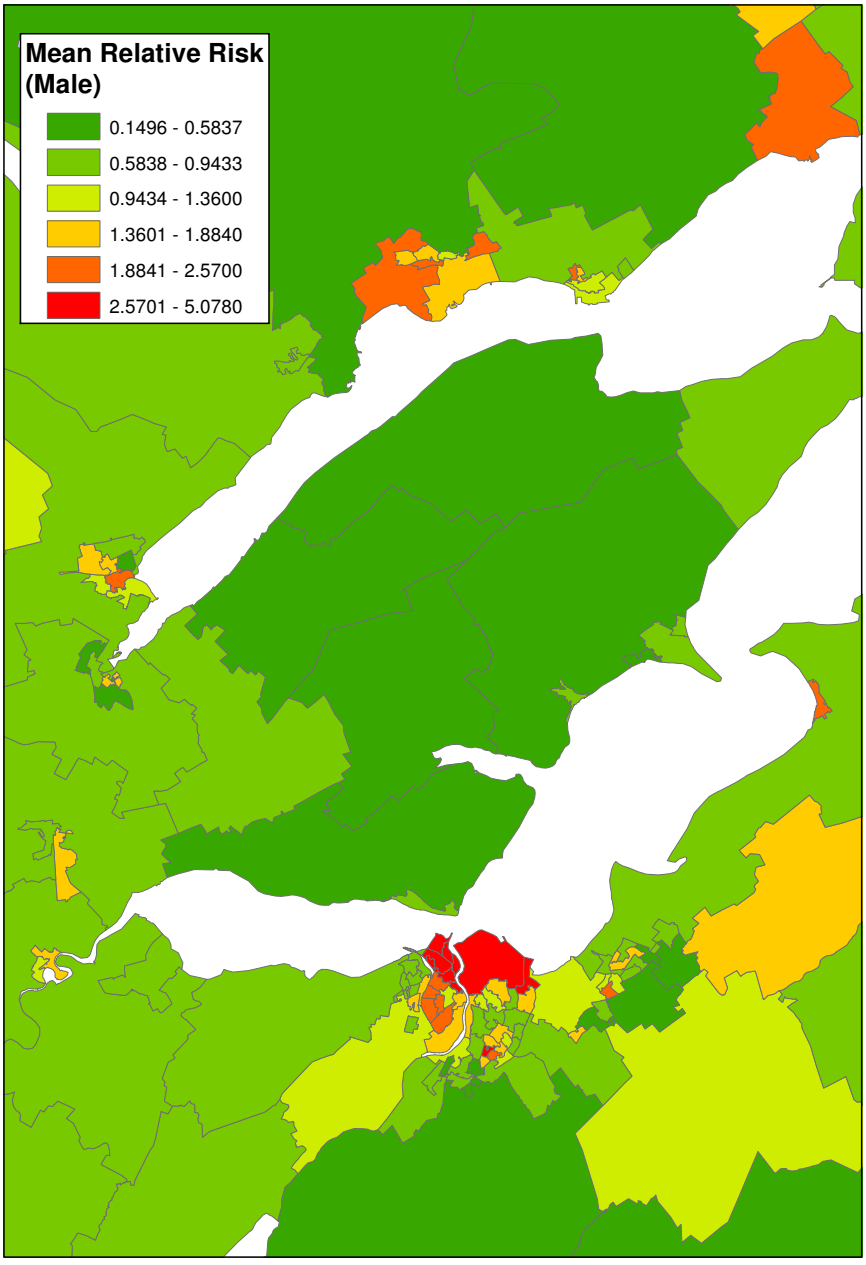


Figure 6.18: Inverness Area Data Zone Map of Mean Male Alcohol-Related Relative Risk

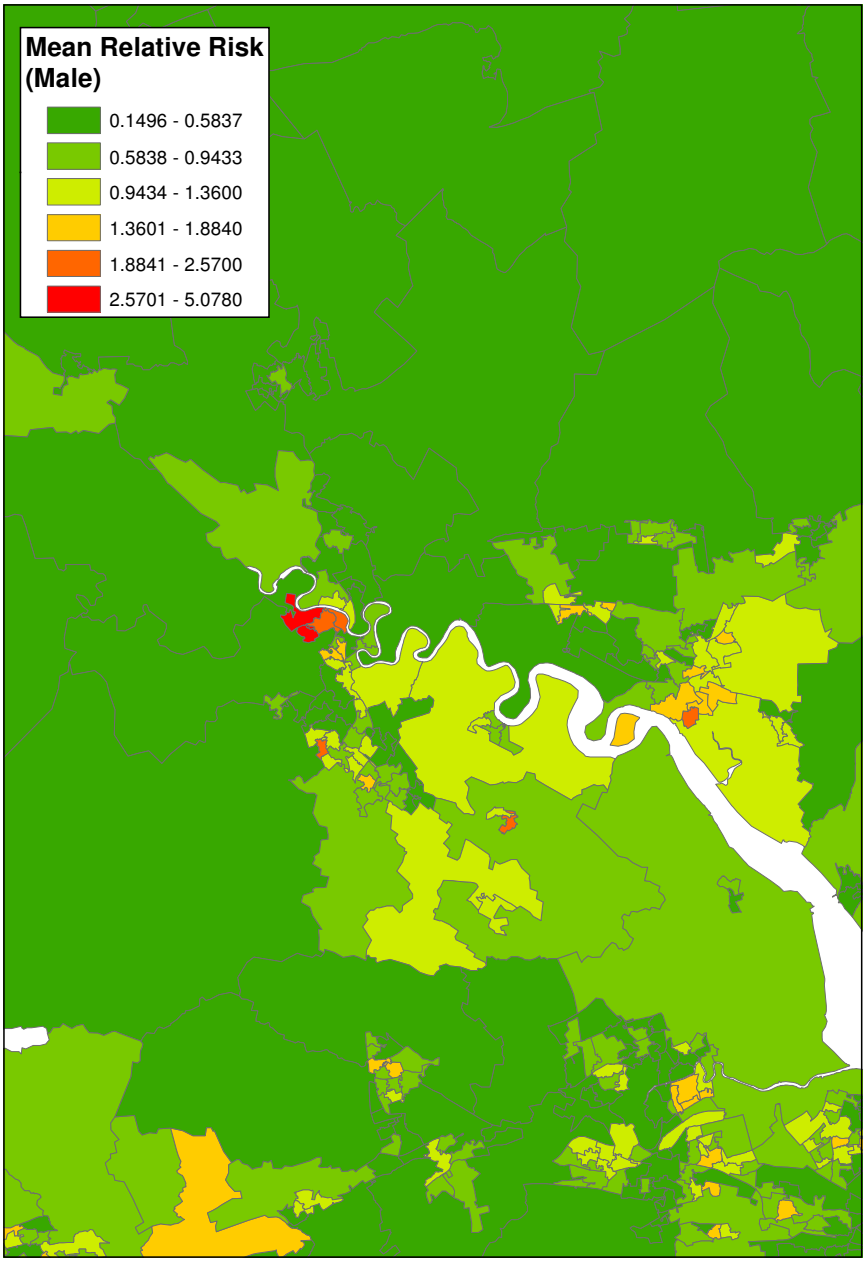


Figure 6.19: Stirling Area Data Zone Map of Mean Male Alcohol-Related Relative Risk

Chapter 7

BYM Models for Female Data

The previous two chapters have used Bayesian spatial models to obtain alcohol-related health-risk estimates both for the whole population and for the male population at the data zone level across Scotland. This chapter aims to use the same model structures to calculate estimates of female alcohol-related health risks in Scotland. Doing so will allow comparisons between risk patterns and any associations with deprivation to be made between the results for each gender.

The data used in this chapter consists of the observed and expected counts of female alcohol-related deaths and hospitalisations in Scotland during the years 2002 to 2006 inclusive. The expected number of female deaths in each data zone during this period has been calculated using indirect age standardisation as discussed in Chapter 3.

7.1 Female Models Considered

This chapter explores nine different models for the female alcohol-related relative risk across Scotland. These models are of exactly the same form as those considered for the combined and male only data in Chapter 5 and Chapter 6 respectively. The models are based on the Besag, York and Mollié model (Besag et al. (1991)) and differ in terms of both fixed and random

effects.

As random effects each model will include either uncorrelated heterogeneity (v), correlated heterogeneity (u) or both ($u + v$). The only fixed effect considered here is a bona fide data zone deprivation score, as discussed in Chapter 2. Area deprivation score has been modelled in two ways; firstly in a linear fashion and secondly by fitting a separate parameter to each of the 10 deprivation scores. Since age standardisation was performed when the expected number of events in each area were calculated it should not be included at the modelling stage.

Table 7.1 gives a summary of the nine models considered in this section, listing the fixed and random effects included in each.

Model Name	**Fixed Effects**	**Random Effects**
Female Model A-v	none	v
Female Model A-u	none	u
Female Model A	none	$u + v$
Female Model B-v	linear deprivation	v
Female Model B-u	linear deprivation	u
Female Model B	linear deprivation	$u + v$
Female Model C-v	non-linear deprivation	v
Female Model C-u	non-linear deprivation	u
Female Model C	non-linear deprivation	$u + v$

Table 7.1: Models for Female Alcohol-Related Relative Risks

Since these models have exactly the same structure as those considered for the combined and male data, the discussion in Section 5.1 regarding model structures, parameters and prior distributions still hold. However, the most important points are restated here.

Vague normal priors with mean zero and precision e^{-5} have been assigned to all deprivation parameters with the exception of β_1 in the non-linear case, which has been arbitrarily set to zero.

Each model specifies, where necessary, a normal prior distribution with

mean zero for the uncorrelated heterogeneity and a conditional autoregressive prior for the correlated heterogeneity. This gives

$$\begin{aligned} v_i &\sim N\left(0, \tau_v^2\right) \text{ and} \\ \left[u_i | u_j, i \neq j, \tau_u^2\right] &\sim N\left(\overline{u_i}, \tau_i^2\right) \end{aligned}$$

where τ_v^2, $\overline{u_i}$ and τ_i^2 are as described in section 3.6 of Chapter 3.

Vague gamma hyperprior distributions have been assigned to the inverse variance hyperparameters of both random effects. In particular

$$\tau_v^2 \sim \text{Gamma}(0.5, 0.0005) \text{ and}$$
$$\tau_u^2 \sim \text{Gamma}(0.5, 0.0005).$$

This hyperprior distribution has been chosen since it is sufficiently vague and commonly used in disease mapping studies where there is no strong prior knowledge.

All female models were run using OpenBUGS and the code for Female Model A, Female Model B and Female Model C is shown in Appendices section 1.1, 1.2 and 1.3 respectively. Note that this is exactly the same code as was used for the equivalent combined and male models, the only difference lies in the data to which they were fitted. Again, the code for all other female models can be derived from this code by deleting any redundant sections; for example delete all code in Female Model B which relates to uncorrelated heterogeneity, v, in order to obtain Female Model B-u.

7.2 Female Convergence

As has been discussed before, the aim of using any of the sampling methods discussed in Chapter 3 is for the joint distribution of the simulated values to converge, or settle, to the joint posterior distribution. A burn-in period of iterations is run for each model until adequate convergence has been reached and then, after discarding existing simulated values, the simulation is continued for a further number of iterations. The length of the burn-in period

and subsequent iterations required varies greatly between different studies and models.

For reasons discussed in section 5.2, it is not practicable to record the simulated value at each iteration for all of the parameters in each model. For all models the female relative risk parameters have been fully monitored for a subset of data zones and a summary monitor has been set for the remaining areas. All other model parameters have been fully monitored. Instead of storing the simulated value of a parameter at every post-burn-in iteration a summary monitor only holds summary statistics about the simulated sample of that parameter. These summary statistics are updated at every iteration but the simulated value itself is then discarded. The 95% credible intervals given by OpenBUGS are exact for fully monitored variables but only approximate for those assigned a summary monitor.

The relative risk parameters chosen to be fully monitored, along with the reasons for doing so are given in Table 8.2

Achieving adequate convergence for the female models proved to be much more difficult than for the combined and male equivalents. In the end, it was decided to use a burn-in period of 150,000 iterations followed by simulating two chains for a further 350,000 iterations. There are still some convergence issues even with these long chain lengths. However, due to the limited time to complete this project longer chains could not realistically be investigated.

Discussed below are various convergence checks and diagnostics which were monitored for the female models. Since the same checks were carried out for all nine models, these will only be discussed in detail for Female Model C-u.

Firstly, the history plots for all fully monitored model parameters will be considered, shown in Figures 7.1, 7.2, 7.3, 7.4, 7.5 and 7.6. History plots give a line plot of simulated parameter values against iteration number, with one simulation chain shown in blue and the other in red. The plots in Figure 7.1 and Figure 7.2 show that the deprivation parameters, β_2 to β_{10}, appear

Data Zone Code	Relative Risk Parameter	Reason Chosen
S01006393	θ_{115}	poor deprivation score
S01006438	θ_{14}	poor deprivation score
S01006490	θ_{1}	good deprivation score
S01006505	θ_{2}	good deprivation score
S01003744	θ_{2521}	rural area
S0100391	θ_{2692}	rural area
S01003380	θ_{3044}	urban/city area
S01002325	θ_{4687}	urban/city area
S01005521	θ_{985}	island / no neighbouring areas
S01000447	θ_{6238}	island / no neighbouring areas
S01003031	θ_{2885}	lowest total population
S01000799	θ_{5792}	highest male population
S01002622	θ_{3557}	highest female population
S01003313	θ_{3046}	highest male SIR
S01006473	θ_{89}	zero female SIR value
S01006341	θ_{172}	zero female SIR value
S01003043	θ_{2889}	very high male SIR

Table 7.2: Data Zones with Fully Monitored Female Relative Risk Estimates

to have converged well; they all form a horizontal band across the history plot where both chains consistently overlap. Figure 7.1 also shows, however, that the α parameter from this model, which was assigned a flat improper prior, has not achieved ideal convergence. The two α chains show similar values, but can be seen to 'weave' above and below each other and hence not consistently overlap.

All of the relative risk parameters monitored appear to exhibit strong convergence (Figures 7.3 to 7.6) with the exception of θ_{985} and θ_{6238}. The similarity of these two history plots with the history plot for α makes it clear that this is due to the lack of convergence in α feeding through to these parameters. Both θ_{985} and θ_{6238} are relative risk parameters for islands, or 'neighbourless', data zones. Female Model C-u includes only correlated heterogeneity, u, which has been assigned a continuous autoregressive prior

distribution. As discussed in Section 3.6 the CAR prior depends on the mean of the bordering, or neighbouring, areas $\overline{u_i}$ and τ_i^2. For areas with no neighbours, then, any relative risks calculated using this model can effectively have no random effects. This is likely to explain the similarities between the patterning observed in the α, θ_{985} and θ_{2386} history plots.

The Gelman-Rubin diagnostic plots as described in section 3.4.3 will also be discussed for the fully monitored parameters of Female Model C-u, shown in Figure 7.7 and Figure 7.8. Again all of the parameter plots shown in these figures suggest that the simulated values have converged to the equilibrium distribution, with the exception of α, θ_{985} and θ_{6238}. Comfortable convergence is exhibited in the majority of the BGR plots since the green line, which shows the width of the central 80% interval of the pooled chains, and the blue line, which shows the average width of the 80% intervals within the idividual chains, are both stable and the red line which represents their ratio is stable at a value of 1. In fact, the intervals are so similar for the individual chains and the pooled chains that the blue line almost completely obscures the green line. The three parameters mentioned which could have converged better are highlighted since the red and blue lines in these plots are not horizontal in appearance and have not settled to any values. For the BGR plots too, the patterns observed for θ_{985} and θ_{6238}, which refer to neighbourless areas, are very similar to that shown for α for the reasons discussed above. This relates to a common problem in spatial statistics known as 'edge effects' where by values for areas which lie at the edge of a map or study region are often less well estimated than those that do not. Although it is a common problem this is the first time that there has been evidence to suggest such problems in this study.

A last visual check of convergence for Female Model C-u will be carried out by looking at the probability density plots of the simulated parameter samples, given in Figure 7.9 and Figure 7.10. Obviously, these plots can only be checked for the parameters which have been fully monitored. In Figure

7.9 we can see that the density plot for α is smooth in appearance, but is not symmetrical. This is a further indication that the convergence of α is not as good as one would have hoped for. The probability density plots for θ_{985} and θ_{6238} show much weaker signs of poor convergence than their history and BGR plots. However, they are less bell shaped and less symmetrical than one would ideally like to see.

I feel it is reasonable to use these female models given that the deviance on which model selection is based has converged very well, and it appears that all other parameters have converged adequately apart from the intercept level of risk and the relative risk estimates for island areas.

Again, as is true for any real-life Bayesian model simulation, it is not possible to say for sure that the parameter estimates have converged to the required posterior distribution. It is possible, but unlikely, that instead of exploring the whole parameter space the simulation chains have become 'stuck' in a certain area.

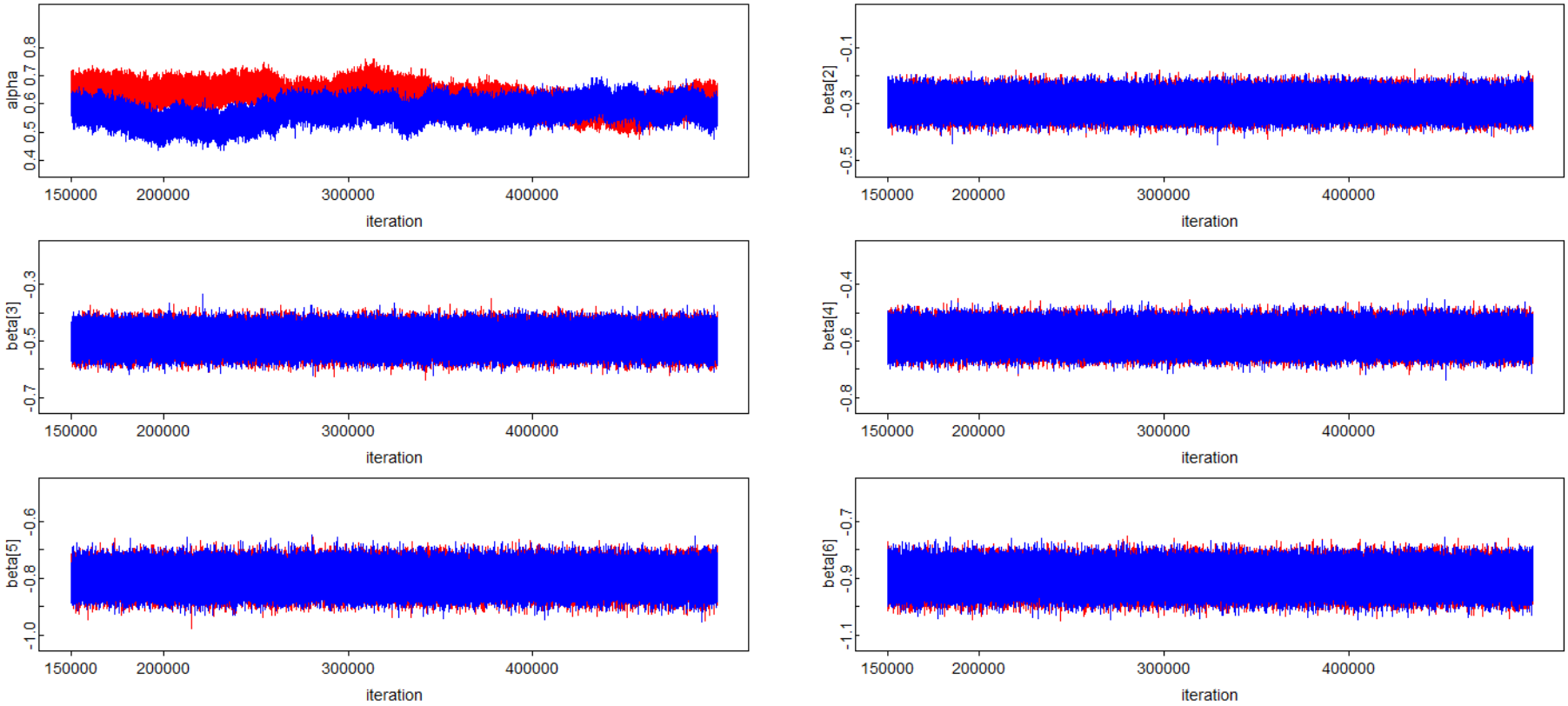


Figure 7.1: Simulation History Plots for a Subset of Female Model C-u Parameters (part 1)

