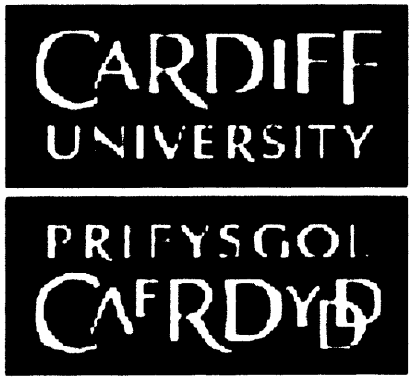
The development of methodology for identifying spatial links between environmental exposure and disease prevalence



Jessica Sian Read

PhD 2007

School of Medicine, Cardiff University

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Declaration

This work has not previously been accepted in substance for any degree and is not concurrently submitted in candidature for any degree.

Signed ..... (candidate) Date 12/10/07

Statement 1

This thesis is being submitted in partial fulfilment of the requirements for the degree of PhD.

Signed ..... (candidate) Date 12/10/07

Statement 2

This thesis is the result of my own independent work/investigation, except where otherwise stated. Other sources are acknowledged by explicit references.

Signed ..... (candidate) Date 12/10/07

Statement 3

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed ..... (candidate) Date 12/10/07

Summary

The recent increased availability of geographically linked individual level health outcome data and improvements in exposure mapping techniques, which furnish point exposure estimates, motivate the development of spatial statistical methodology that takes full advantage of individual level data. Kernel density estimation is a powerful tool for mapping the risk of a health outcome that uses individual level data. Development of kernel density methodology has provided a global significance test for regions of elevated relative risk and a test for the spatial association between a health outcome and environmental exposure. Comparisons with some existing spatial statistical techniques highlight the strengths of the kernel density based methods. Moreover, simulation exercises indicate that the kernel density test for spatial association is a more powerful testing procedure than the most popular standard test proposed by Stone. Kernel density estimation and the global significance test for regions of elevated relative risk are illustrated for congenital malformations around a landfill site and sex ratios in Cardiff and the Vale of Glamorgan. The application of these methodologies revealed that both birth outcomes had a statistically significant heterogeneous spatial pattern over the relevant study regions even after adjustment for known confounders. Good quality, high resolution environmental exposure data was unavailable and prevented a direct application of the kernel density test for spatial association with a health outcome. However, the test can be applied to any two relative risk/density surfaces and was used to compare the spatial patterns of chromosomal and non-chromosomal anomalies in the region of the Nant y Gwyddon landfill site. It was concluded that the spatial patterns for the two sets of anomalies were different. The test was also used to assess the quality of the adjustment for confounders when producing expected risk surfaces and the adjustment was found to be adequate.

Acknowledgements

Many thanks go to my supervisors without whom the completion of this thesis would not have been possible. I am most grateful to Barry Nix for passing on some of his statistical knowledge and answering my continual stream of questions. Thanks to Ian Matthews for his valuable advice and support. Thank you to Louise James and John Evans for preparing datasets for my use. Finally, gratitude goes to my family and friends for their support throughout the last 3 years. I would like to acknowledge the financial support provided by the MRC.

Aims and Objectives

The main aims of this thesis relate to developing new methodology for the spatial analysis of health outcomes that takes advantage of increased availability of better quality, high resolution exposure data and data for health outcomes based on address point. These aims are as follows:

- to derive a hypothesis testing procedure for detecting the spatial association between the relative risk surface for a given health outcome and mapped exposure to a specified risk factor;
- to establish a hypothesis test for assessing the global statistical significance of regions over which the relative risk surface for a given health outcome, determined by kernel density techniques, is elevated.

The applications that highlight the uses of the new and existing methodologies discussed in this thesis include:

- exploring the relative risk of congenital malformations in proximity to the Nant y Gwyddon landfill site;
- comparing the spatial distributions of chromosomal and non-chromosomal anomalies around the Nant y Gwyddon landfill site with the view to illustrating how the methodologies discussed can be used to investigate two health outcomes;
- investigating the temporal and spatial patterns in the sex ratio for Cardiff and the Vale of Glamorgan.