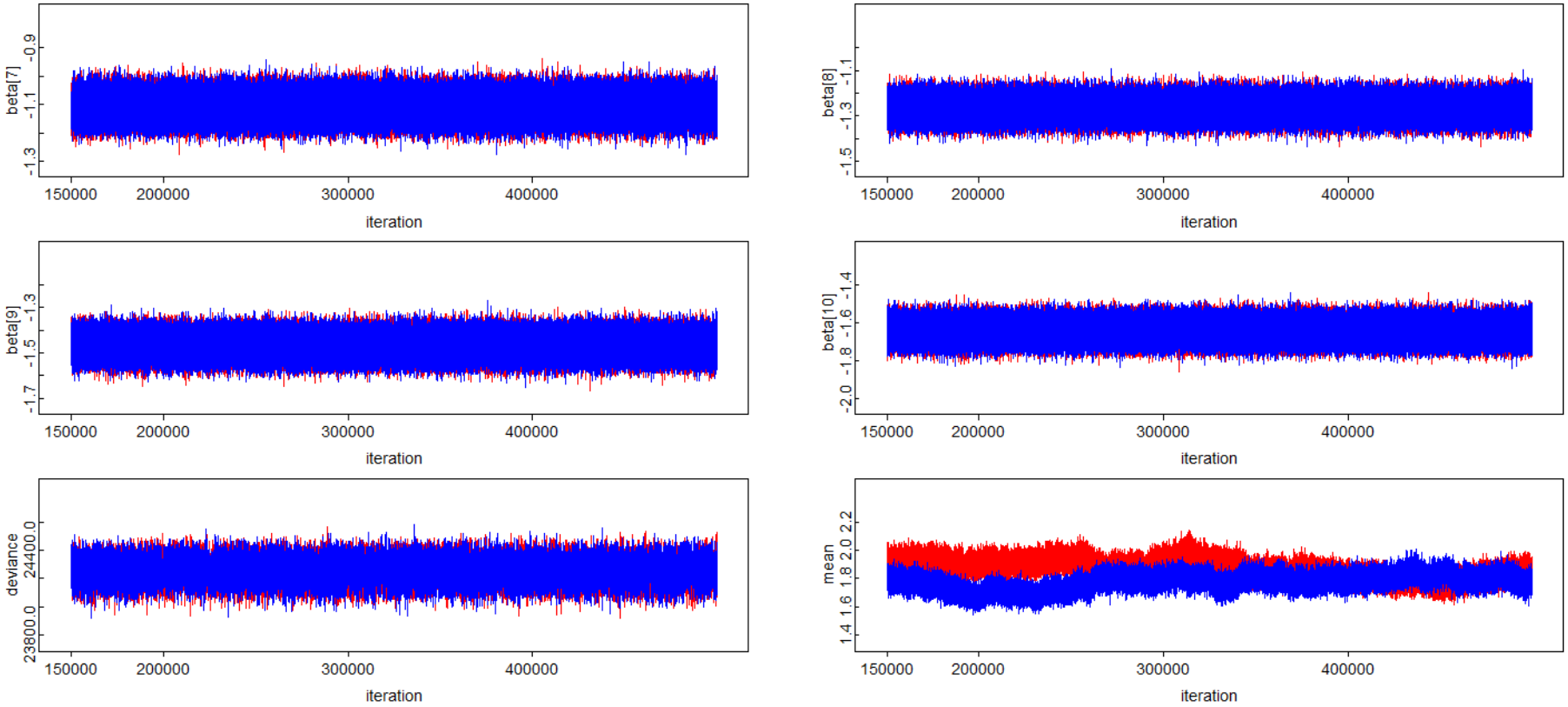


Figure 7.2: Simulation History Plots for a Subset of Female Model C-u Parameters (part 2)

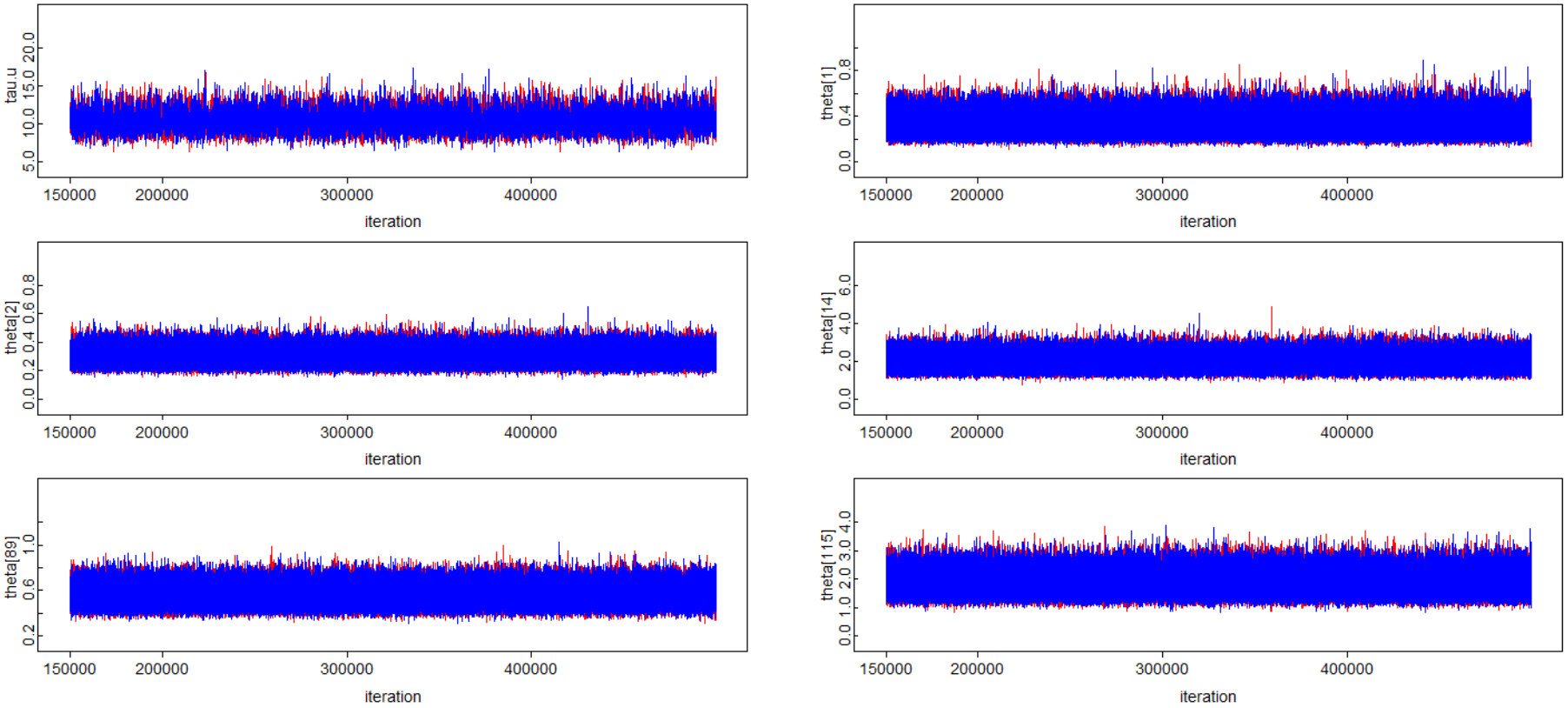


Figure 7.3: Simulation History Plots for a Subset of Female Model C-u Parameters (part 3)

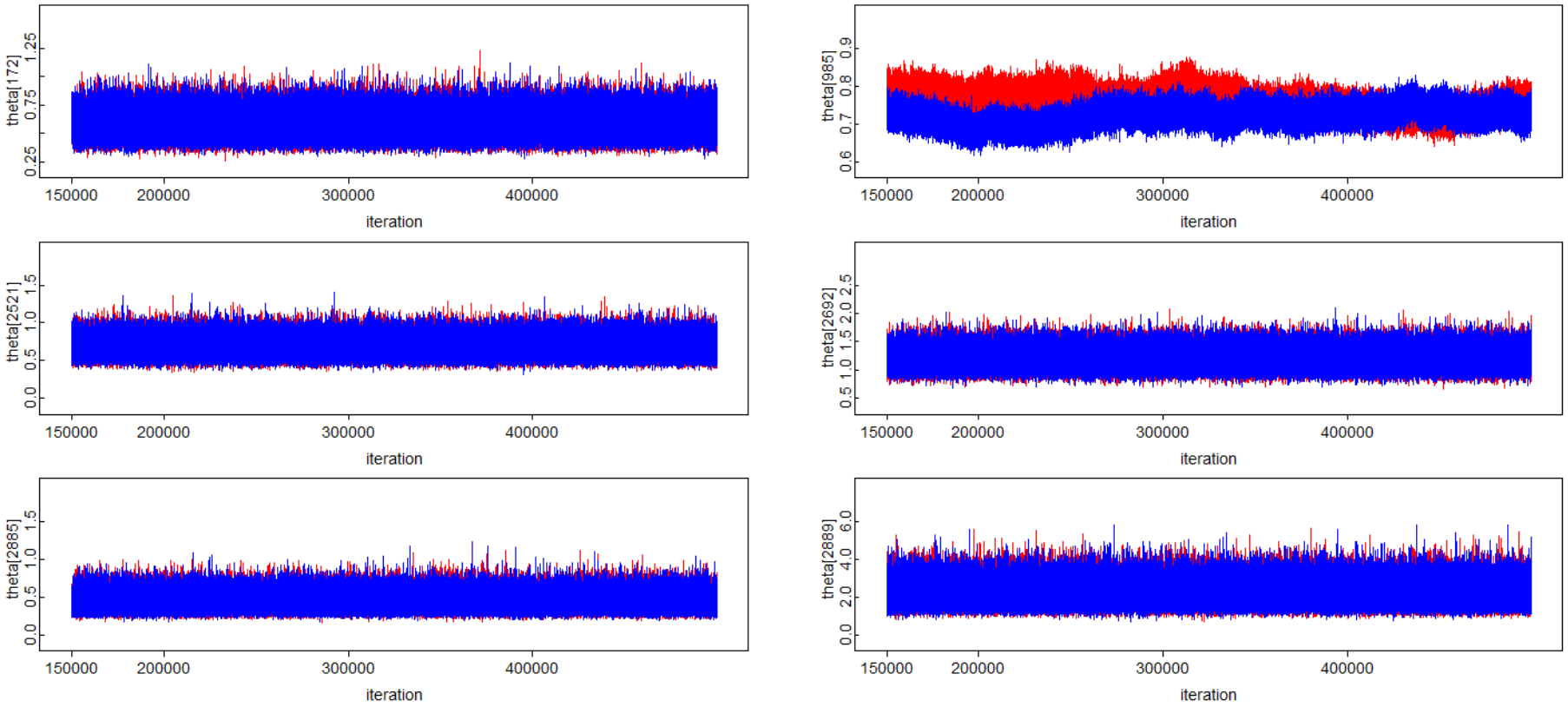


Figure 7.4: Simulation History Plots for a Subset of Female Model C-u Parameters (part 4)

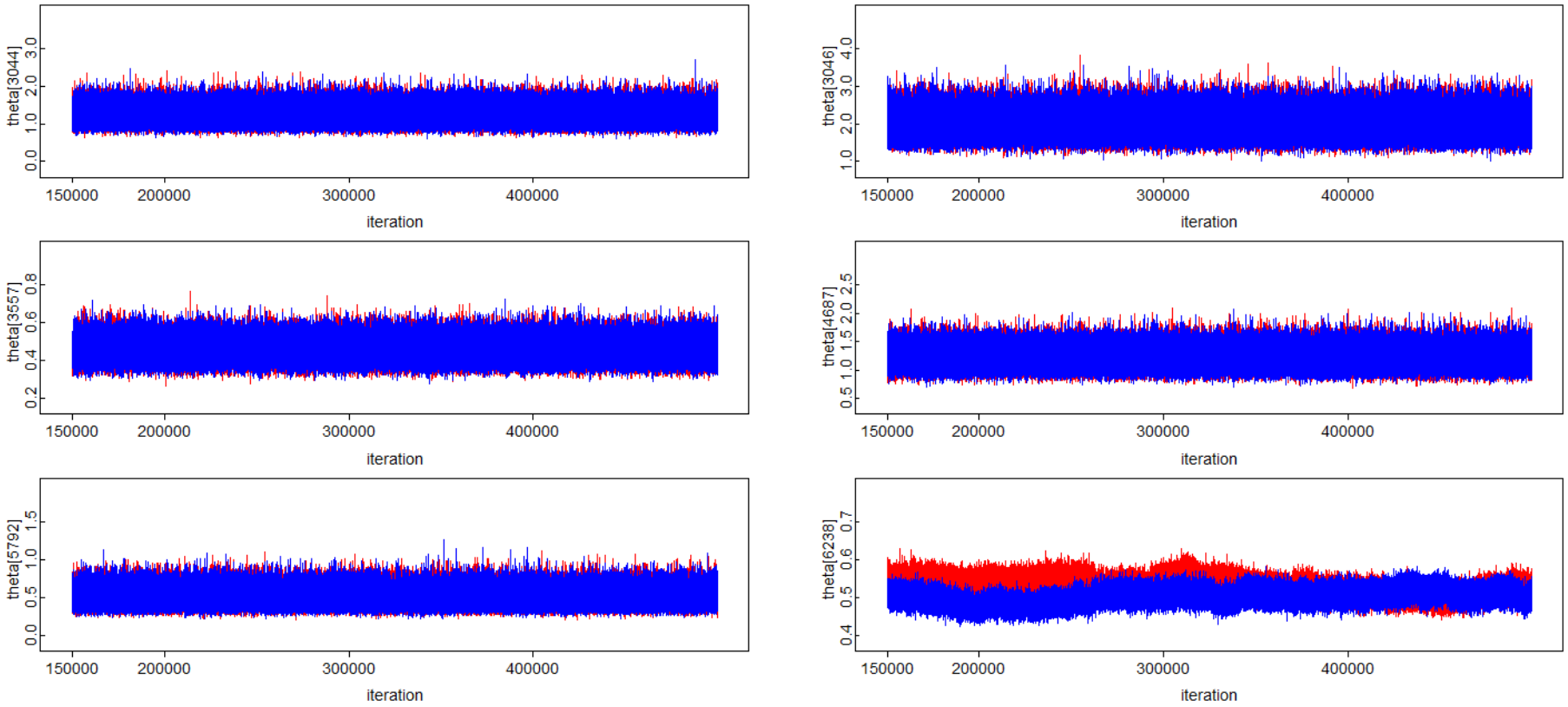


Figure 7.5: Simulation History Plots for a Subset of Female Model C-u Parameters (part 5)

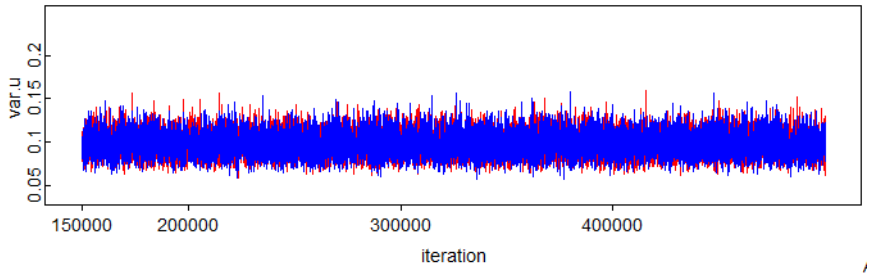


Figure 7.6: Simulation History Plots for a Subset of Female Model C-u Parameters (part 6)

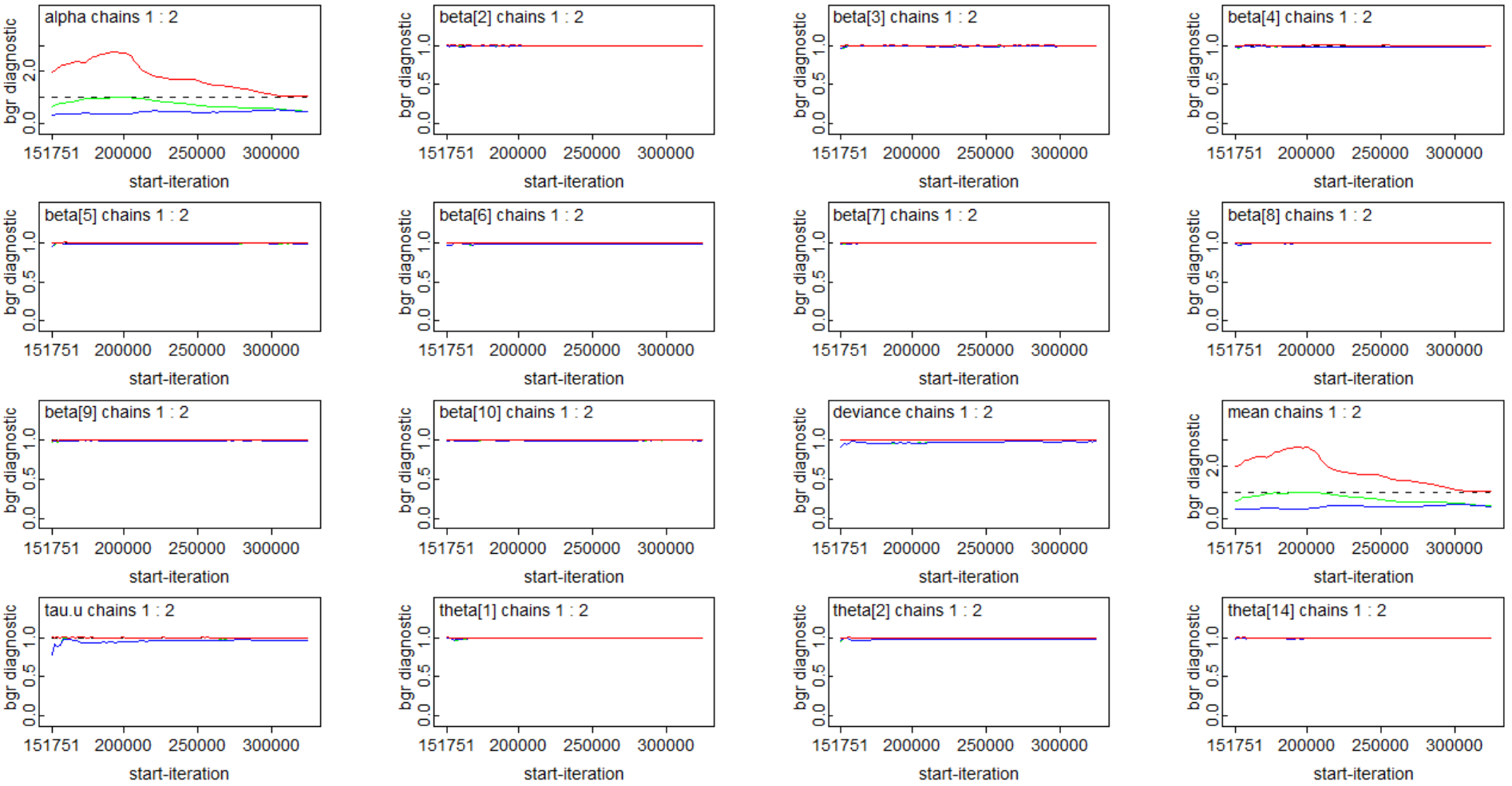


Figure 7.7: BGR Diagnostic Plots for a Subset of Female Model C-u Parameters (part 1)

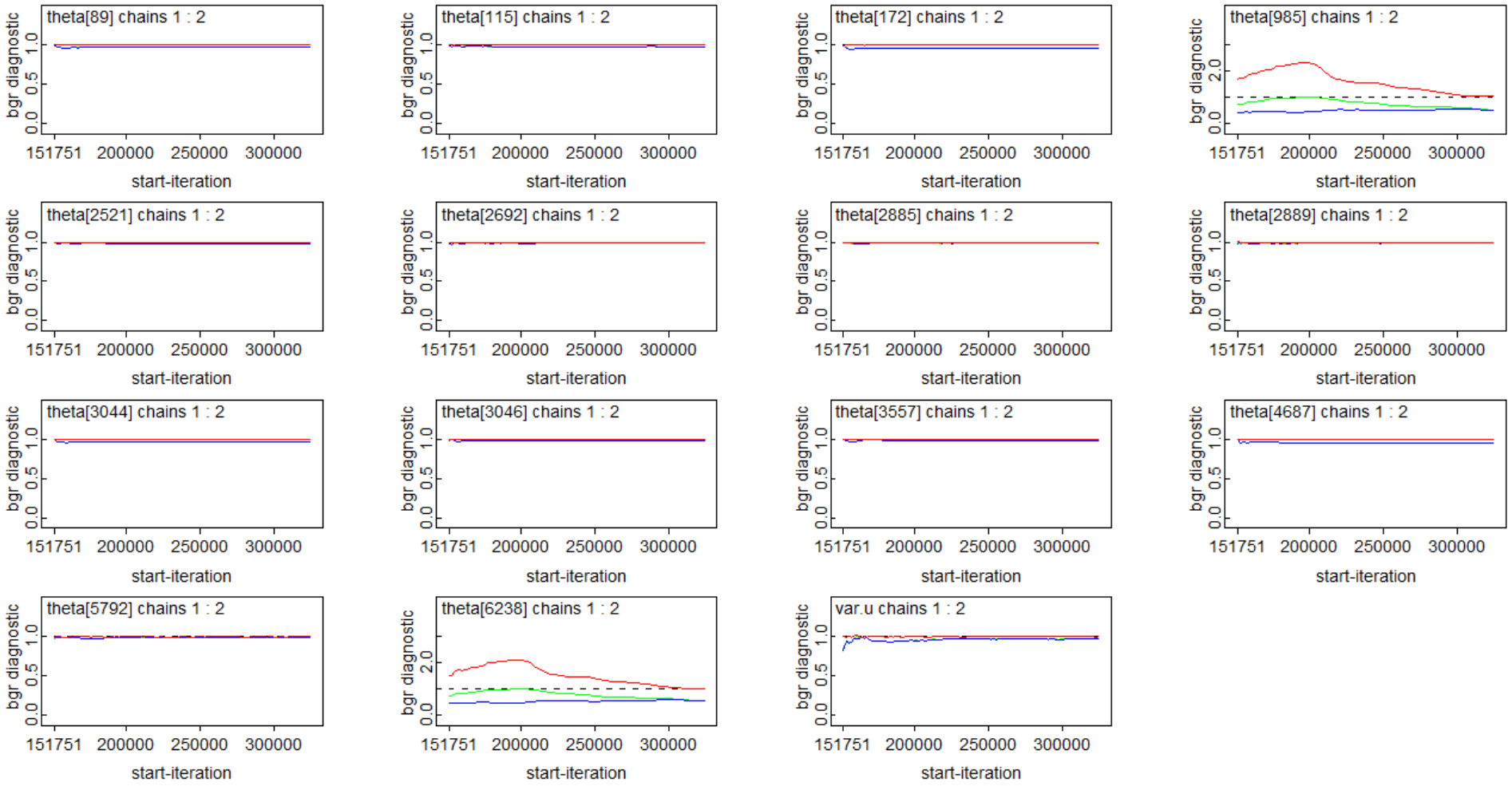


Figure 7.8: BGR Diagnostic Plots for a Subset of Female Model C-u Parameters (part 2)

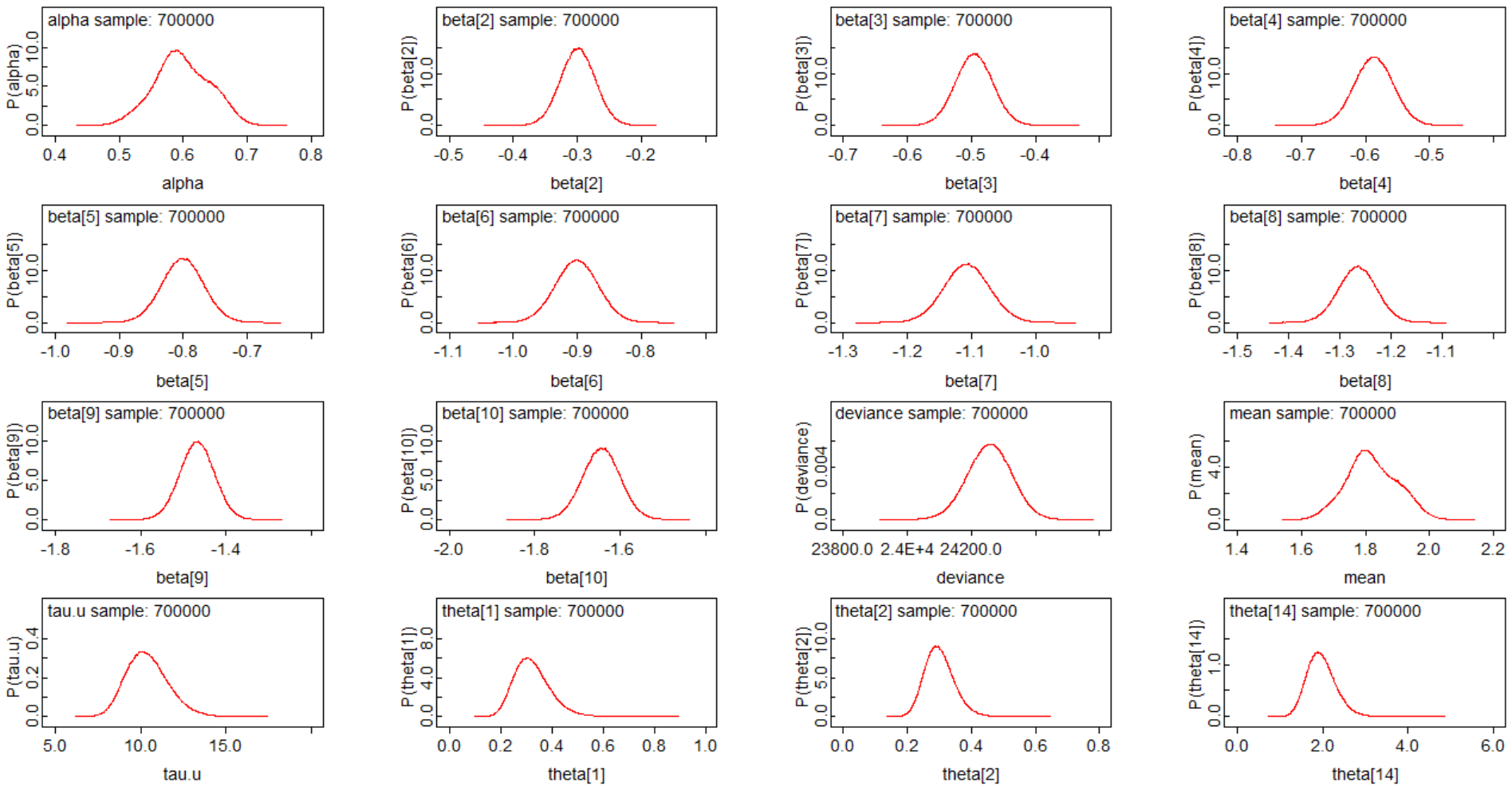


Figure 7.9: Posterior Density Plots for a Subset of Female Model C-u Parameters (part 1)

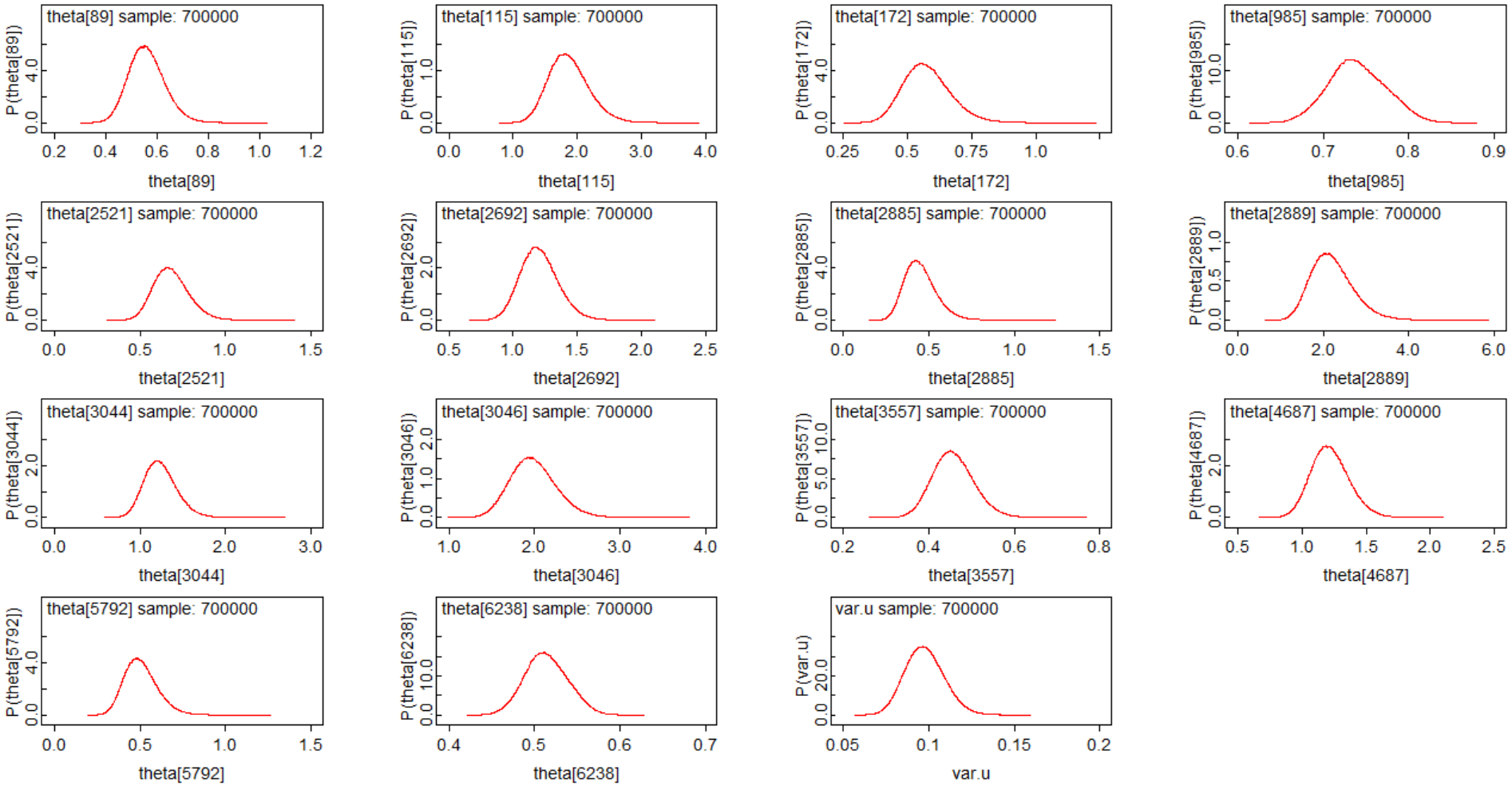


Figure 7.10: Posterior Density Plots for a Subset of Female Model C-u Parameters (part 2)

7.3 Female DIC

As discussed in previous chapters, the DIC is a measure of goodness-of-fit which is commonly used to choose between Bayesian models of differing complexities. The DIC values have been calculated for all nine female models using the pD method and the results are given in Table 7.3. The values in this table have been split according to the random and fixed effects in each model; for fixed effects either none, linear deprivation or non-linear deprivation and for random effects either correlated heterogeneity only (u), uncorrelated heterogeneity only (v) or both ($u + v$).

Table 7.3 shows that negative pD values are experienced for some of the female models as in the case of the combined and male relative risk models. Again, these possible but undesirable negative values only occur for models which contain spatial random effects (u). The lowest female DIC value calculated using the pD method is shown to correspond to a model which contains no fixed effects. Given the strong similarities in patterning exhibited between the female SIR maps and the deprivation score maps it is very unlikely that deprivation score does not explain a significant amount of the variation in relative risks. Due to the dubious model choice and the undesirable negative pD values, it was decided to instead calculate DIC using the p*D method used in the previous two chapters and discussed in Chapter 3.

The female DIC values calculated using the p*D method are shown in Table 7.4. The lowest female DIC value using the p*D method corresponds to the model which includes non-linear deprivation and only correlated heterogeneity (u), Female Model C-u. The p*D method of calculating DIC therefore leads to a much more intuitive model choice.

7.4 Female Model Selection

It would be normal practice to select Female Model C-u as the 'best' model for female alcohol-related relative risks in Scotland since it gave the lowest DIC value. The selection of this model is consistent with the subjective impressions made in the female SIR section of Chapter 4. Here it was predicted that there would be a link between female alcohol-related relative risk and area deprivation score, but that the relationship may not be linear. The lack of linearity was suggested since there was noticed to be a larger increase in average female SIR value between the deprivation scores of 1 and 2 than between any other pair of consecutive scores.

As always though, it must be remembered that these Bayesian models are based on a set of assumptions chosen by the modeller and incorporated via the prior and hyperprior distributions used. Although Female Model C-u has been chosen when using the current priors, this might not always be the case. It is necessary to investigate how sensitive the female model selection and parameter estimates are to the choice of prior distributions used. Unfortunately, due to the limited time available to complete this project a fully comprehensive sensitivity analysis will not be possible. The sensitivity to the choice of hyperpriors assigned to τ_u^2 and τ_v^2 will, however, be examined.

7.5 Female Hyperprior Sensitivity Analysis

The nine female models given in Table 7.1 will be re-run with different hyperprior distributions given to τ_u^2 and τ_v^2. Similarly to previous chapters, these sensitivity models will use the same names as those given in Table 7.1 but with 'Sens' appended at the end, for example 'Female Model A-Sens'. Running these models with different hyperpriors allows one to see whether the choice of distribution will affect the model selection and to see by how much the estimated alcohol-related relative risk estimates are affected.

The alternative hyperpriors used are the same as those used for the com-

bined and male models. Where necessary, each female sensitivity model uses the following hyperpriors:

$$\tau_u^2 \sim \text{Gamma}(1,1) \text{ and}$$
$$\tau_v^2 \sim \text{Gamma}(1,1).$$

These distributions are a lot more restrictive than the original gamma distributions used and would not be an ideal first choice. However, it is hoped that, even though the hyperpriors used are so different, the model choice and relative risk estimates will be similar between the original and sensitivity female models. If this is true then this indicates that the model results are not too sensitive to hyperprior choice.

Convergence of all nine sensitivity models was monitored and all were found to have converged at least as well as Female Model C-u discussed above after a burn-in period of 150,000 iterations followed by two chains of 350,000 simulations.

Once again, when all female sensitivity models had been run, the DIC values were calculated. Table 7.5 shows the DIC values for the female sensitivity models calculating using the pD method. This table shows that negative pD values are also experienced by some of the female sensitivity models which contain correlated heterogeneity. The lowest DIC value is experienced for the model which contains linear deprivation and both correlated and uncorrelated random effects, Female Model B-Sens. However, due to the negative pD values and to be consistent with the model selection methods used for the original models, DIC calculated using the p*D method will be used instead. Table 7.6 gives the DIC values calculated using the p*D method for the female sensitivity models. Here the lowest DIC value corresponds to the model with non-linear deprivation and only u, Female Model C-u-Sens.

The original and sensitivity female models therefore both lead to the same model structure being chosen. It is now of interest to see how the estimated parameter values compare between the two models. Table 7.7 gives

the estimate and 95% credible interval for all fully monitored parameters in both Female Model C-u and Female Model C-u-Sens. From this table it is clear that the female models are not too sensitive to hyperprior choice, since the parameter estimate from each model lies within the corresponding credible interval from the other.

7.6 Female Model Results

Since the results of the selected female model and female sensitivity model were so similar, it has been decided that Female Model C-u will be considered the final model since it uses more appropriate and vague hyperpriors.

The Female Model C-u results in Table 7.7 show that none of the credible intervals for the deprivation parameters, β_2 to β_{10}, contain zero and there are only two instances where they slightly overlap: for β_3 and β_4 and for β_5 and β_6. Boxplots of the deprivation parameter samples have been produced and are given below in Figure 7.11. The boxplots show that the β values appear to be less linear than those shown for the combined and male models discussed previously, but they still appear fairly linear. The reason that non-linear as opposed to linear deprivation proved to be included in the best fitting female model is likely to be the large decrease in estimated value between β_1 and β_2. These results support the model choice; if all credible intervals were found to contain zero or overlap, this would suggest that non-linear deprivation was not necessary and an alternative model may be more appropriate.

All of the deprivation score parameter estimates are less than or equal to zero and the value of β_{d_i} gets progressively smaller as d_i increases from a score of 1 to 10. This is as one would expect, since it suggests that more deprived ares will have a smaller decrease in relative risk than less deprived areas. In fact, Female Model C-u suggests that it is highly likely that, on average, the least deprived data zones with a deprivation score of 10 have an alcohol-related relative risk which is between only 0.177 and 0.211 times

that of the most deprived data zones.

If the fully monitored relative risk parameters are compared between the final female results (Table 7.7) and the male results (Table 6.7) it can be seen that several areas appear to show significant differences between male and female estimates. So, although both the male and female risks are estimated using the same model structure, the data has resulted in different estimates being produced.

The structure of Female Model C-u ensures that 100% of the variance which remains after fitting non-linear deprivation is assigned to spatial effects, or correlated heterogeneity. This means that the chosen model assumes that the female alcohol-related relative risk in each data zone depends on the risk estimates of its neighbouring areas. Thus, estimates for island or coastal areas, or any areas with a small number of bordering areas for that matter, may have poorer less reliable estimates than those with many neighbours. This phenomenon is known as the 'edge effects' and has been extensively studied in papers such as Rodeiro & Lawson (2005) and Yamada (2009). The effective lack of random effects in the chosen model for neighbourless areas means that all neighbourless areas with the same deprivation score will have the same relative risk estimates.

The ten highest and ten lowest alcohol-related relative risk estimates calculated using Female Model C-u are given in Table 7.8. First of all this table shows that unlike the female SIR values given in Table 4.3 there are now no zero female risk estimates. It is also clear that the model-based female risk estimates have much a smaller range than the female SIR values. This indicates that the modelling process has successfully dealt with, or smoothed over, the SIR problem of extreme risk estimates.

Another clear issue in this table is that the credible intervals for the four lowest risk estimates contain negative values. Obviously, by the nature of relative risks and the models used, no negative values could have been simulated for these parameters. The reason for the negative value is that

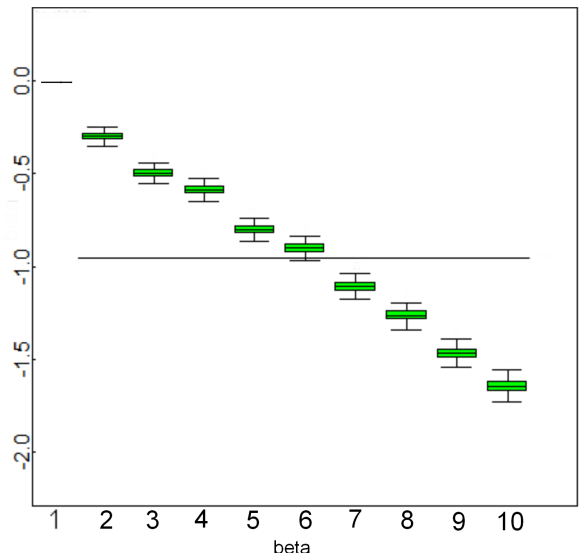


Figure 7.11: Boxplots of Deprivation Parameters for Female Model C-u

when parameters are assigned a summary monitor rather than a full monitor, which is the case for the majority of parameters in this research, the credible intervals produced by OpenBUGS are only approximate. This is the first time this issue has arisen and I believe it to be because the lowest estimated female risks are so close to zero.

Although it has not been possible to fully monitor all of the female relative risk parameters, those listed in Table 8.2 have been. For these parameters the exact 95% credible intervals given by the full monitor have been compared with their approximate equivalents given by the summary monitor. Apart from for a few intervals for estimates which are extremely close to zero, they were almost exactly the same.

If we look further at the results in Table 7.8 it is again the case that the ten highest female risk estimates from Female Model C-u correspond to data zones with a deprivation score of 1, whereas only half of the lowest risk estimates are for areas with the least deprived score of 10.

Although the 20 most extreme model-based risk estimates for females (Table 7.8) do not correspond to the same data zones as those for males (Table 6.8) there are some strong similarities. Data zones in the following

intermediate geographies appear in the highest ten risk estimates for both males and females: Ayr North Harbour, Wallacetown and Newton South and Inverness Merkinch. In a comparison of the ten lowest male and female model-based risk estimates in these tables, it can be seen that none of the intermediate geographies are the same, but both sexes experienced low estimated values in Renfrewshire and Argyll & Bute.

7.7 Female Alcohol-Related Relative Risk Maps

A much simpler way to examine the patterning of the Female Model C-u alcohol-related relative risk estimates is to map them at the data zone level across Scotland.

Similarly to previous chapters, the female relative risk estimates will be mapped at the data zone level of geography across the whole of Scotland (Figure 7.12). Magnified areas of this map will also be shown, due to the very small area of many data zones, for Aberdeen (Figure 7.13), Ayrshire (Figure 7.14), Dundee (7.15), Edinburgh (7.16), Glasgow (7.17), the Inverness area (Figure 7.18) and Stirling (Figure 7.19).

When comparing the full map of female relative risk estimates (Figure 7.12) with the female SIR map (Figure 4.24) it is clear that the modelling process has helped to create a much smoother map of female risk estimates. However, it must be remembered that the risk cut points used in each are not the same. The modelled results show large blocks of colour which represent clusters of areas which fall into the same risk category. The appearance of higher female risk values in the North West Isles of Scotland appear to be even stronger for the modelled risk estimates.

The model-based female risk maps also share very similar patterns to the deprivation maps shown in Chapter 2, as one would expect since deprivation score has been included in the modelling process. This added smoothness makes it much easier to pick out 'hotspots' of relatively high or low female

alcohol-related risk; this is shown especially well if the female relative risk and female SIR maps of Edinburgh and Glasgow are compared (Figure 7.16, Figure 4.28, Figure 7.17 and Figure 4.29).

The patterns exhibited by the female model-based risk estimates are very similar to those shown in the male maps in Chapter 6; however, they are some differences. Most notably the high alcohol-related health risks experienced in the South and East of Glasgow City appear to be much more extreme for males than for females. Also, females have been estimated to have relatively higher risks in North and South Uist compared to males.

Again, there is the possibility that some would say Female Model C-u forced the female relative risk maps to be overly-smooth by only including spatial random effects in the model. This issue is discussed further in Section 5.7.