Contents

List of Appendices

Chapter One: Introduction

Investigations into the spatial relationship between exposure to environmental pollutants and specified health outcomes are a hot topic of discussion in the literature. Spatial statistical analysis of this nature will only become increasingly popular in the future with visual appeal, increased availability of geographical techniques and the desire for a cleaner environment [1]. In the absence of individual level data, a commonly used approach involves mapping the risk of the health outcome for pre-specified administrative boundaries and subsequently testing for association between this and the distance from a putative point source of pollution, where distance provides a proxy for exposure to environmental toxicants. However, assessment of exposure is improving with geographic information systems (GIS) [2] and the efforts of the National Environmental Public Health Tracking Network (EPHTN) [3], the Small Area Health Statistics Unit (SAHSU) and the European Health and Environment Information System (EUROHEIS) [4] in developing frameworks to assess the impact of environmental agents on health outcomes. This better quality, higher resolution exposure data coupled with the increased availability of point data for the health outcome based on address point [1] motivates the need for more advanced methodology that can take full advantage of this individual level data. The development of this more advanced methodology for the spatial analysis of a health outcome and how it relates to exposure to environmental contaminants is the focus of the work presented in this thesis.

Landfill sites have caused much concern for local residents in terms of adverse health impacts. A landfill site is a location on land where waste is disposed. Landfills vary in size, age and the level to which they are controlled with some still in operation whilst others are not. The waste contained in each landfill site is different with the contents of some landfill sites being more hazardous than others. Additionally, site variations can lead to differences between the amount and composition of chemicals emitted from the site, the way in which they migrate from the site and the extent to which the local population can come into contact with the contaminants. All these factors can make investigations into the health impact of residence near a landfill site complex with multiple factors interacting. A detailed description of waste types, putative exposure routes and management practices can be found in a thesis by Vrijheid [5].

Infamously, the Love Canal was drained for use as a landfill site for chemical by-products which were sealed in metal drums [6]. Materials seeped out of the drums into basements of nearby homes and a school and ran into the storm sewers which fed into Nigara River. Anecdotal evidence suggested that there may be a higher rate of congenital malformations amongst exposed families. However, overall it was judged that there was no clear evidence of an increased risk of congenital malformations in women living next to the canal. Despite this lack of evidence . the Love Canal incident led to the risk of congenital malformations in the vicinity of landfill sites becoming the focus of numerous studies [7].

Congenital malformations are physiological or structural abnormalities that originate before or during birth and can also be referred to as congenital anomalies, congenital abnormalities and birth defects [8]. In his extensive discussion of congenital malformations, Kalter iterates that even today the causes of congenital abnormalities are still largely unknown [6]. In fact, the cause of far less than half of congenital anomalies has been determined. Congenital anomalies cause 1 in 5 of the deaths that occur in the first year of infancy and 1 in 3 of the deaths in the first month. If further understanding of birth defects was obtained then it may be possible to prevent their development and consequently avoid these deaths. The lack of knowledge combined with the adverse nature of congenital anomalies motivates countless studies of birth defects including that presented in this thesis.

The preoccupation with landfill sites and congenital malformations triggered by the Love Canal incident motivated the use of the risk of congenital malformations in proximity to a landfill site in the development of the new more advanced methodology presented in this thesis. The spatial relationship between congenital malformations and exposure to landfill sites has long been debated yet the methodology that has previously been used to investigate the link is not without flaws. The use of this existing statistical methodology may have led to incorrect conclusions concerning the link between the risk of congenital malformations and exposure to emissions from landfill sites. Therefore, the proposed more advanced spatial statistical methodology is required to increase our understanding of the impact of landfill emissions on the risk of having a congenitally malformed birth.

The landfill site selected for use in the development of the methodology presented in this thesis was Nant y Gwyddon, a Welsh landfill site in the Rhondda valley. Fielder et al. performed a study on the Nant y Gwyddon landfill site from 1983 to 1997 [9]. The

investigation revealed that the incidence of all congenital anomalies was raised in the exposed wards both before and after the opening of the site. This suggests that the opening of the Nant y Gwyddon landfill site did not influence the risk of congenital malformations. Subsequently, a paper by James et al. identified regions of elevated risk of congenital anomalies in the area surrounding Nant y Gwyddon between 1983 and 1997. This discovery motivated further investigation into whether these elevations in the risk of congenital malformations is a direct result of exposure to emissions from the landfill site [10].