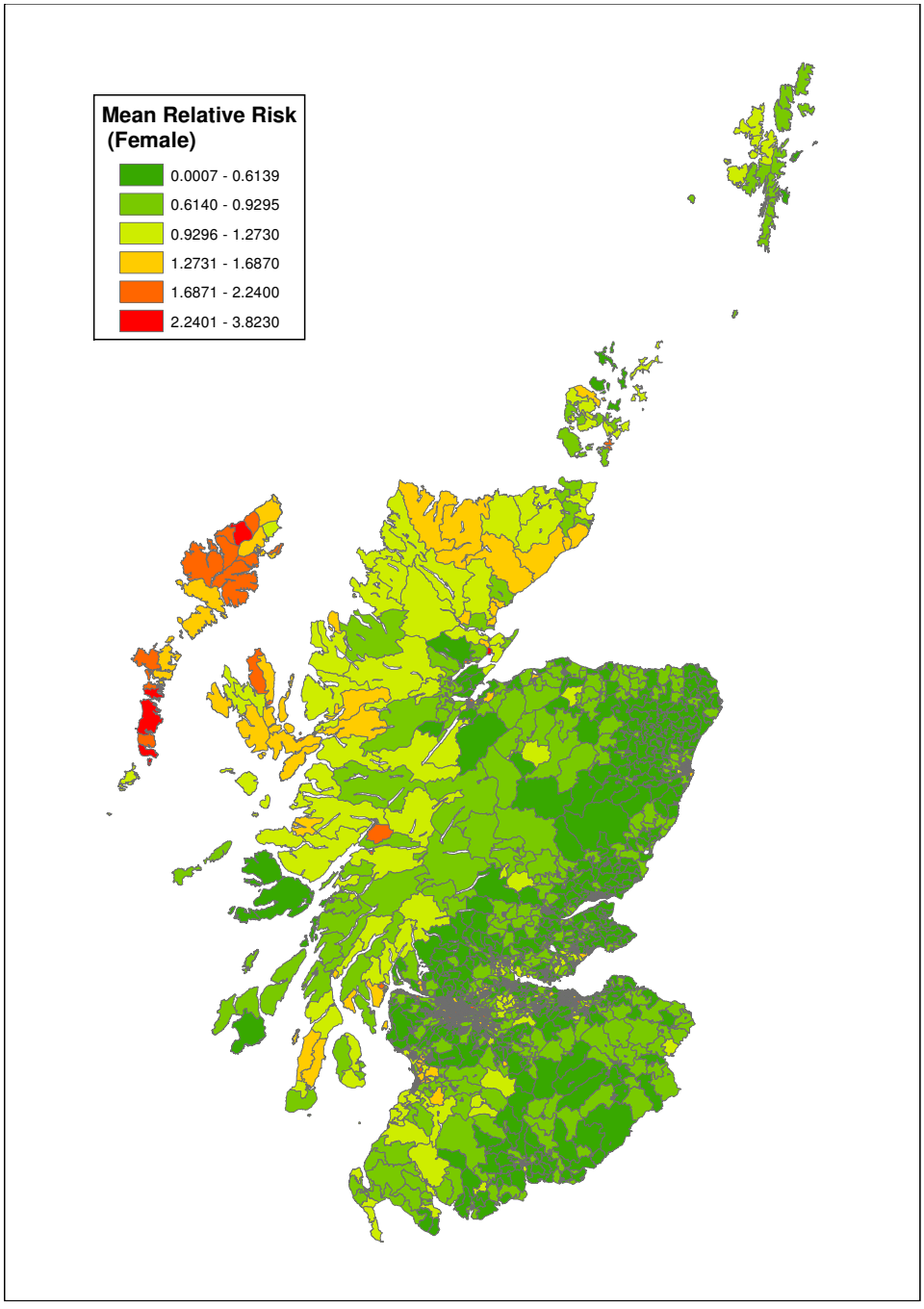


Figure 7.12: Data Zone Map of Mean Female Alcohol-Related Relative Risk

	No Covariates				Non-linear Deprivation				Linear Deprivation			
Random Effects	$\overline{D}$	$\hat{D}$	**pD**	**DIC**	$\overline{D}$	$\hat{D}$	**pD**	**DIC**	$\overline{D}$	$\hat{D}$	**pD**	**DIC**
$\mathbf{v_i}$	24320	21440	2877	27190	24190	22960	1230	25420	24190	22930	1255	25440
$\mathbf{u_i}$	24300	24290	12.48	24310	24260	24330	-68.69	24190	24260	24220	41.6	24300
$\mathbf{u_i + v_i}$	24280	25670	-1388	22900	24200	24290	-89.33	24110	24210	24260	-47.53	24170

Table 7.3: Deviance Statistics and DIC using pD Method for Female Models

	No Covariates				Non-linear Deprivation				Linear Deprivation			
Random Effects	$\overline{\mathbf{D}}$	$\widehat{\mathbf{var}}\{\mathbf{D}\}$	**p*D**	**DIC**	$\overline{\mathbf{D}}$	$\widehat{\mathbf{var}}\{\mathbf{D}\}$	**p*D**	**DIC**	$\overline{\mathbf{D}}$	$\widehat{\mathbf{var}}\{\mathbf{D}\}$	**p*D**	**DIC**
$\mathbf{v_i}$	24320	12387.7	6193.85	30513.9	24190	10962.1	5481.0	29671.0	24190	11004.0	5502.0	29692.0
$\mathbf{u_i}$	24300	9594.2	4797.1	29097.1	24260	4727.9	2364.0	26624.0	24260	4820.5	2410.3	26670.3
$\mathbf{u_i + v_i}$	24280	10363.2	5181.6	29461.6	24200	7353.1	3676.5	27876.5	24210	7334.2	3667.1	27877.1

Table 7.4: Deviance Statistics and DIC using p*D Method for Female Models

	No Covariates				Non-linear Deprivation				Linear Deprivation			
Random Effects	$\overline{D}$	$\hat{D}$	**pD**	**DIC**	$\overline{D}$	$\hat{D}$	**pD**	**DIC**	$\overline{D}$	$\hat{D}$	**pD**	**DIC**
$\mathbf{v_i}$	24310	21420	2883	27190	24120	22820	1300	25420	24120	22790	1322	25440
$\mathbf{u_i}$	24300	24250	40.48	24340	24210	24260	-53.37	24150	24210	24220	-11.71	24200
$\mathbf{u_i + v_i}$	24080	23970	107.4	24190	24200	24290	-89.33	24110	23800	23550	247.2	24050

Table 7.5: Female Deviance Statistics and DIC using pD Method (Sensitivity Models)

	No Covariates				Non-linear Deprivation				Linear Deprivation			
Random Effects	$\overline{\mathbf{D}}$	$\widehat{\mathbf{var}}\{\mathbf{D}\}$	**p*D**	**DIC**	$\overline{\mathbf{D}}$	$\widehat{\mathbf{var}}\{\mathbf{D}\}$	**p*D**	**DIC**	$\overline{\mathbf{D}}$	$\widehat{\mathbf{var}}\{\mathbf{D}\}$	**p*D**	**DIC**
$\mathbf{v_i}$	24310	12365.4	6182.7	30492.7	24120	9982.0	4991.0	29111.0	24120	9980.0	4990.0	29110.0
$\mathbf{u_i}$	24300	9615.8	4807.9	29107.9	24210	4538.7	2269.4	26479.4	24210	4703.2	2351.6	26561.6
$\mathbf{u_i + v_i}$	24080	9862.5	4931.2	29011.2	23800	6107.4	3053.7	26853.7	23800	6190.5	3095.3	26895.3

Table 7.6: Female Deviance Statistics and DIC using p*D Method (Sensitivity Models)

	Female Model C-u		Female Model C-u-Sens	
Parameter	**Estimate**	**95% Credible Interval**	**Estimate**	**95% Credible Interval**
α	0.5986	(0.514, 0.6809)	0.6047	(0.4524, 0.6785)
β_1	0	NA	0	NA
β_2	-0.2991	(-0.352, -0.2461)	-0.2981	(-0.3516, -0.2446)
β_3	-0.4957	(-0.5526, -0.4388)	-0.4947	(-0.5522, -0.4373)
β_4	-0.5867	(-0.6461, -0.5274)	-0.5859	(-0.6458, -0.5259)
β_5	-0.8008	(-0.8645, -0.7372)	-0.8004	(-0.8648, -0.7359)
β_6	-0.9009	(-0.9663, -0.8356)	-0.9003	(-0.9663, -0.8344)
β_7	-1.107	(-1.177, -1.037)	-1.106	(-1.176, -1.035)
β_8	-1.265	(-1.339, -1.192)	-1.264	(-1.338, -1.191)
β_9	-1.468	(-1.547, -1.389)	-1.466	(-1.545, -1.387)
β_{10}	-1.643	(-1.729, -1.558)	-1.642	(-1.729, -1.557)
τ_u^2	10.39	(8.252, 13.03)	9.245	(7.509, 11.34)
var(u)	0.09753	(0.07676, 0.1212)	0.1094	(0.08817, 0.1332)
θ_1	0.3239	(0.207, 0.4816)	0.3224	(0.2012, 0.4882)
θ_2	0.3	(0.2203, 0.3981)	0.2974	(0.2151, 0.3997)
θ_{14}	1.965	(1.385, 2.692)	1.967	(1.367, 2.73)
θ_{89}	0.5594	(0.4351, 0.7065)	0.5562	(0.4274, 0.7099)
θ_{115}	1.872	(1.33, 2.552)	1.875	(1.313, 2.583)
θ_{172}	0.5758	(0.4171, 0.7733)	0.5748	(0.4105, 0.7813)
θ_{985}	0.7398	(0.6774, 0.806)	0.7452	(0.6401, 0.805)
θ_{2521}	0.6833	(0.5044, 0.9047)	0.6861	(0.4987, 0.92)
θ_{2692}	1.204	(0.9423, 1.513)	1.203	(0.9309, 1.525)
θ_{2885}	0.4498	(0.2959, 0.6572)	0.4563	(0.2933, 0.6787)
θ_{2889}	2.207	(1.388, 3.305)	2.218	(1.365, 3.377)
θ_{3044}	1.241	(0.9103, 1.651)	1.25	(0.9035, 1.683)
θ_{3046}	1.988	(1.512, 2.558)	1.987	(1.494, 2.581)
θ_{3557}	0.4573	(0.3701, 0.5579)	0.4562	(0.3661, 0.5607)
θ_{4687}	1.22	(0.9542, 1.534)	1.22	(0.9453, 1.549)
θ_{5792}	0.5057	(0.3428, 0.7175)	0.507	(0.3378, 0.7291)
θ_{6238}	0.514	(0.4673, 0.5637)	0.518	(0.4427, 0.5641)

Table 7.7: Selection of Parameter Results from Female Model C-u and Female Model C-u-Sens

Rank	Data Zone	Intermediate Geography	Local Authority	Deprivation	Mean RR	95% Credible Interval
Lowest 10 RRs	S01000751	Lomond Shore	Argyll & Bute	9	0.000671	(-0.001195, 0.002018)
	S01005515	North and East Isles	Shetland Islands	8	0.0007	(-0.001196, 0.001972)
	S01005516	North and East Isles	Shetland Islands	8	0.000704	(-0.001198, 0.001977)
	S01000753	Lomond Shore	Argyll & Bute	6	0.00117	(-0.001196, 0.002166)
	S01005360	Bishopton	Renfrewshire	10	0.08064	(0.009389, 0.232)
	S01005350	Bishopton	Renfrewshire	10	0.08157	(0.009443, 0.2348)
	S01005359	Bishopton	Renfrewshire	10	0.08161	(0.009487, 0.2353)
	S01005356	Bishopton	Renfrewshire	10	0.0823	(0.009712, 0.234)
	S01005354	Bishopton	Renfrewshire	10	0.08385	(0.009702, 0.2415)
	S01005358	Bishopton	Renfrewshire	6	0.1678	(0.01959, 0.479)
Highest 10 RRs	S01003302	Laurieston and Tradeston	Glasgow City	1	3.192	(2.297, 4.311)
	S01005590	Ayr North Harbour, Wallacetown and Newton South	South Ayrshire	1	3.22	(2.236, 4.448)
	S01003855	Inverness Merkinch	Highland	1	3.347	(2.334, 4.595)
	S01005425	Langlee	Scottish Borders	1	3.366	(2.433, 4.525)
	S01000748	Dunoon	Argyll & Bute	1	3.429	(2.378, 4.752)
	S01000749	Dunoon	Argyll & Bute	1	3.437	(2.361, 4.79)
	S01000717	Campbeltown	Argyll & Bute	1	3.551	(2.403, 4.989)
	S01005592	Ayr North Harbour, Wallacetown and Newton South	South Ayrshire	1	3.563	(2.665, 4.659)
	S01005594	Ayr North Harbour, Wallacetown and Newton South	South Ayrshire	1	3.751	(2.715, 5.019)
	S01005598	Ayr North Harbour, Wallacetown and Newton South	South Ayrshire	1	3.823	(2.842, 5.014)

Table 7.8: Table of Fitted Female Model C-u Alcohol-Related Relative Risks

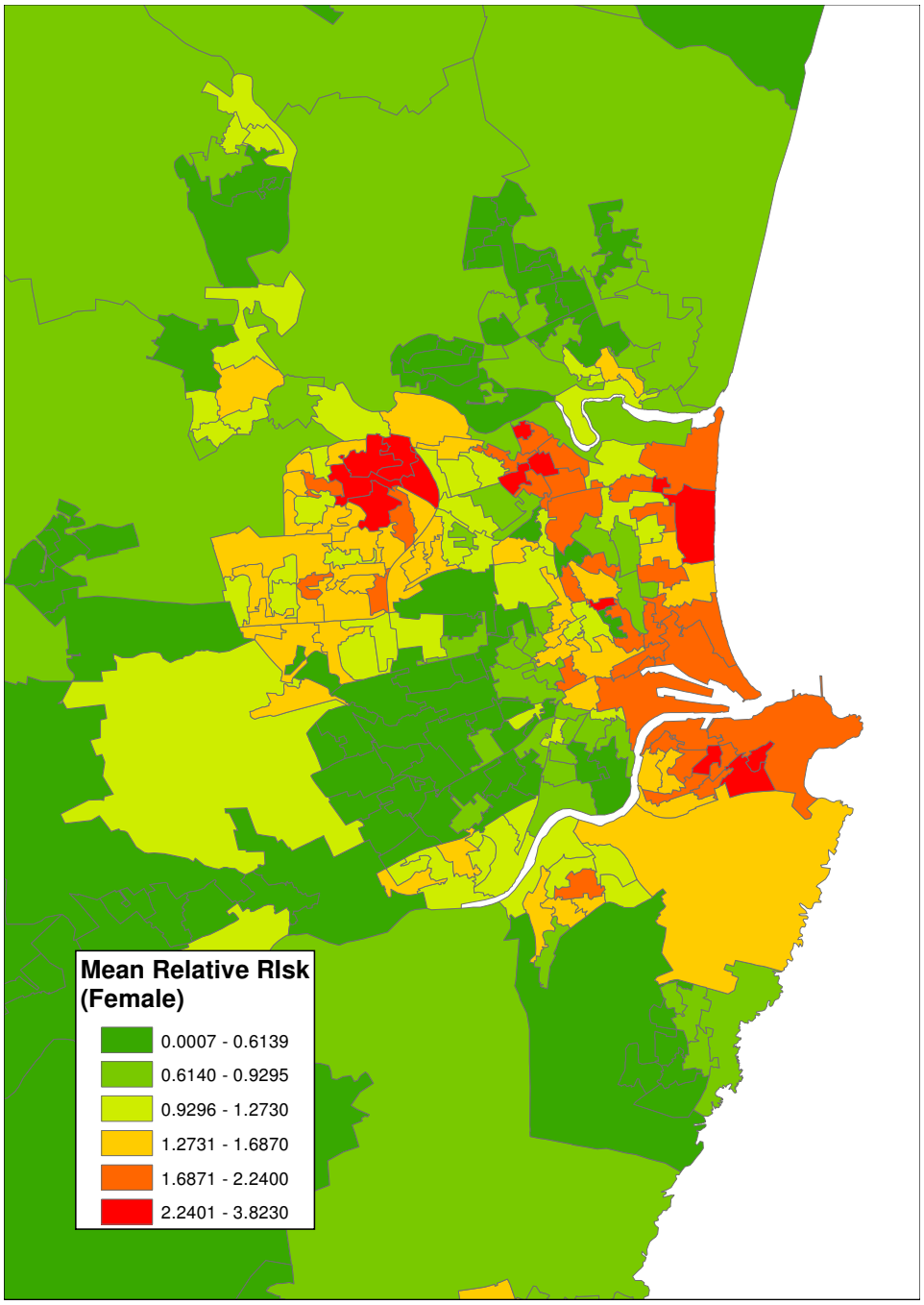


Figure 7.13: Aberdeen Area Data Zone Map of Mean Female Alcohol-Related Relative Risk

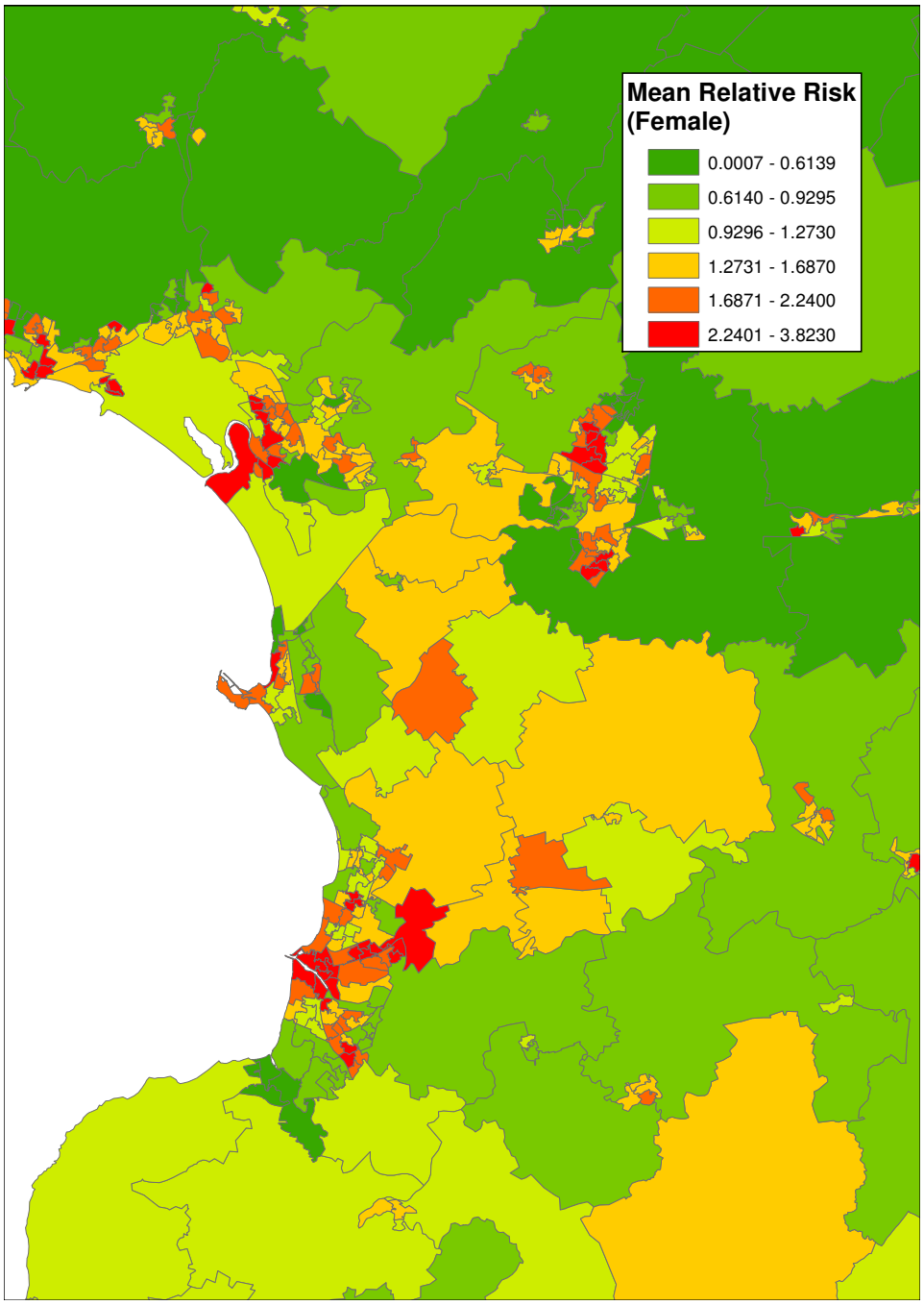


Figure 7.14: Ayrshire Area Data Zone Map of Mean Female Alcohol-Related Relative Risk

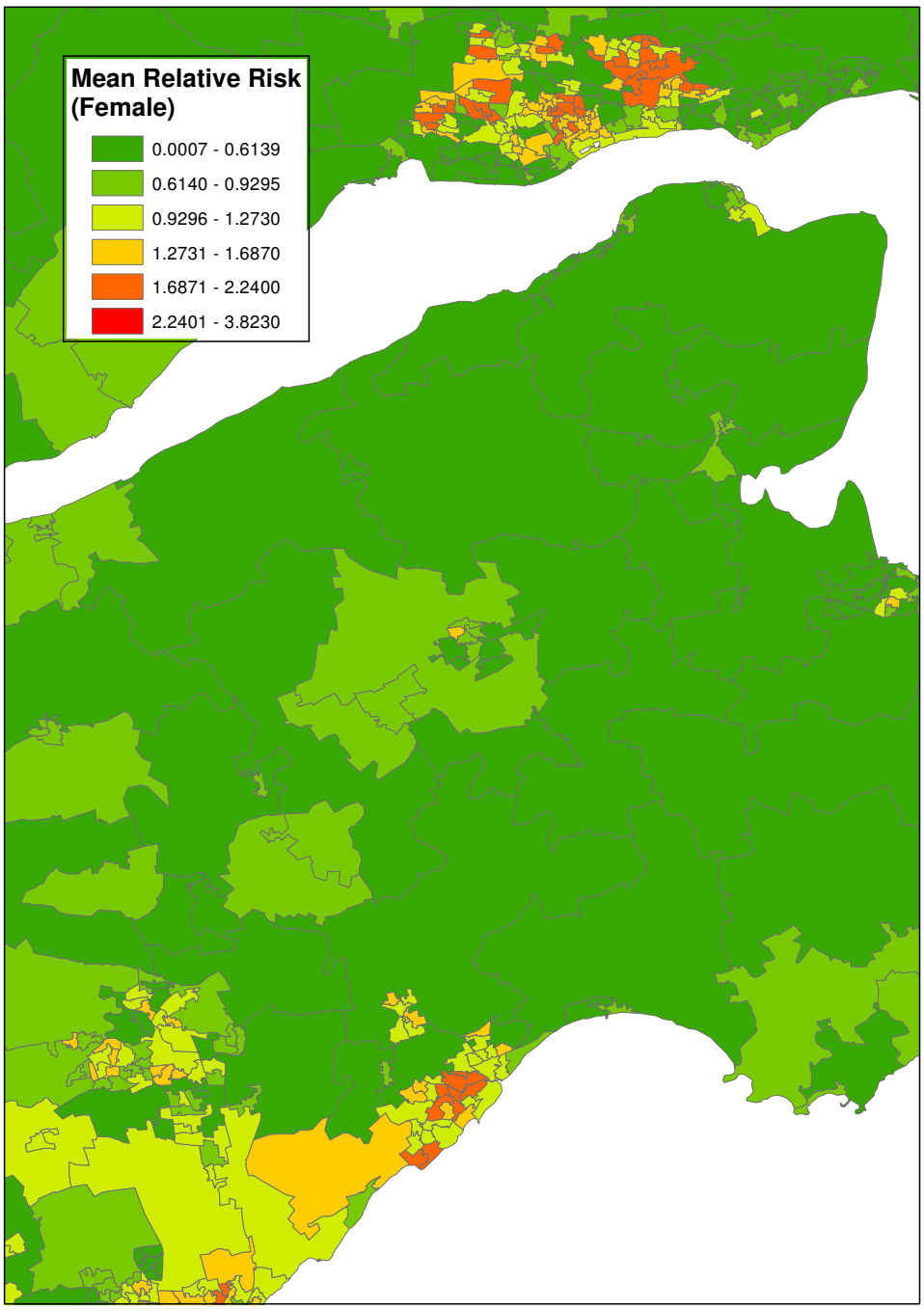


Figure 7.15: Dundee Area Data Zone Map of Mean Female Alcohol-Related Relative Risk

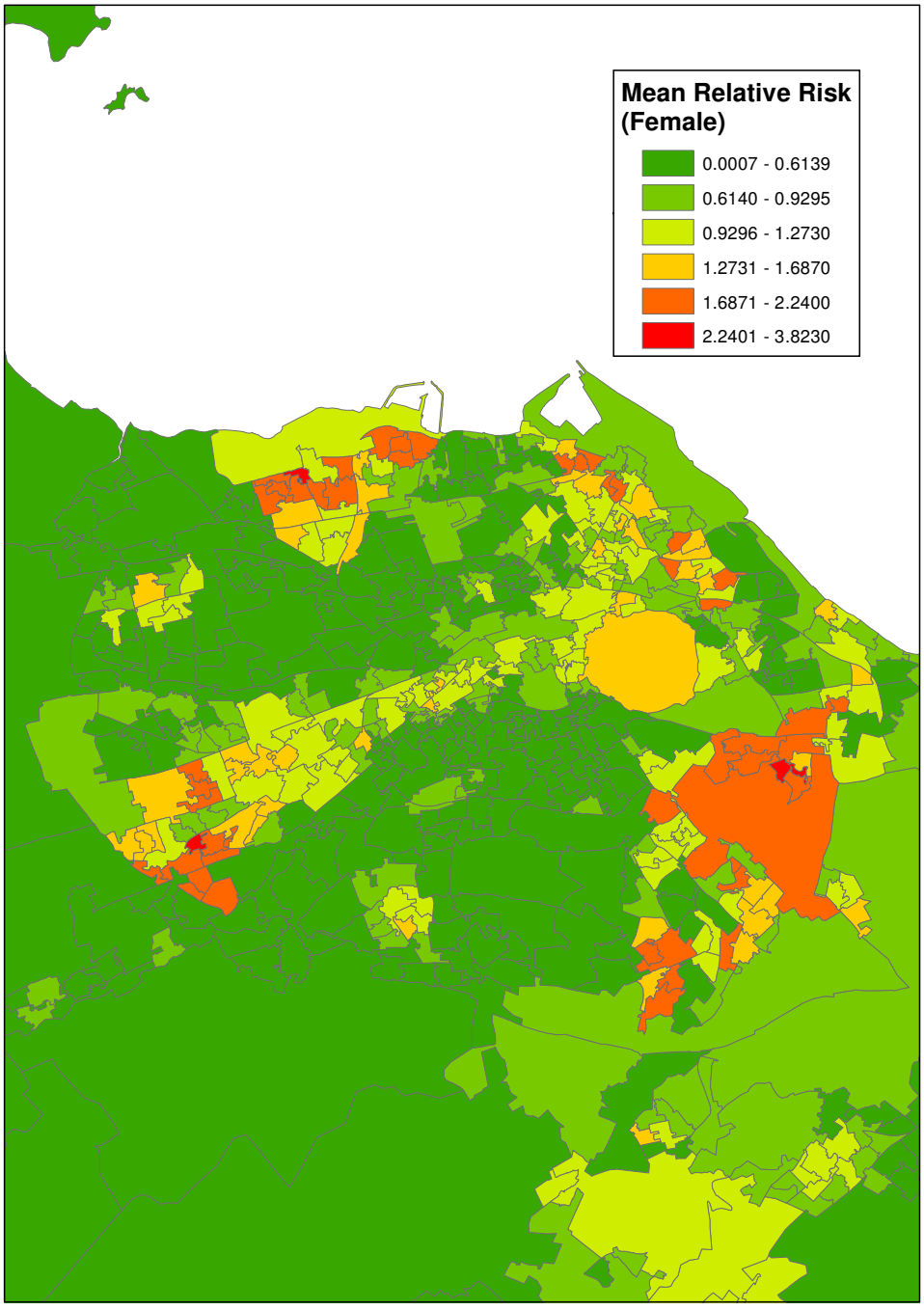


Figure 7.16: Edinburgh Area Data Zone Map of Mean Female Alcohol-Related Relative Risk

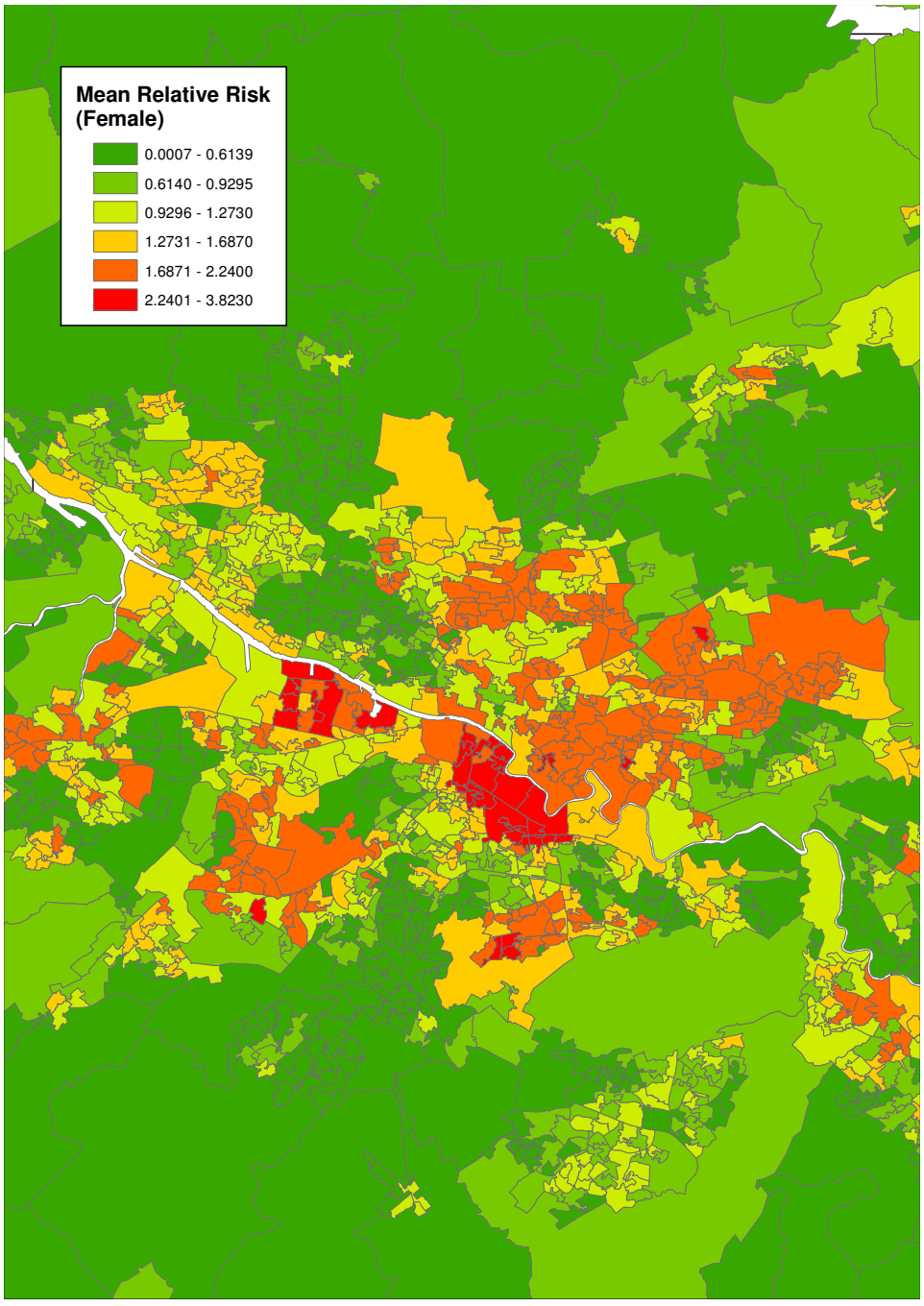


Figure 7.17: Glasgow Area Data Zone Map of Mean Female Alcohol-Related Relative Risk

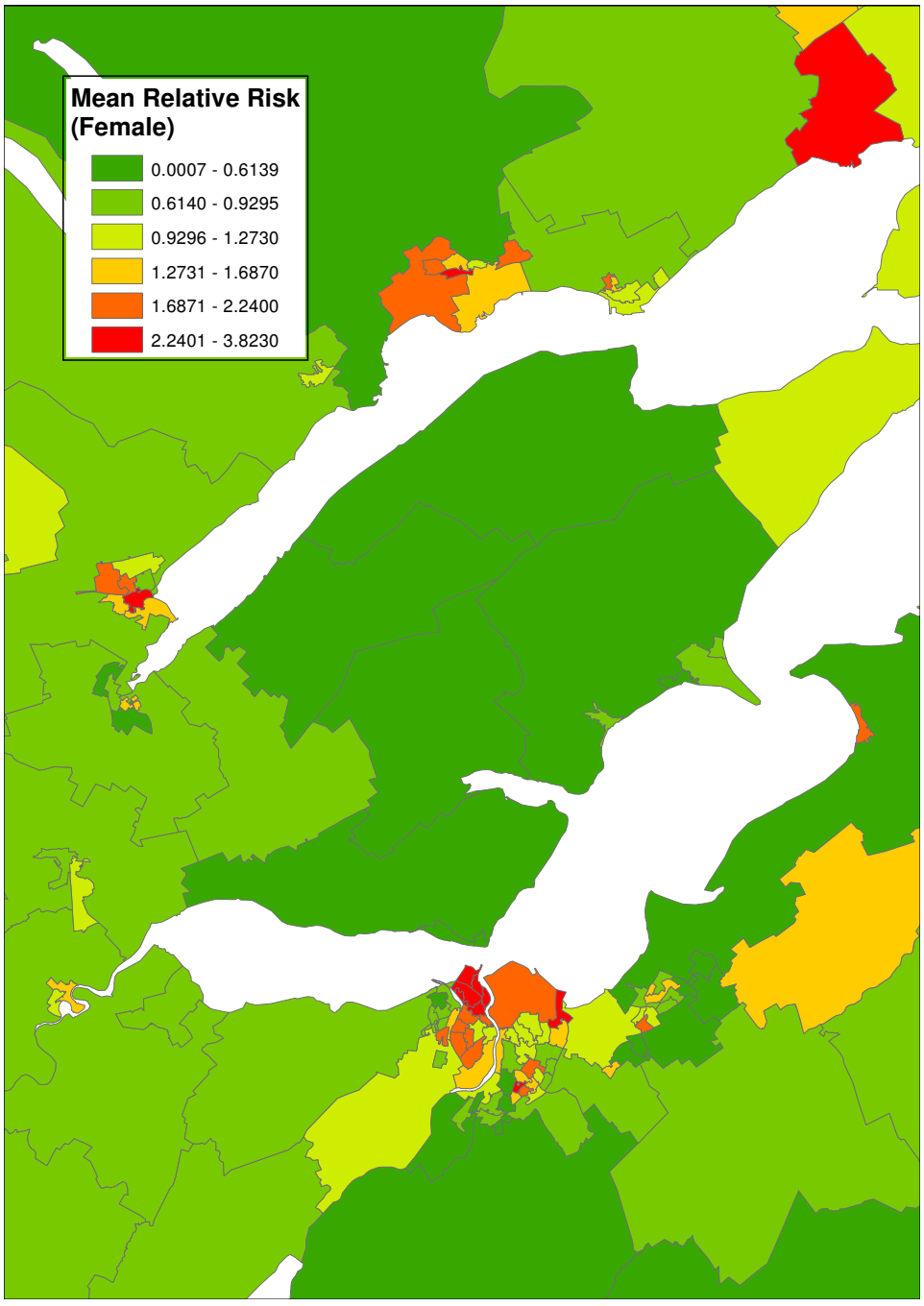


Figure 7.18: Inverness Area Data Zone Map of Mean Female Alcohol-Related Relative Risk

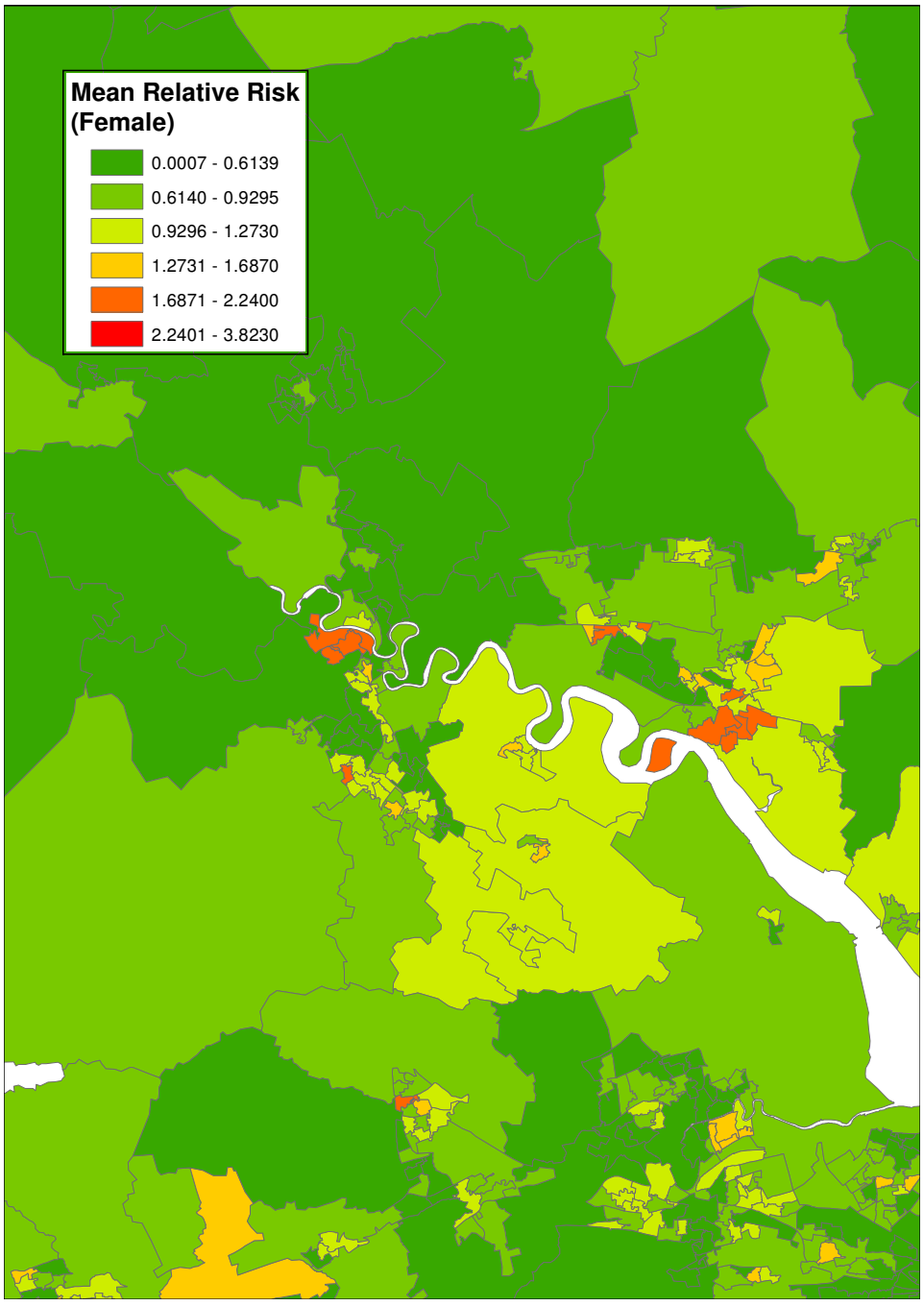


Figure 7.19: Stirling Area Data Zone Map of Mean Female Alcohol-Related Relative Risk

Chapter 8

Distance Models for Combined Male and Female Data

So far this thesis has modelled alcohol-related health risks across Scotland while accounting for area deprivation levels and the age and sex structure of the population. All previous modelling chapters find that data zone deprivation score is a significant fixed effect. However, in each of the chosen models some variation in relative risk values persists after the deprivation variable is fitted. This chapter will consider a further fixed effect, in addition to area deprivation score, in an attempt to explain some of the remaining variation in Scottish alcohol-related health risks.

The additional covariate considered here is the minimum Euclidian distance from each data zone centroid to a single malt whisky distillery. Further discussion on how these distances are estimated is given in Chapter 2. It is well documented that Scotland has a strong history in whisky production and it is possible that close proximity to such distilleries may have an effect on alcohol-related risk. Such a link may arise because staff discounts are offered to employees of the distilleries.

It is also known that alcohol consumption levels are influenced by social and cultural attitudes towards the substance. A more remote possibility is that, given the old age of many of the distilleries, their existence in a given

area may prove to be an indicator of the historical and general attitude towards alcohol within that area.

It must be remembered that any potential link found between alcohol-related risk and proximity to a single malt whisky distillery will not necessarily be causal. If the results show that there is a significant relationship between alcohol-related risk and distillery proximity, it will not be possible to say that being close to a distillery causes higher or lower risk, just that there is an association between either wide/close proximity and high risk rates.

Distance models will only be examined for the combined male and female data in the first instance. The data used is exactly the same as that used for the combined models in Chapter 5, but with the addition of the estimated minimum distance from each data zone to a single malt whisky distillery.

8.1 Models Considered

This Chapter will fit 15 new models to the combined male and female death and hospitalisation data. Like previous chapters these models are based on the Besag, York and Mollié model which is a Bayesian relative risk model that incorporates spatial random effects. The models will vary in terms of both fixed and random effects. As random effects each model will contain either correlated heterogeneity (u), uncorrelated heterogeneity (v) or a convolution prior ($u + v$).

Two fixed effects are considered in these models; area deprivation score and the minimum distance to a distillery. Since the expected count data used here has already been standardised for age and sex it is not necessary to include these variables in the modelling process. Deprivation has been modelled in two different ways; firstly in a linear manner and secondly by assigning a separate parameter to each of the ten bona fide deprivation scores. In all 15 models an additive-link distance effect has been included, which will be discussed below. Models which consider deprivation alone for the

combined data have already been explored in Chapter 5 and they will not be run again in this section.

The first check for a spatial association between alcohol-related health risk and single malt whisky distilleries is to see if there is a decline in relative disease risk with increased distance from a distillery. It has been decided to use an additive-link distance effect to fit the distance covariate in all models since for radial distance decline the background rate of risk is believed to be unaffected at large distances. Discussions of similar model structures are given in section 7.7.1 of Lawson et al. (2003*a*).

Since the models in this Chapter examine two fixed effects it is necessary to consider appropriate interaction terms. This is because any association between the distance to a distillery and alcohol-related risk may not be the same for areas with differing deprivation levels.

Table 8.1 names all 15 distance models explored and gives a summary of the fixed and random effects included in each.

Model Name	**Fixed Effects**	**Random Effects**
Distance Model A-v	distance	v
Distance Model A-u	distance	u
Distance Model A	distance	$u+v$
Distance Model B-v	distance & linear deprivation	v
Distance Model B-v-Int	distance, linear deprivation & interaction term	v
Distance Model B-u	distance & linear deprivation	u
Distance Model B-u-Int	distance, linear deprivation & interaction term	u
Distance Model B	distance & linear deprivation	$u+v$
Distance Model B-Int	distance, linear deprivation & interaction term	$u+v$
Distance Model C-v	distance & non-linear deprivation	v
Distance Model C-v-Int	distance, non-linear deprivation & interaction term	v
Distance Model C-u	distance & non-linear deprivation	u
Distance Model C-u-Int	distance, non-linear deprivation & interaction term	u
Distance Model C	distance & non-linear deprivation	$u+v$
Distance Model C-Int	distance, non-linear deprivation & interaction term	$u+v$

Table 8.1: Distance Model Names and Descriptions

As explained in section 3.6, the Besag, York and Mollié model assumes

that the relative risk in area i, θ_i, is given by

$$\theta_i = exp(\alpha + u_i + v_i)$$

where $exp(\alpha)$ is the baseline or overall level of relative risk. The models considered here are based on this but also incorporate the fixed effects discussed above.

Distance Model A fits only distance as a fixed effect, giving

$$\theta_i = exp\left(\alpha + log(1 + exp(-\alpha_2 m_i)) + u_i + v_i\right)$$

where m_i is the minimum distance from the centroid of area i to a single malt whisky distillery measured in meters and α_2 is a parameter. Given this structure the prior for α_2 must be constrained as numerical instability can arise if a vague prior distribution is used. All 15 models fitted in this chapter assign α_2 a normal prior with mean zero and variance 1, so

$$\alpha_2 \sim N(0, 1).$$

The justification of such an additive model is related to the behaviour of the multiplicative model when the distance m_i becomes large. Consider the alternative of a multiplicative model in which

$$\theta_i = exp(\alpha_0 + other_covariates - \alpha_1 * m_1), \tag{8.1}$$

then at large distances the whole risk level tends to zero. This is strictly not appropriate since the risk could easily be high at distance from a distillery due to reasons unrelated to the source. An additive link is therefore used since it keeps the background risk unaltered when m_i is large. The α_2 parameter is a distance decline parameter and when it is significant and positive this implies that there is a significant distance decline.

Distance Model B includes both linear deprivation and the distance factor, giving

$$\theta_i = exp\left(\alpha + log(1 + exp(-\alpha_2 m_i)) + \beta d_i + u_i + v_i\right)$$

where d_i represents the deprivation score in data zone i. In this model, and all models which include linear deprivation, the β parameter is assigned a diffuse normal prior with mean zero and precision e^{-5},

$$\beta \sim N(0, e^{-5}).$$

Since Distance Model B contains two fixed effects it is necessary to fit a second model for this combination which also includes a term to represent a possible interaction between these effects. Distance Model B-Int contains such an interaction term and models the relative risk in area i as

$$\theta_i = exp\left(\alpha + log(1 + exp(-\alpha_2 m_i)) + \beta d_i + \beta_2 d_i m_i + u_i + v_i\right)$$

where β_2 is given a vague normal prior distribution, $N(0, e^{-5})$.