Before embarking on any investigation into the aetiology of a given health outcome it is important to understand aspects of the health outcome that are already known. Congenital anomalies develop during the development of an unborn child. Consequently, in precedence to discussing in detail known aspects of congenital malformations the development of an unborn child is described from the moment that the gametes are created up to the point that the foetus is ready for birth.

1.1 Embryonic and Foetal Development

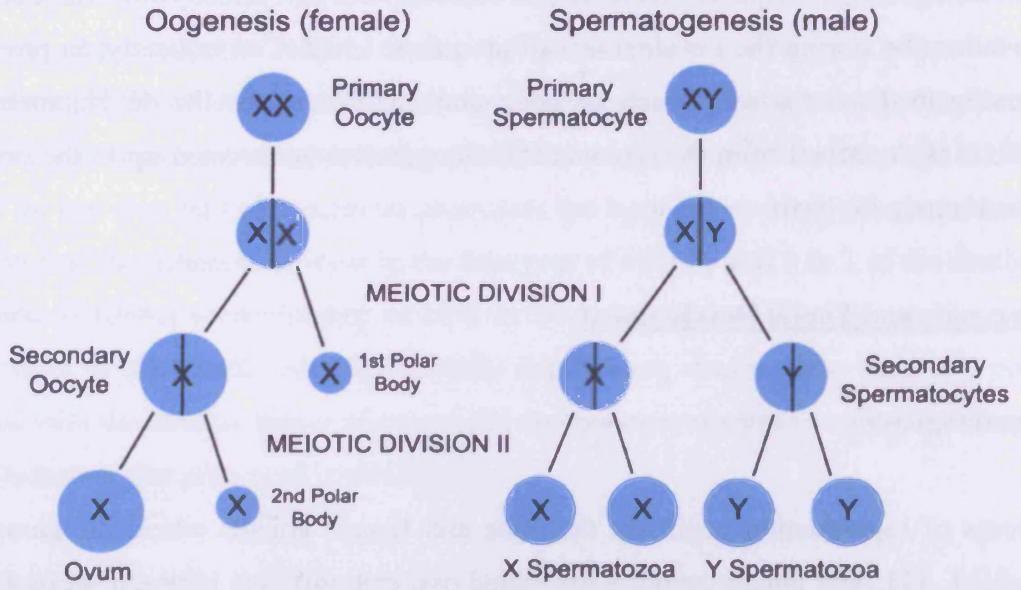
1.1.1 Gametogenesis

The process of reproduction begins in the male and female gonads where the gametes are produced [11, 12]. The female gametes are called ova and only one is produced each month compared to the millions of male gametes, or spermatozoa, which are generated in the same time frame. These gametes are the product of gametogenesis (see Figure 1.1) which begins with stem cells, with a total of 46 chromosomes, in the ovary or testes. The chromosomes comprise of DNA which contains the hereditary information required to create new life and is passed on from cell to cell [13]. The stem cells divide by mitosis to produce two diploid daughter cells.

46 is the correct number of chromosomes for a human being and for a female two of these will be X chromosomes, whereas a male has an X and a Y chromosome [11, 12]. Therefore, the offspring only requires 23 chromosomes from each parent, which is achieved when the daughter cells undergo meiosis producing haploid secondary oocytes (female) or spermatocytes (male), each with 22 autosomal chromosomes and one sex chromosome.

Spermatogenesis occurs in the testes from puberty producing spermatozoa daily until about 60 years of age. The spermatozoa shrink and develop a tail during spermatogenesis to enable motility. In contrast oogenesis occurs in two separate stages with the first meiotic division occurring at the birth of the mother in her ovaries. The second stage, however, does not occur until the ovum is fertilised in the fallopian tube. A further disparity between gametogenesis in males and females is the final product, with four spermatozoa generated from spermatogenesis, whereas oogenesis only produces one ovum, the remaining haploid daughter cells are polar bodies which either degenerate or do not survive.

Figure 1.1: the process of gametogenesis in males and females



Secondary oocytes can only contain an X chromosome whereas the secondary spermatocytes contain either an X or a Y chromosome. In order to restore the full 46 chromosomes required to begin a new life one ovum and one spermatozoon must fuse together in a process called fertilisation.