Distance Model C also includes both distance and deprivation effects, but unlike Distance Model B non-linear deprivation is considered. Distance Model C fits the relative alcohol-related risk in data zone i as

$$\theta_i = exp\left(\alpha + log(1 + exp(-\alpha_2 m_i)) + \beta_{d_i} + u_i + v_i\right)$$

where there is a separate parameter, β_1 to β_{10}, for each of the ten deprivation scores. The parameter for the worst deprivation score of 1 has been arbitrarily set to zero and the remaining 9 parameters are given vague normal prior distributions, $\mathrm{N}(0, e^{-5})$. Therefore, for Distance Model C we have

$$\begin{aligned} \beta_1 &= 0 \text{ and} \\ \beta_j &\sim N(-, e^{-5}), \end{aligned}$$

for j in 2:10.

Again, there must be a further model fitted to the combined data which includes an interaction term between the two fixed effects included in Distance Model C. An appropriate interaction term is included in Distance Model C-Int, which gives the relative risk in area i as,

$$\theta_i = exp\left(\alpha + log(1 + exp(-\alpha_2 m_i)) + \beta_{d_i} + \beta_{2_{d_i}} m_i + u_i + v_i\right).$$

Distance Model C-Int introduces a further 10 parameters, β_{2_1} to $\beta_{2_{10}}$, which apply to the interaction term. The parameter β_{2_1} which corresponds to the worst deprivation score of 1 is arbitrarily set to zero and β_{2_2} to $\beta_{2_{10}}$ are set to follow the same vague normal prior distributions as β_2 to β_{10}.

All models discussed in this chapter have been simulated using the OpenBUGS software. The OpenBUGS code for Distance Model A, Distance Model B, Distance Model B-Int, Distance Model C and Distance Model C-Int is given in Appendix sections 10.4 to 10.8 respectively. The code for all 15 distance models considered can be derived from these scripts by omiting the redundant parts of the code, for example by deleting all references to u in the Distance Model A code to obtain the code for Distance Model A-v.

8.2 Convergence

The distance models investigated simulate a separate relative risk parameter and in some cases two random effects for every single area. Given that there are 6505 data zones in the study it proved impossible to record these parameter values at every iteration due to computer memory limitations. For all 15 distance models a summary monitor has been set for all relative risk and random effect parameters. A subset of the relative risk parameters have also been fully monitored in order to assess convergence and are given in Table 8.2. All other model parameters have been fully monitored.

All previous models considered for the combined alcohol-related relative risk, discussed in Chapter 5, were found to display strong evidence of adequate convergence after a burn-in period of 10,000 iterations. Two post burn-in chains of 150,000 simulations were run for each of these models.

It is desirable to be able to compare the fit of the distance models to the earlier deprivation-only models of combined relative risk. For this reason all 15 distance models have also been run for 10,000 burn-in iterations followed by two simulation chains of length 150,000 from different starting points.

Data Zone Code	Relative Risk Parameter	Reason Chosen
S01006393	θ_{115}	poor deprivation score
S01006438	θ_{14}	poor deprivation score
S01006490	θ_{1}	good deprivation score
S01006505	θ_{2}	good deprivation score
S01003744	θ_{2521}	rural area
S0100391	θ_{2692}	rural area
S01003380	θ_{3044}	urban/city area
S01002325	θ_{4687}	urban/city area
S01005521	θ_{985}	island / no neighbouring areas
S01000447	θ_{6238}	island / no neighbouring areas
S01003031	θ_{2885}	lowest total population
S01000799	θ_{5792}	highest male population
S01002622	θ_{3557}	highest female population
S01003313	θ_{3046}	highest male SIR
S01006473	θ_{89}	zero female SIR value
S01006341	θ_{172}	zero female SIR value
S01003043	θ_{2889}	very high male SIR

Table 8.2: Data Zones with Fully Monitored Relative Risk Estimates

Convergence of these models was checked in the same way as previous chapters. Very similar levels of convergence were observed for all 15 models, so only the results for Distance Model B-u will be discussed in detail.

Firstly, history plots for a selection of the relative risk and other parameters from Distance Model B-u are shown in Figures 8.1, 8.2, 8.3 and 8.4. These plot all post burn-in simulated parameter values against the corresponding iteration number, showing both chains on the same plot. Colour is used to distinguish between the different sets of results, with one chain being shown in red and the other in blue.

With the exception of α_2, these plots all exhibit extremely good convergence characteristics with the points forming a horizontal band across each plot which exhibit no patterns or trends. Although the convergence of α_2 is not as good as for the other non-relative-risk parameters, there are no obvious patterns or trends in the history plot and the values of both chains consistently overlap. It appears that this parameter would benefit if the model was run for a longer period. However, given the limited time available

and the fact that all distance and previous combined relative risk models would have to be re-run with a longer chain length, coupled with the fact that convergence currently looks fairly reasonable for this variable, it has been decided to use these results. The reason that all models should be run with the same chain length is that the p*D method of calculating DIC is based on the variance of the simulated deviance values. While under perfect convergence the posterior variance should not vary with the number of chain iterations, under only adequate convergence it is possible that the posterior variance may decrease as the number of simulations increases, all be it very slightly.

The history plots also show that all of the relative risk parameters have converged very well, with the exception of θ_{985} (or RR[985]) which corresponds to an island/ neighbourless area. This parameter still appears to have achieved convergence since both chains consistently overlap and do not exhibit any trends or patters, but like α_2 it appears that it would benefit from a longer chain length.

For the same subset of Distance Model B-u parameters the Gelman-Rubin diagnostics, as discussed in section 3.4.3, are given in Figure 8.5 and Figure 8.6. All of these plots suggest adequate model convergence. The green line, which represents the width of the central 80% interval of the pooled chains, and the blue line, which shows the average width of the 80% intervals within the individual chains, are both stable and the red line which represents their ratio is stable at a value of 1. In fact, the intervals are so similar for the individual and the pooled chains, that the blue line almost completely obscures the green line.

The density plots from Distance Model B-u were also examined for these parameters and are shown in Figures 8.7 and 8.8. These plots add further evidence that the model has converged adequately. The plots show a very smooth density for all fully monitored parameters apart from α_2 and θ_{985}. Although not completely smooth, the densities for these parameters do closely

resemble normal densities and do not appear to be particularly multimodal.

Given the above evidence, it has been assumed that all 15 distance models of combined alcohol-related relative risk have a achieved adequate convergence. However, with more time it would be better to run these models with a longer simulation chain length.

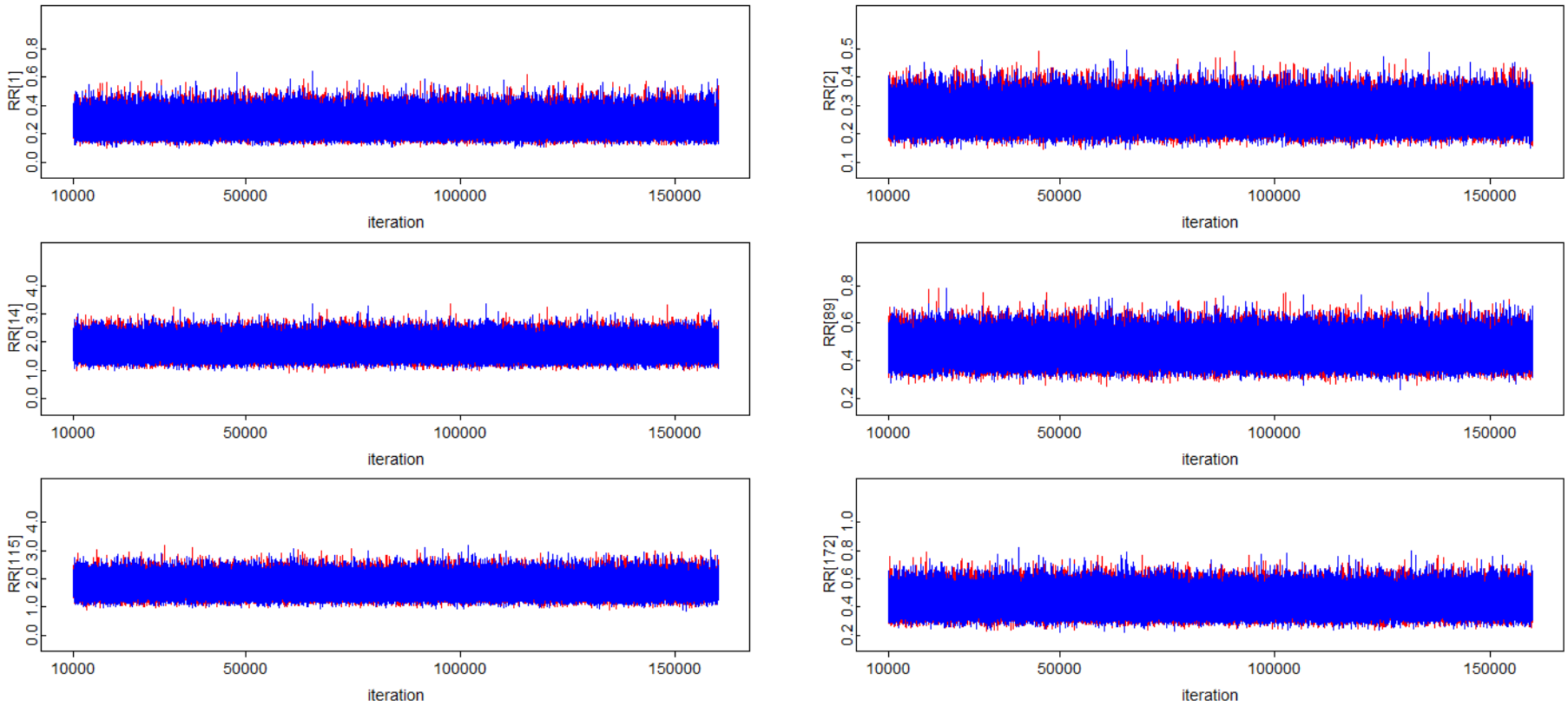


Figure 8.1: Simulation History Plots for a Subset of Distance Model B-u Parameters (part 1)

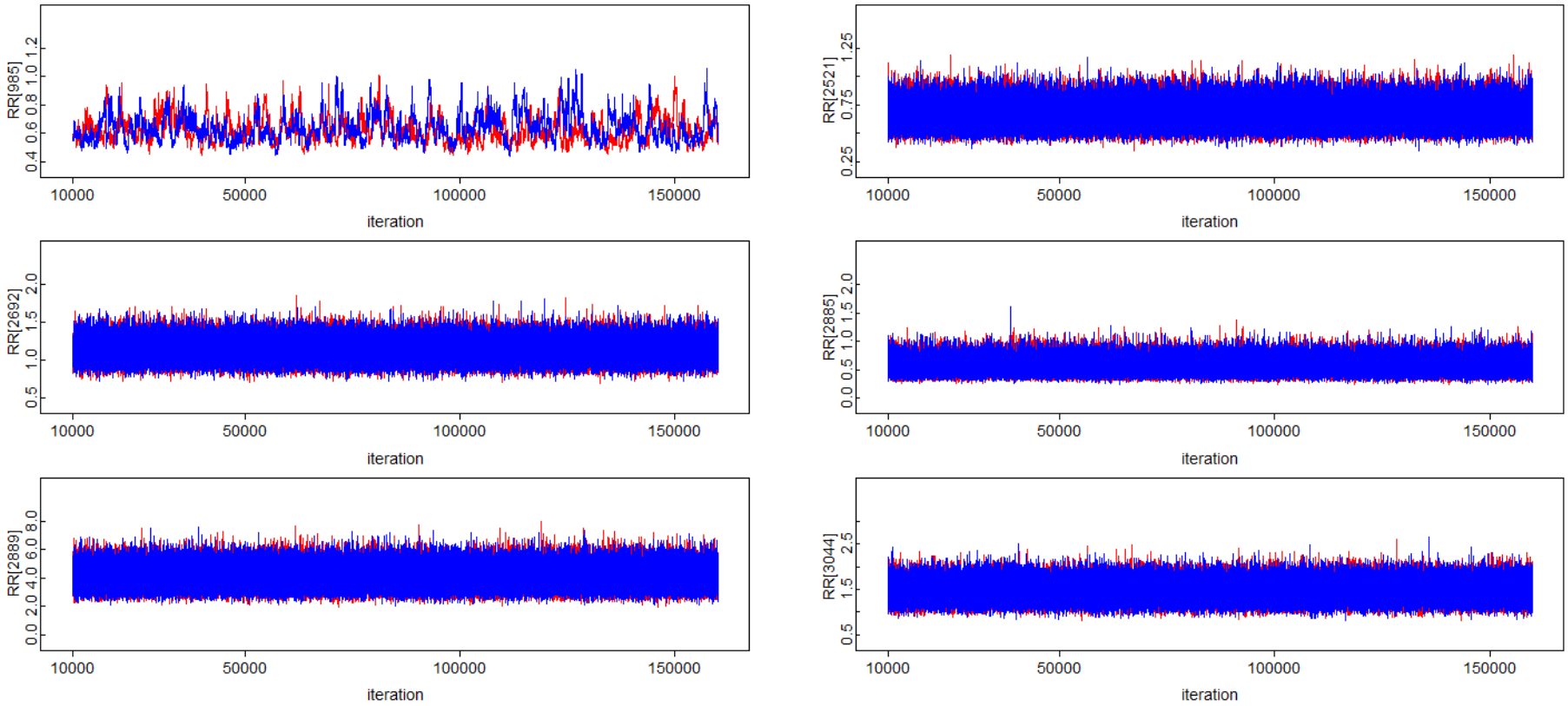


Figure 8.2: Simulation History Plots for a Subset of Distance Model B-u Parameters (part 2)

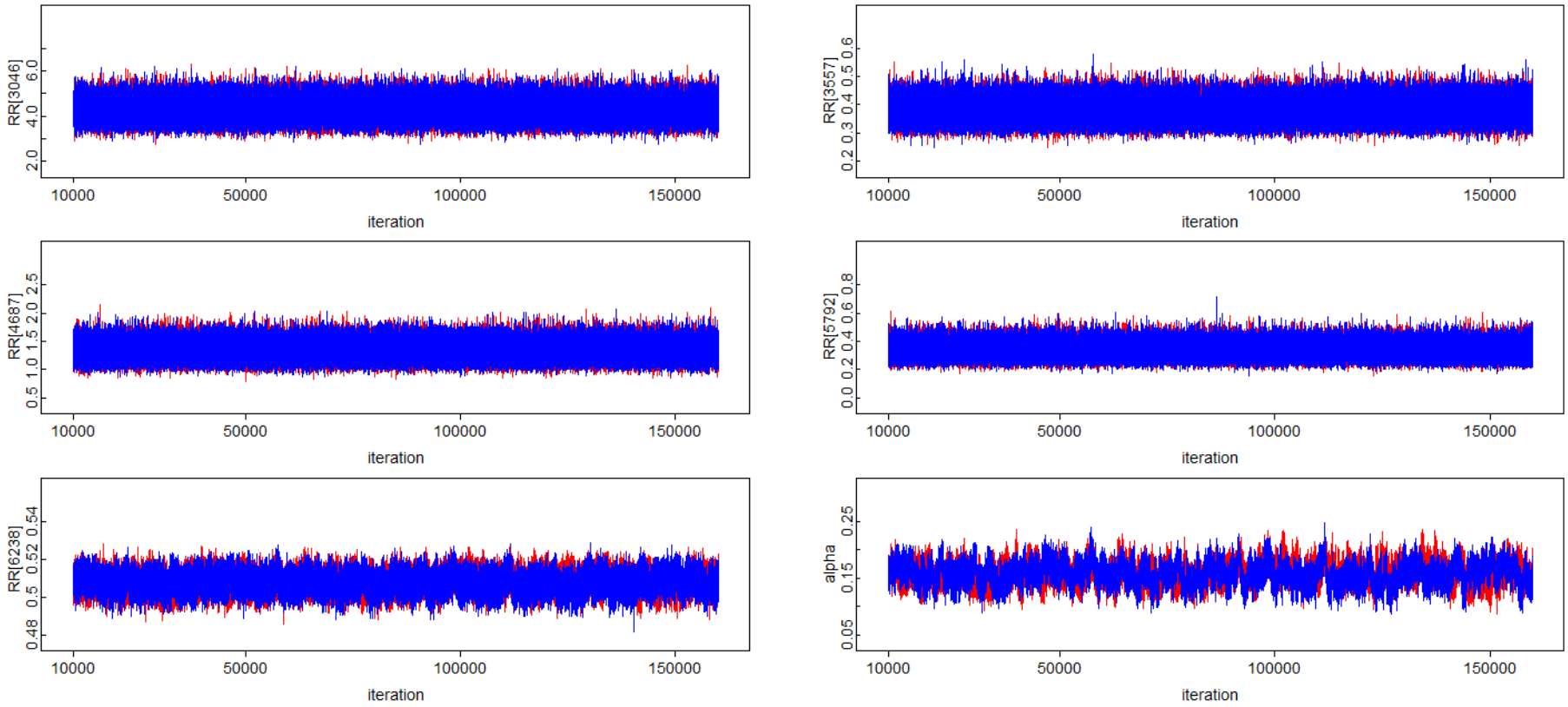


Figure 8.3: Simulation History Plots for a Subset of Distance Model B-u Parameters (part 3)

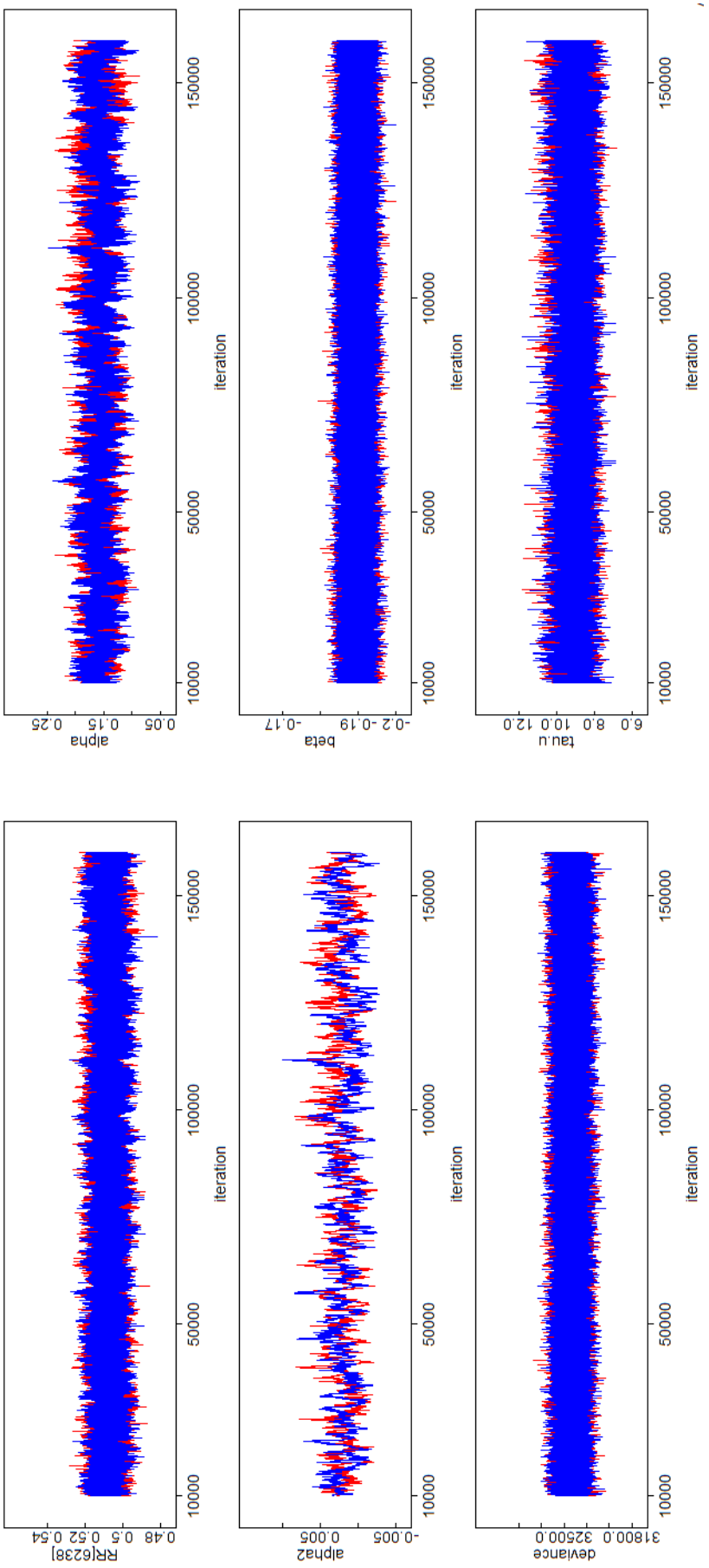


Figure 8.4: Simulation History Plots for a Subset of Distance Model B-u Parameters (part 4)

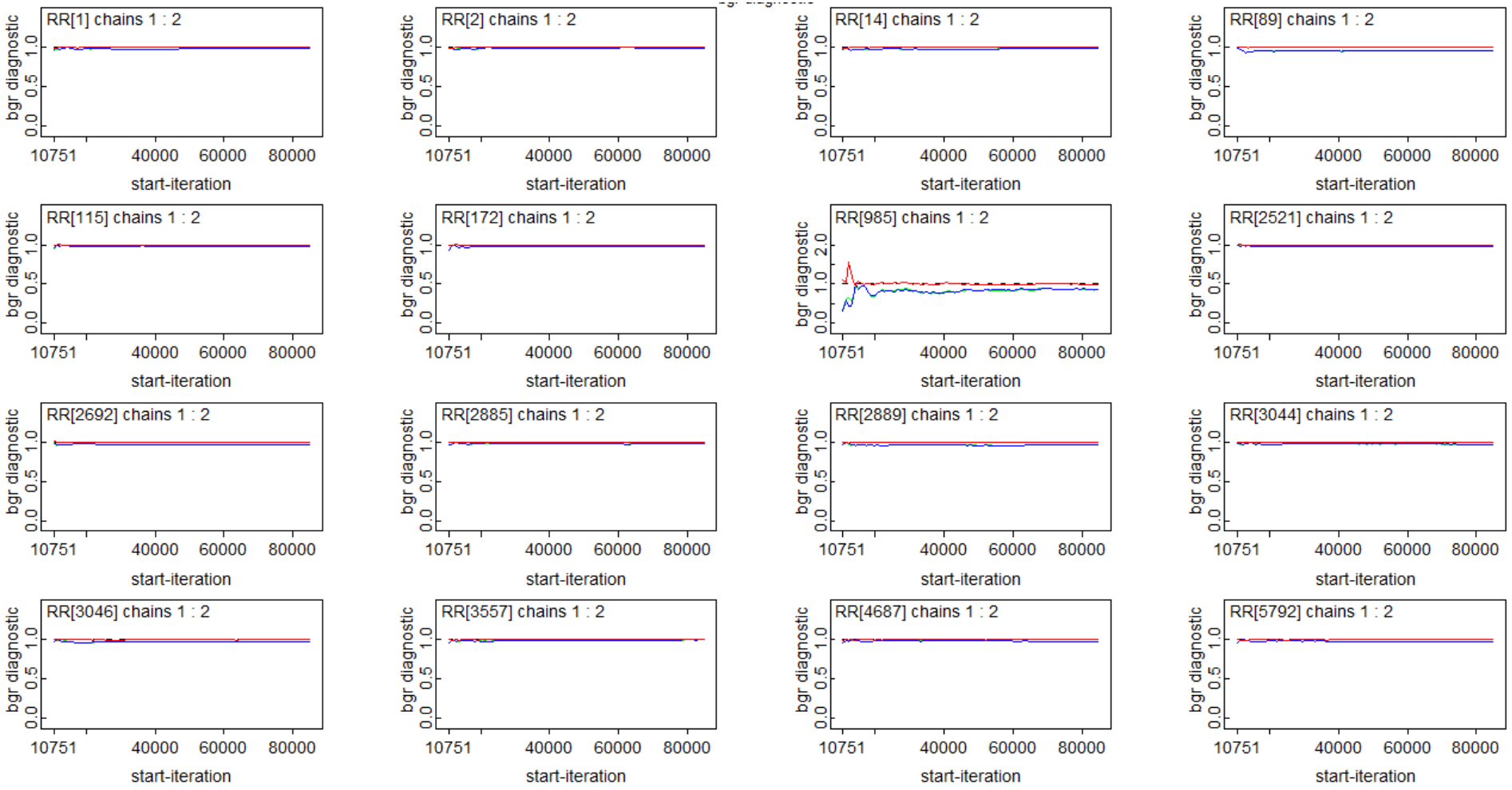


Figure 8.5: BGR Diagnostic Plots for a Subset of Distance Model B-u Parameters (part 1)

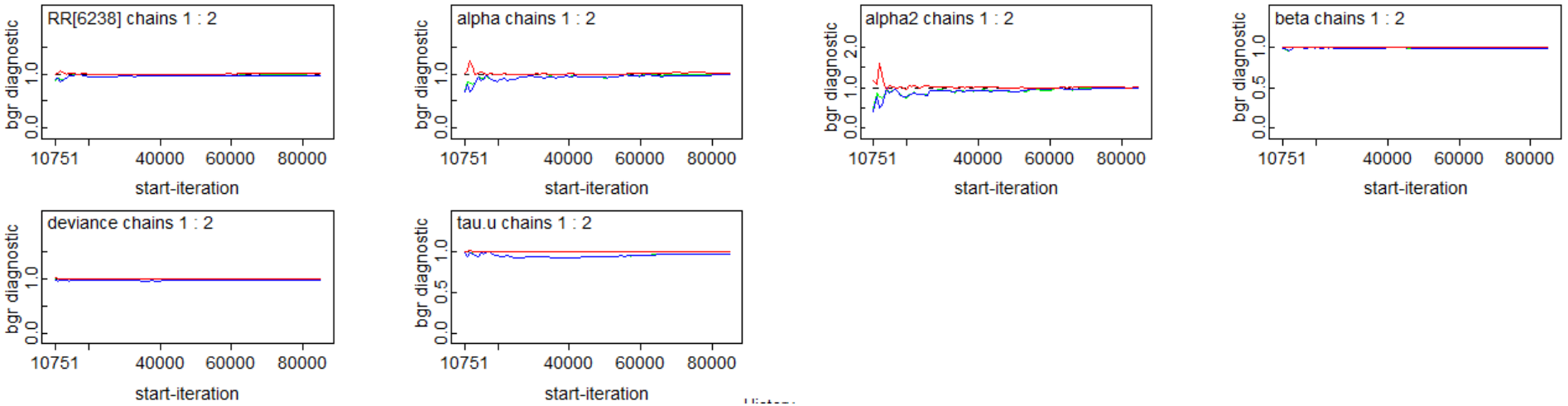


Figure 8.6: BGR Diagnostic Plots for a Subset of Distance Model B-u Parameters (part 2)

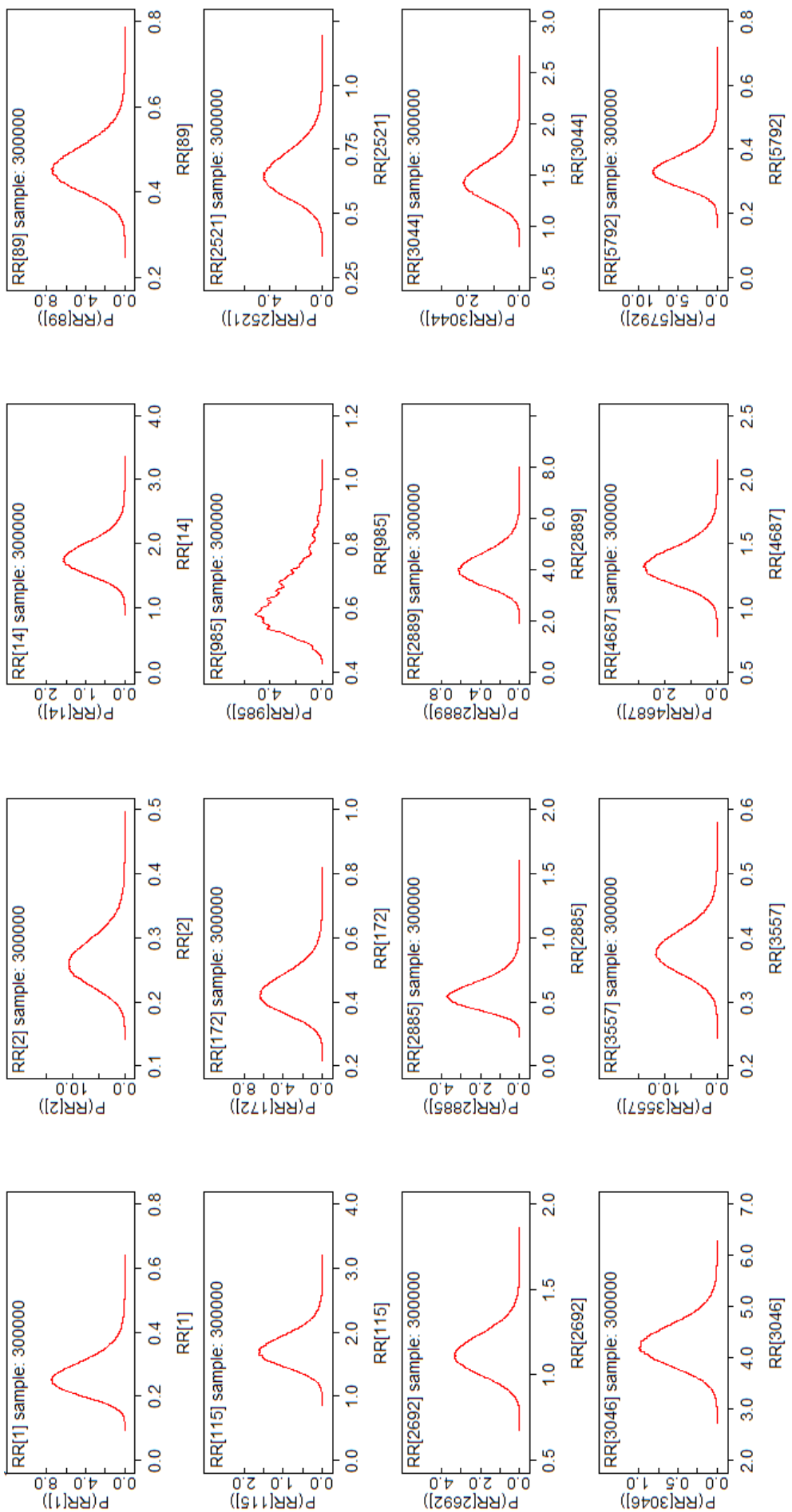


Figure 8.7: Posterior Density Plots for a Subset of Distance Model B-u Parameters (part 1)

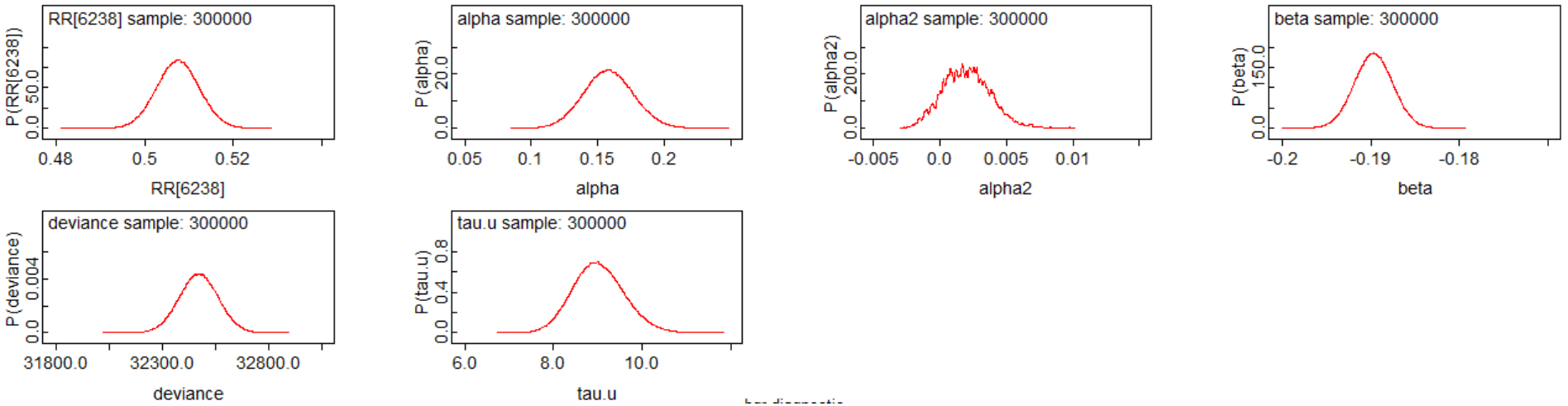


Figure 8.8: Posterior Density Plots for a Subset of Distance Model B-u Parameters (part 2)

8.3 DIC

The DIC measure of goodness-of-fit will be used to choose the best fitting distance model of combined alcohol-related relative risk. However, it may be the case that the earlier models for the combined relative risks provide a better fit to the data. The DIC values will therefore be compared between all 15 distance models discussed in this Chapter and all 9 deprivation models considered in Chapter 5.

Since the distance model DIC results will be compared with the DIC results in Chapter 5, it has been decided to start with the p*D method of calculating DIC. The DIC values, calculated using the p*D method, for all distance models are given in Table 8.3.

Model Name	$\overline{\mathbf{D}}$	$\widehat{\mathbf{var}}\{\mathbf{D}\}$	$\mathbf{p^*D}$	**DIC**
Distance Model A-v	32090	12927.7	6463.9	38553.9
Distance Model A-u	32330	12100	6050	38380
Distance Model A	32150	12633.8	6316.9	38466.9
Distance Model B-v	32070	11837.44	5918.72	37988.7
Distance Model B-v-Int	32080	11794	5897	37977
Distance Model B-u	32420	8363.1	4181.6	36601.6
Distance Model B-u-Int	32420	8374.1	4187.1	36607.1
Distance Model B	32060	10774.4	5387.2	37447.2
Distance Model B-Int	32060	10983.0	5491.5	37551.5
Distance Model C-v	32090	11772.3	5886.2	37976.2
Distance Model C-v-Int	32100	11837.4	5918.7	38018.7
Distance Model C-u	32430	8306.5	4153.3	36583.3
Distance Model C-u-Int	32430	8302.9	4151.4	36581.4
Distance Model C	32070	10836.8	5418.4	37488.4
Distance Model C-Int	32070	10629.6	5314.8	37384.8

Table 8.3: DIC for Distance Models using p*D

The lowest distance model DIC value is 36601.6. This corresponds to Distance Model B-u which includes distance decline and linear deprivation fixed effects and a correlated heterogeneity random effect. However, if we compare this value to the DIC values given in Table 5.4 the DIC for Model C-u is lower.

8.4 Model Selection

It is common practice to select the model with the lowest DIC value as the 'best' model. Using this method, the results obtained suggest that Model C-u is the best model for combined alcohol-relative risk that has been considered in this thesis. Previous discussion of Model C-u is given in sections 5.4 and 5.6, but most notably for this chapter it does not incorporate the distance effect. This suggests that the minimum distance between data zone centroid and a single malt whisky distillery, fitted as a distance-decline effect, does not explain a significant amount of the variation which remains after area deprivation is considered.

It must be remembered that all of these models are based on assumptions which are set by the investigator by way of the prior and hyperprior distributions for the parameters and hyperparameters. Although proximity to a single malt distillery has proved to be insignificant here, this may not be the case under different assumptions. With limited time to complete this project a comprehensive sensitivity analysis is not possible. However, sensitivity to the gamma(0.5,0.0005) hyperpriors assigned to τ_u^2 and τ_v^2 will be examined.

8.5 Hyperprior Sensitivity Analysis

Every distance model listed in Table 8.1 will be re-run with alternative hyperprior distributions assigned to τ_u^2 and τ_v^2. These sensitivity models will use the same names as those given in Table 8.1 but with 'Sens' appended at the end, for example 'Distance Model A-Sens'. The reason for running the models with alternative hyperpriors is to see whether the choice of hyperprior will affect model selection. If a distance model is chosen under the different assumptions, then the risk estimates from that model will be compared with those from Model C-u.

The alternative hyperpriors considered in each model (where required)

are

$$\tau_u^2 \sim \text{Gamma}(1,1) \text{ and}$$
$$\tau_v^2 \sim \text{Gamma}(1,1).$$

These are the same alternative priors used for the earlier combined relative risk models and, as discussed in section 5.5, they are much less vague and very different from the original hyperpriors used.

Again, since the sensitivity results for the distance models will be compared with those for the previous combined models of relative risk in chapter 5, the only DIC values considered will be those calculated using the p*D method. For each distance sensitivity model the DIC value and related statistics are given in Table 8.4.

Model Name	$\overline{\mathbf{D}}$	$\widehat{\mathbf{var}}\{\mathbf{D}\}$	$\mathbf{p^*D}$	**DIC**
Distance Model A-v-Sens	32080	13018.8	6509.4	38589.4
Distance Model A-u-Sens	32330	12078.0	6039.0	38369.0
Distance Model A-Sens	32070	12122.0	6061.0	38131.0
Distance Model B-v-Sens	32050	11556.3	5778.1	37828.1
Distance Model B-v-Int-Sens	32060	11664.0	5832.0	37892
Distance Model B-u-Sens	32390	8208.4	4104.2	36494.2
Distance Model B-u-Int-Sens	32390	8210.2	4105.1	36495.1
Distance Model B-Sens	31820	8602.6	4301.3	36121.3
Distance Model B-Int-Sens	31820	8617.4	4308.7	36128.7
Distance Model C-v-Sens	32060	11642.4	5821.2	37881.2
Distance Model C-v-Int-Sens	32070	11642.4	5821.2	37891.2
Distance Model C-u-Sens	32400	8184.8	4092.4	36492.4
Distance Model C-u-Int-Sens	32400	8040.7	4020.4	36420.4
Distance Model C-Sens	31820	8613.7	4306.8	36126.8
Distance Model C-Int-Sens	31820	8561.8	4280.9	36100.9

Table 8.4: DIC for Distance Sensitivity Models using p*D

Of all the distance relative risk sensitivity models, the lowest DIC value of 36100.9 was observed for Distance Model C-Int-Sens. This model contains non-linear deprivation, a distance decline effect, an interaction term between deprivation and distance along with both correlated and uncorrelated heterogeneity. If we compare this to Distance Model B-u which is the distance model with the lowest DIC when run using the original hyperpriors it can be

seen that their structures are quite different. Distance Model B-u contains terms for distance decline, linear deprivation and correlated heterogeneity. So the chosen distance models under both hyperprior distributions differ in terms of both random effects, interaction terms and fixed effect structure chosen.

These differences in model structure seem to be quite substantial. However, it may be the case that, although very different models have been assumed, the resulting relative risk estimates would be similar between these models. This, however, will not be investigated. This is because if the DIC values for the distance sensitivity models in Table 8.4 are compared with those for the earlier combined sensitivity models in Table 5.6 the lowest DIC value of 36092.6 is actually observed for Model C-Sens.

Thus, under both the original and alternative hyperprior distributions, it is suggested that a distance decline effect should not be included in the model for combined alcohol-related relative risks in Scotland.

A comparison of the Model C-u and Model C-Sens parameter estimates is given in section 5.5.

8.6 Model Results

Since under both sets of assumptions the lowest DIC value corresponds to a non-distance model of combined relative risk, fitted in chapter 5, the resulting risk estimates are the same as those discussed for Model C-u in section 5.6.

The maps of combined male and female alcohol-related relative risk across Scotland given in section 5.7 give the same results as this chapter since distance was not included in the chosen model.

It should be remembered that these results do not mean that there is no relationship between alcohol-related risk and proximity to a single malt whisky distillery. Instead they indicate that fitting the above distance-decline

effect does not appear to account for a significant amount of the unexplained variation in relative risk which remains after allowing for age, sex and area deprivation score. It may prove that using a different form of distance fixed effect, fitting the same model to similar but different data or running exactly the same models for longer chain lengths would yield different results.

Chapter 9

Discussion

This chapter aims to discuss the merits and shortcomings of this research as well as various areas in which future research could be carried out.

9.1 Summary of Results

The results of this study suggest that area deprivation score is significantly associated with alcohol-related health risk for both males and females. Both the selected male and female models suggest that the relative risks are best described by a non-linear area deprivation score effect and correlated heterogeneity. More precisely, it is suggested that both male and female alcohol-related health risk is higher in more deprived areas and that the risk in any given area is related to the risk in its neighbouring areas.

The given analysis offers insufficient evidence to suggest that there is an association between the combined male and female alcohol-related health risk and proximity to a single malt whisky distillery.

9.2 Merits of Project

This thesis has allowed alcohol-related health relative risk maps of Scotland to be produced at a much finer level of geography than ever before. This

allows the patterns in relative risk to be examined in much more detail and with less loss of information due to aggregation. Being able to map these health risks at a small area level may prove useful when trying to allocate alcohol-related funding appropriately between small community NHS centres such as general practice surgeries.

I feel that the inclusion of both deaths and hospitalisations due to alcohol as opposed to just deaths gives a much better indication of the patterns of problem drinking across Scotland. For example, many people who are heavy drinkers are also heavy smokers, but when aiming to identify the areas with the highest levels of problem drinking it makes sense to count someone who has been hospitalised due to cirrhosis of the liver even if they later died of lung cancer. Many previous studies focus on alcohol-related deaths only.

This thesis has produced disease maps and estimates of alcohol-related health relative risks across Scotland separately for males and females. This makes comparisons between male and female risk estimates straightforward and the use of colour-coded maps makes such comparisons very accessible to non-statisticians.

This project also ventured into new territories in Bayesian spatial modelling. I can find no existing papers which consider Bayesian spatial models similar to those used in this project with as many areas. The use of 6505 data zone areas caused some problems while running the models in OpenBUGS. I contacted the BUGS project at Cambridge regarding these problems and developments were made to the OpenBUGS software which fixed these issues.

9.3 Persisting Issues of Project

There are some issues highlighted throughout this work which would be worth some further investigation.

The negative pD values obtained for some of the combined, male and female models of relative risk are undesirable. It would be of interest to

carry out further investigations into why these values have arisen.

The complex spatial geography of Scotland, with its many islands, also presents a major challenge for the type of models used in this thesis. The difficulty is related to potential 'edge effects'. All final models for the alcohol-related relative risks in Scotland contain spatial heterogeneity, which means that the risk in any given area depends on the risk in its neighbouring areas. The selected models will therefore produce poorer relative risk estimates for areas which do not share a border with any other areas.

The default option in most GIS mapping packages, and in OpenBUGS, is to assume that if two areas are separated by a physical boundary such as a river or sea they are not neighbours. However, the neighbourhood structure used for the spatial modelling is intended to represent correlations in the underlying alcohol-realted health risk, due to similarities in environmental and socioeconomic risk factors, rather than just physical proximity. It could be argued that many of the Scottish islands may have similar levels of risk to other nearby islands, yet they would not be regarded as neighbours by default in the current analysis. Further research could be carried out in which these physical boundaries are treated differently, hence defining the neighbouring areas differently.