1.1.2 Fertilisation

Fertilisation is the primary stage in the development of an unborn child. A woman ovulates once in each menstrual cycle, this is the process where the mature ovum leaves the ovary

where it was produced [11-15]. The ovum is protected by a jelly-like coating as it is ovulated. The ovum moves down the fallopian tube by a ciliary action and peristalsis.

The spermatozoa are ejaculated into the woman's vagina during intercourse. After ejaculation the spermatozoa travel through the uterus and up the fallopian tube by contractions. Only one spermatozoon penetrates through the ovum's protective membrane and further spermatozoa are prevented from entering by the hardening of the ovum's outer surface. Once inside the ovum the spermatozoon undergoes changes where the tail is separated from the head and absorbed whilst the head forms the male pronucleus. The ovum also generates the female pronucleus by undergoing its second meiotic division. The nuclei are drawn together and fuse to create a single cell, called the zygote. The zygote contains two copies of each chromosome, one from the female and the other from the male, which align with each other.

Later in this chapter a second birth outcome, the sex ratio, is considered. The baby's gender is decided from this point as are aspects of their appearance, for example eye, hair and skin colour. The gender of the baby is dependant on the spermatozoon that fertilised the ovum, if it is a Y-bearing spermatozoon then the offspring is male whereas an X-bearing spermatozoon produces a female child.

1.1.3 Pre-Implantation

The zygote now undergoes rapid cell division or cell cleavage by mitosis, which involves the duplication of the 46 chromosomes and the subsequent separation of the cell to produce two diploid daughter cells [11-16]. These daughter cells should adhere to each other, when they do not adhere twinning occurs. Meanwhile the zygote moves Down's the fallopian tube and after four days the conceptus reaches the uterus and is now referred to as the blastocyst. The blastocyst is free in the uterine or endometrial cavity for 4 or 5 days whilst it prepares for implantation [17].

1.1.4 Implantation

Implantation is the process where the blastocyst embeds itself into the uterine wall or endometrium [11-16]. After contact with the endometrium the conceptus erodes through the lining and sinks into the endometrium's deeper layers taking two days to burrow through.

Implantation is complete 14 days after fertilization with the conceptus completely submerged in the upper part of the uterine wall. A fibrin plug marks the site of the implantation and is the only part that is exposed. Once it is embedded it will remain here for the 9 months of its development gaining the nutrients and oxygen it requires.

1.1.5 Embryonic Period

The embryonic period occurs after implantation and continues until 8 weeks after fertilization, during this time the unborn child is referred to as the embryo. There is a wide range of literature on the development of unborn children and they do contradict each other in many ways which may in part be a result of beliefs stemming from pro-life or pro-choice views [11, 12, 14-16, 18-22]. The cells in the embryo continue to divide and each newly developed cell now has a specific role for the development of muscular, neural, reproductive, skeletal, digestive and circulatory functions.

Initially the embryonic disk develops as the inner cell mass flattens. The embryonic disk differentiates into three layers each with different purposes. The ectoderm generates the nervous system, skin and its appendages and the sensory organs. Bones, muscles, the vascular system and urogenital system is a product of the mesoderm. Finally the endoderm is the seed for the gastrointestinal tract and derivative organs, thyroid and parathyroid glands, the thymids and lungs. In the first week of the embryonic period, 3 weeks after fertilization, the backbone, spinal column and central nervous system begin to develop. The faster rate of growth in the embryo's back results in the embryo forming a C-shape.

In the 4^{th} week organogenesis begins with the development of the heart. The cardiac activity begins as early as 22 days after fertilization. The central nervous system has developed into the primary brain vesicles, that is, the forebrain, midbrain and hindbrain and the spinal column fuses. The respiratory, renal and digestive primitive elements form at this time culminating at the end of the week with pancreatic, renal, ureteric and lung buds. The organ systems at this stage can be identified despite being immature.

The 5^{th} week is the time when cells form which will ultimately provide muscles, bones and nerves and when the arm and leg buds appear. The legs and arms with toes and fingers are fully formed by week 6. The eyes, ears and nasal organs start to form at this point and the

circulatory system, thyroid gland, liver, pancreas and gall bladder are present. At this time brain waves are detectable and the mouth, lips, fingernails and skeletal system materialise. The ossification or hardening of the bones begins.