More importantly, the island/neighbourhood structure of Scotland may be at the heart of the convergence problems found for some models, which may in turn explain the problems found with the negative pD estimates in the model comparison criteria.

The alternative measure of pD, p$*$D, which has been used in the model selection process throughout this thesis, tends to over penalise more complex hierarchical models (as discussed in Chapter 3). Other model selection possibilities could have been explored. For example, since I had a prior belief that deprivation level should be incorporated into the final models of alcohol-related relative risk a more subjectivist approach could have been explored. However, since I had no strong *a priori* opinion about the specific form that

any such deprivation level relationship should take this would have proved difficult.

Limitations in computing capabilities have meant that not all parameters in the fitted models could be fully monitored. Parameters which were not fully monitored were assigned a summary monitor. Summary monitors output exact mean and standard deviation values based on the sample of simulated values for the monitored parameter, but only approximate 95% credible intervals. Ideally every parameter would be fully monitored in order to avoid these approximate intervals.

9.4 Areas for Further Research

I feel that there are many more areas in which this research could be expanded in the future.

The current research considers an association between proximity to a single malt whisky distillery and alcohol-related health risks. The current methods use the approximate minimum Euclidian distance between each data zone and a distillery, which is 'as the bird flies'. It would be desirable to try and estimate more 'real world' distances, possibly by considering the minimum length of road between a data zone and a distillery.

As well as investigating the effects of physical distance between areas and distilleries, it would be beneficial to look at the proportion of the population in each area that live within a certain distance of a distillery and/or that work in a distillery if such information is available.

A further element regarding a possible link between whisky production and alcohol-related health risks in Scotland would be to include all distilleries in Scotland and, possibly most importantly, whisky bottling plants. It is probably more common for workers to buy or take whisky home from places where it is held in bottles rather than casks or distillers. A further improvement which could be made is to develop and include a measure of

the scale of production or staff numbers at each distillery and bottling plant.

As well as investigating the effects of physical distance between areas and distilleries, it would be beneficial to look at the proportion of the population in each area that live within a certain distance of a distillery and the proportion of the population that work in a distillery if such information is available.

On top of age, sex, deprivation and proximity to whisky production there are other factors which are worth consideration in a model for alcohol-related health relative risk.

It would be interesting to incorporate the proportion of adults in each area that are in longterm relationships. Such data may not be readily available, but the proportion of adults who are married or in civil partnerships could be used as a proxy if more accurate data is not available.

A further area of interest is the number of premises with a late licence within a certain distance of a data zone. This would allow investigation into whether alcohol risk rates are higher or lower in areas with late licences. Being able to buy alcohol for longer may cause people to drink more, but on the other hand shorter legal drinking hours may increase the incidence of binge drinking before the premises close and gatherings in people's homes after closure in which they can drink for as long as they like.

When considering such small areas it may also be worth including factors which describe the religious population in each area. For example, without including such a factor there may be small areas that have a much lower alcohol-related risk estimate than expected given neighbouring risk values and deprivation level if there is a high muslim population in the area.

It could also be of interest to further the investigation into a link between alcohol-related health risks and deprivation by examining the individual components of deprivation separately.

As well as considering additional variables within the existing model structures, there are several alternative Bayesian spatial model structures

which could be explored including shared components models for males and females together, multivariate CAR models and mixture models.

Chapter 10

Appendices

10.1 Model A - OpenBUGS Code

```
model                                                                    (1)
{                                                                        (2)
for (i in 1:m)                                                           (3)
{                                                                        (4)
    # Poisson Likelihood for Observed Counts                             (5)
    y[i]~dpois(mu[i])                                                    (6)
    log(mu[i])<-log(e[i])+alpha+u[i]+v[i]                                (7)
    # Relative Risk                                                      (8)
      theta[i]<-exp(alpha+u[i]+v[i])                                     (9)
    # Prior distribution for the uncorrelated heterogenity              (10)
      v[i]~dnorm(0,tau.v)                                               (11)
}                                                                       (12)
eps<-1.0E-6                                                             (13)
#CAR distribution for the spatial correlated heterogenity               (14)
u[1:m]~car.normal(adj[],weights[],num[],tau.u)                          (15)
# Weights                                                               (16)
for (k in 1:sumNumNeigh)                                                (17)
{                                                                       (18)
    weights[k]<-1                                                       (19)
}                                                                       (20)
# Improper distribution for the mean relative risk in the study region  (21)
alpha~dflat()                                                           (22)
mean<-exp(alpha)                                                        (23)
# Hyperprior distributions on inverse variance paramenters of random effects (24)
tau.u~dgamma(0.5,0.0005)                                                (25)
tau.v~dgamma(0.5,0.0005)                                                (26)
var.u <- 1/tau.u                                                        (27)
var.v<- 1/tau.v                                                         (28)
}                                                                       (29)
```

10.2 Model B - OpenBUGS code

```
model                                                                    (1)
{                                                                        (2)
for (i in 1:m)                                                           (3)
{                                                                        (4)
    # Poisson likelihood for observed counts                             (5)
    y[i] ~ dpois(mu[i])                                                  (6)
    log(mu[i]) <- log(e[i]) + alpha + v[i] + u[i] +  beta*d[i]           (7)
    # Relative Risk                                                      (8)
    theta[i] <- exp(alpha + v[i] + u[i] + beta*d[i])                     (9)
    # Prior distribution for the uncorrelated heterogenity               (10)
    v[i] ~ dnorm(0, tau.v)                                               (11)

eps <- 1.0E-6                                                            (12)

# CAR prior distribution for spatial correlated heterogenity             (13)
u[1:m] ~ car.normal(adj[], weights[], num[], tau.u)                      (14)

# Weights                                                                (15)
for(k in 1:sumNumNeigh)                                                  (16)
{                                                                        (17)
    weights[k] <- 1                                                      (18)
}                                                                        (19)

# Improper prior distribution for the mean relative risk in the study region (20)
alpha ~ dflat()                                                          (21)
mean <- exp(alpha)                                                       (22)

# Prior on regression coefficients                                       (23)
beta ~ dnorm(0.0, 1.0E-5)                                                (24)

# Hyperprior distribution on inverse variance parameter of random effects (25)
tau.u ~ dgamma(0.5, 0.0005)                                              (26)
tau.v ~ dgamma(0.5, 0.0005)                                              (27)
var.u <- 1/tau.u                                                         (28)
var.v<- 1/tau.v                                                          (29)
}                                                                        (30)
```

10.3 Model C - OpenBUGS code

```
model                                                                    (1)
{                                                                        (2)
for (i in 1:m)                                                           (3)
{                                                                        (4)
    # Poisson likelihood for observed counts                             (5)
    y[i] ~ dpois(mu[i])                                                  (6)
    log(mu[i]) <- log(e[i]) + alpha + v[i] + u[i] +  beta[d[i]]          (7)
    # Relative Risk                                                      (8)
    theta[i] <- exp(alpha + v[i] + u[i] + beta[d[i]])                    (9)
    # Prior distribution for the uncorrelated heterogenity               (10)
    v[i] ~ dnorm(0, tau.v)                                               (11)
}                                                                        (12)

eps <- 1.0E-6                                                            (13)

# CAR prior distribution for spatial correlated heterogenity             (14)
u[1:m] ~ car.normal(adj[], weights[], num[], tau.u)                      (15)

# Weights                                                                (16)
for(k in 1:sumNumNeigh)                                                  (17)
{                                                                        (18)
    weights[k] <- 1                                                      (19)
}                                                                        (20)

# Improper prior distribution for the mean relative risk in the study region (21)
alpha ~ dflat()                                                          (22)
mean <- exp(alpha)                                                       (23)

# Prior on beta coefficients                                             (24)
beta[1] <- 0                                                             (25)
for (k in 2:10)                                                          (26)
    {                                                                    (27)
    beta[k] ~ dnorm(0.0, 1.0E-5)                                         (28)
    }                                                                    (29)

# Hyperprior distribution on inverse variance parameter of random effects (30)
tau.u ~ dgamma(0.5, 0.0005)                                              (31)
tau.v ~ dgamma(0.5, 0.0005)                                              (32)
var.u <- 1/tau.u                                                         (33)
var.v<- 1/tau.v                                                          (34)
}                                                                        (35)
```

10.4 Distance Model A - OpenBUGS code

```
model    {
                u[1:m] ~ car.normal( adj[], weights[], num[], tau.u)

                for ( i in 1:m )
                {
                    # Poisson likelihood for observed counts
                    y[i] ~ dpois( mu[i])
                    f[i] <- ( 1 + exp( -alpha2*dist[i] ) )
                    # log(f[i]) is an additive-link distance effect. If the estimate of alpha2 is positive
                    # then there is a decline with distance, which might be interpereted as
                    # significant if the alpha2 is well estimated.
                    log( mu[i] ) <- log(e[i]) + alpha + log(f[i]) + v[i] + u[i]
                    RR[i] <- mu[i]/e[i]
                    # Prior distribution for the uncorrelated heterogenity
                    v[i] ~ dnorm(0.0, tau.v)
                }

                eps <- 1.0E-6

                tau.u ~ dgamma( 0.5, 0.0005)
                tau.v ~ dgamma( 0.5, 0.0005)
                alpha2 ~ dnorm(0.0, 1)
                alpha ~ dflat()

                for ( k in 1:sumNumNeigh)
                {
                    weights[k] <- 1
                }
            }
```

10.5 Distance Model B - OpenBUGS code

```
model    {
                u[1:m] ~ car.normal( adj[], weights[], num[], tau.u)

                for ( i in 1:m )
                {
                    # Poisson likelihood for observed counts
                    y[i] ~ dpois( mu[i])
                    f[i] <- ( 1 + exp( -alpha2*dist[i] ) )
                    # log(f[i]) is an additive-link distance effect. If the estimate of alpha2 is positive
                    # then there is a decline with distance, which might be interpereted as significant
                    # if the alpha2 is well estimated.
                    log( mu[i] ) <- log(e[i]) + alpha + beta*dep[i] + log(f[i]) + v[i] + u[i]
                    RR[i] <- mu[i]/e[i]
                    # Prior distribution for the uncorrelated heterogenity
                    v[i] ~ dnorm(0.0, tau.v)
                }

                eps <- 1.0E-6

                tau.u ~ dgamma( 0.5, 0.0005)
                tau.v ~ dgamma( 0.5, 0.0005)
                beta ~ dnorm(0.0, 1.0E-5)
                alpha2 ~ dnorm(0.0, 1)
                alpha ~ dflat()

                for ( k in 1:sumNumNeigh)
                {
                    weights[k] <- 1
                }
            }
```

10.6 Distance Model B-Int - OpenBUGS code

```
model    {
                u[1:m] ~ car.normal( adj[], weights[], num[], tau.u)

                for ( i in 1:m )
                {
                    # Poisson likelihood for observed counts
                    y[i] ~ dpois( mu[i])
                    f[i] <- ( 1 + exp( -alpha2*dist[i] ) )
                    # log(f[i]) is an additive-link distance effect. If the estimate of alpha2 is positive
                    # then there is a decline with distance, which might be interpereted as significant
                    # if the alpha2 is well estimated.
                    log( mu[i] ) <- log(e[i]) + alpha + beta*dep[i] + log(f[i]) + beta2*dep[i]*dist[i] + v[i] + u[i]
                    RR[i] <- mu[i]/e[i]
                    # Prior distribution for the uncorrelated heterogenity
                    v[i] ~ dnorm(0.0, tau.v)
                }

                eps <- 1.0E-6

                tau.u ~ dgamma( 0.5, 0.0005)
                tau.v ~ dgamma( 0.5, 0.0005)
                beta ~ dnorm(0.0, 1.0E-5)
                beta2 ~ dnorm(0.0, 1.0E-5)
                alpha2 ~ dnorm(0.0, 1)
                alpha ~ dflat()

                for ( k in 1:sumNumNeigh)
                {
                    weights[k] <- 1
                }
            }
```

10.7 Distance Model C - OpenBUGS code

```
model    {
                u[1:m] ~ car.normal( adj[], weights[], num[], tau.u)

                for ( i in 1:m )
                {
                    # Poisson likelihood for observed counts
                    y[i] ~ dpois( mu[i])
                    f[i] <- ( 1 + exp( -alpha2*dist[i] ) )
                    # log(f[i]) is an additive-link distance effect. If the estimate of alpha2 is positive
                    # then there is a decline with distance, which might be interpereted as significant
                    # if the alpha2 is well estimated.
                    log( mu[i] ) <- log(e[i]) + alpha + beta[dep[i]] + log(f[i]) + v[i] + u[i]
                    RR[i] <- mu[i]/e[i]
                    # Prior distribution for the uncorrelated heterogenity
                    v[i] ~ dnorm(0.0, tau.v)
                }

                eps <- 1.0E-6

                tau.u ~ dgamma( 0.5, 0.0005)
                tau.v ~ dgamma( 0.5, 0.0005)

                # Prior on beta coefficients
                beta[1] <- 0
                for (j in 2:10)
                {
                    beta[j] ~ dnorm(0.0, 1.0E-5)
                }

                alpha2 ~ dnorm(0.0, 1)
                alpha ~ dflat()

                for ( k in 1:sumNumNeigh)
                {
                    weights[k] <- 1
                }
            }
```

10.8 Distance Model C-Int - OpenBUGS code

```
model    {
                u[1:m] ~ car.normal( adj[], weights[], num[], tau.u)

                for ( i in 1:m )
                {
                    # Poisson likelihood for observed counts
                    y[i] ~ dpois( mu[i])
                    f[i] <- ( 1 + exp( -alpha2*dist[i] ) )
                    # log(f[i]) is an additive-link distance effect. If the estimate of alpha2 is positive
                    # then there is a decline with distance, which might be interpereted as significant
                    # if the alpha2 is well estimated.
                    log( mu[i] ) <- log(e[i]) + alpha + beta[dep[i]] + log(f[i]) + beta2[dep[i]]*dist[i] + v[i] + u[i]
                    RR[i] <- mu[i]/e[i]
                    # Prior distribution for the uncorrelated heterogenity
                    v[i] ~ dnorm(0.0, tau.v)
                }

                eps <- 1.0E-6

                tau.u ~ dgamma( 0.5, 0.0005)
                tau.v ~ dgamma( 0.5, 0.0005)


                # Prior on beta coefficients
                beta[1] <- 0
                for (j in 2:10)
                {
                    beta[j] ~ dnorm(0.0, 1.0E-5)
                }

                beta2[1] <- 0
                for (j in 2:10)
                {
                    beta2[j] ~ dnorm(0.0, 1.0E-5)
                }

                alpha2 ~ dnorm(0.0, 1)
                alpha ~ dflat()

                for ( k in 1:sumNumNeigh)
                {
                    weights[k] <- 1
                }
        }
```

Bibliography

Arab, A., Hooten, M. B. & Wikle, C. K. (2007), *Encyclopedia of Geographical Information Science*, Springer, chapter Hierarchical Spatial Models.

Bailey, N., Flint, J., Goodlad, R., Shucksmith, M., Fitzpatrick, S. & Pryce, G. (2003), Measuring deprivation in scotland: Developing a long-term strategy, Technical report, Scottish Centre for Research on Social Justice, Universities of Glasgow and Aberdeen.

Bernardo, J. & Smith, A. F. M. (1994), *Bayesian Theory*, John Wiley and Sons Inc.

Besag, J., York, J. & Mollié, A. (1991), 'Bayesian image restoration with two applications in spatial statistics', *Annals of the Institute of Statistical Mathematics* **43**, 1–59.

Best, N., Richardson, S. & Thomson, A. (2005), 'A comparison of bayesian spatial models for disease mapping', *Stat Methods in Medical Research* **14**, 35–39.

Brooks, S. & Draper, D. (1999), Comparing the efficiency of MCMC samplers, Technical report, Department of Mathematical Sciences, University of Bath.

Carlin, B. P. & Louis, T. A. (2009), *Bayesian Methods for Data Analysis*, third edn, Chapman and Hall/CRC.

Clayton, D. & Kaldor, J. (1987), 'Empirical bayes estimates of age-standardized relative risks for use in disease mapping.', *Biometrics* **43**(3).

Congdon, P. (2003), *Applied Bayesian Modelling*, John Wiley and Sons Ltd.

Congdon, P. D. (2010), *Applied Bayesian Hierarchical Methods*, Chapman and Hall/CRC.

Cowles, M. K. & Carlin, B. P. (1996), 'Markov chain monte carlo convergence diagnostics: A comparative review', *Journal of the American Statistical Association* **91**(434), 883–904.

Cressie, N. A. C. (1993), *Statistics for Spatial Data (revised edition)*, Wiley.

Davies, C. A. (2005), Spatial Multilevel Modelling of Cancer Mortality in Europe, PhD thesis, Department of Statistics, University of Glasgow.

Diggle, P. j., Tawn, J. & Moyeed, R. (1998), 'Model-based geostatistics', *Journal of the Royal Statistical Society* **47**, 299–350.

Donnelley, R. (2008), 'Changing Scotland's relationship with alcohol: a discussion paper on our strategic approach', Technical report, Scottish Government.

Emslie, C. & Mitchell, R. (2009), 'Are there gender differences in the geography of alcohol-related mortality in scotland? an ecological study', *BMC Public Health* **9**, 58.

Flowerdew, R., Graham, E. & Feng, Z. (2004), The production of an updated set of data zones to incorporate 2001 census geography and data, Technical report, School of Geography and Geosciences, University of St Andrews.

Gelman, A., Carlin, J. B., Stern, H. S. & Rubin, D. B. (2004*a*), *Bayesian Data Analysis*, second edn, Chapman and Hall/CRC.

Gelman, A., Carlin, J. B., Stern, H. S. & Rubin, D. B. (2004*b*), *Bayesian Data Analysis*, Chapman and Hall / CRC, chapter Model checking and improvement, pp. 157–192.

Gilks, W., Richardson, S. & Spiegelhalter, D. (1996), *Markov Chain Monte Carlo in Practice*, Chapman and Hall.

Glaister, J. (1886), 'The epidemic history of Glasgow during the century 1783-1883', Philosophical Society of Glasgow.

Henderson, R., Shimakura, S. & Gorst, D. (2002), 'Modeling spatial variation in leukimia survival data', *Journal of the American Statistical Association* (97), 965 to 972.

Jackson, M. (1999), *Malt Whisky Companion*, fourth edn, Dorling Kindersley Limited, 80 Strand, London,.

Lawson, A. B. (2009*a*), *Bayesian Disease Mapping: Hierarchical modelling in spatial epidemiology*, Chapman and Hall / CRC.

Lawson, A. B. (2009*b*), *Bayesian Disease Mapping: Hierarchical modelling in spatial epidemiology*, Chapman and Hall / CRC.

Lawson, A. B., Browne, W. J. & Rodeiro, C. L. V. (2003*a*), *Disease Mapping with WinBUGS and MLwiN*, 1st edn, John Wiley and Sons Ltd.

Lawson, A. B., Browne, W. J. & Rodeiro, C. L. V. (2003*b*), *Disease Mapping with WinBUGS and MLwiN*, John Wiley and Sons Ltd., chapter Disease Mapping Basics, pp. 1–15.

Lawson, A., Biggeri, A., Boehning, D., Lesaffre, E., Viel, J. & Clark, A. (2000), 'Disease mapping models: an emprical evaluation', *Satistics in Medicine* **19**.

Lawson, A., Biggeri, A., Böhning, D., Lesaffre, E., Viel, J.-F. & Bertollini, R. (1999), *Disease Mapping and Risk Assessment for Public Health*, John Wiley and Sons Ltd.

Leroux, B. G. (2000), 'Modelling spatial disease rates using maximum likelihood', *Statistics in Medicine* **19**(17 to 18), 2321 to 2332.

Leyland, A. H. & Davies, C. A. (2005), 'Empirical Bayes methods for Disease Mapping', *Statistical Methods in Medical Research* **14**, 17–34.

Leyland, A. H., Dundas, R., McLoone, P. & Boddy, F. A. (2007), *Inequalities in mortality in Scotland*, number 16 *in* 'Occasional Paper Series', MRC Social and Public Health Sciences Unit.

Marshall, R. J. (1991), 'A review of methods for the statistical analysis of spatial patterns of disease', *Journal of the Royal Statistical Society. Series A (Statistics in Society)* **154**(3), 421–441.

McLoone, P. (2003), 'Increasing mortality among adults in scotland 1981 to 1999', *European Journal of Public Health* **13**, 230–234.

Mollié, A. (1999), *Disease Mapping and Risk Assessment for Public Health*, John Wiley and Sons Ltd., chapter Bayesian and Empirical Bayes Approaches to Disease Mapping, pp. 15 – 29.

Raftery, A. & Lewis, S. (1992), *Bayesian Statistics 4*, Oxford University Press, chapter How many iterations in the Gibbs Sampler?, pp. 763–773.

Research Unit in Health, Behaviour & Change (2007), Explaining Scotland's high mortality, Technical report, School of Clinical Sciences and Community Health, University of Edinburgh.

Rodeiro, C. L. V. & Lawson, A. B. (2005), 'An evaluation of the edge effects in disease map modelling', *Computational Statistics and Data Analysis* **49**(1), 45–62.

Spiegelhalter, D. (2006), 'Some dic slides', MRC Biostatistics Unit, Cambridge.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P. & van der Linde, A. (2002), 'Bayesian deviance, the effective number of parameters and the comparison of arbitarily complex models.', *Journal of the Royal Statistics Society* **B**(64), 583–640.

Tsai, S. P. & Wen, C. P. (1986), 'A review of methodological issues of the standardized mortality ratio (smr) in occupational cohort studies', *International Journal of Epidemiology* **15**(1).

Yamada, I. (2009), *International Encyclopedia of Human Geography*, Vol. ISBN - 978-0-08-044910-4, Elsevier Ltd, University of Utah, Salt Lake City, UT, USA, chapter Edge Effects, pp. 381–388.