In week 7 the lungs separate from the gullet. The baby begins to move which helps develop the muscles and bones. Eyelids, toes and nose begin to grow in the 7th week. Additionally, the liver takes over the production of blood cells during this time. In the 8th week bones start to grow and replace the cartilage as arms and legs lengthen. Also the external genitalia and external ears develop. In this final week of the embryonic period the kidneys develop, the urinary and rectal passages separate, the anal membrane ruptures, facial clefts close and endocrine glands begin to function.

All the organs are in place but may not be mature or functioning yet and the baby grows to half an inch in size by the conclusion of the embryonic period. The embryo is no longer bent over as the growth of the front has caught up with that of the back. The head and facial features have grown dramatically.

1.1.6 Foetal Period

The foetal period comprises the remainder of the first trimester and the second and third trimesters. This phase is one of rapid growth and development as the foetus prepares for survival outside the uterus. The mother transfers immunity to the child and there is rapid brain, movement and organ development over this time to give the child independence after birth. Ossification of the bones occurs where the cartilage that initially made up the skeletal system hardens. At this stage the nervous system begins to function at a primitive level, the teeth develop beneath the gums and the bone marrow takes over blood production from the kidneys. The sex of the child can be determined in the 4th month. Movements of the baby develop pathways to the brain. The skin grows with sweat glands and hair. Fat fills out on the body so that the limbs become smooth and plump. The eyes fully open and hair grows on the head. Ultimately by the end of this period the unborn child should be fully prepared for birth.

If a problem occurs during the development of an unborn child then it can lead to a congenital malformation. Therefore, exposure to important risk factors is meaningful between gametogenesis and birth particularly during the first trimester when most of the organogenesis

occurs [23]. Congenital anomalies are discussed in detail in the next subsection and known aspects of the aetiology of birth defects are described.

1.2 Congenital Malformations

There are two types of congenital anomalies: chromosomal and non-chromosomal abnormalities. Chromosomal anomalies are the result of an error in chromosomes which contain DNA, the information required to produce a human being. Non-chromosomal anomalies occur when there is an error in one or more of the stages in the development and growth of the unborn child not relating to chromosomes.

1.2.1 Chromosomal Anomalies

Chromosomal anomalies are abnormalities relating to chromosomes and they are a result of errors that occur during meiosis [11, 24]. There are two main errors that lead to chromosomal anomalies. The first is where one pair of chromosomes does not separate in time leaving one daughter cell with two copies and the other with no copies of the chromosome. The second is where a chromosome lags and gets omitted from both the daughter cells. Both these errors lead to unequal separation of the chromosomes resulting in spermatozoa or ova with fewer or more chromosomes than normal. The abnormality is passed on to the offspring during fertilisation. Too few or too many chromosomes result in either abnormal or failure of foetal development. 40-50% of spontaneous abortions are caused by chromosomal anomalies [11]. Chromosomal anomalies are typically related to the maternal age of the mother with older women at greater risk of having an affected child. This phenomenon is thought to relate to the ability to miscarry an affected foetus which is reduced with increased age [11].

Autosomal abnormalities

Foetuses with too few autosomes cannot develop as too much information is absent. However it is possible for a foetus to have too many chromosomes and survive. Trisomies originate during meiosis, when three copies of a given chromosome are produced as opposed to two [17, 25-29]. The majority of trisomies lead to miscarriage except for Patau syndrome (trisomy 13), Edwards syndrome (trisomy 18) and Down's syndrome (trisomy 21).

Patau and Edwards syndromes are associated with serious structural defects that are easily identified before birth and it is uncommon for an affected foetus to survive. Any affected children that are born live no longer than two years. Edwards syndrome occurs in 1 in 3000 births. Typical symptoms for trisomy 18 include hypoplastic lungs, flexion deformities, rocker-bottom feet, clenched hands, craniofacial abnormalities and other congenital malformations. On the other hand Patau syndrome is far less common. Symptoms that can occur with Patau syndrome include defects of the face, eyes and forebrain, cleft lip and palate, severe mental retardation, deafness, rocker-bottom feet, super-numerous digits, congenital heart defects and cryptorchidism.

Down's syndrome is less severe and people affected by it are more likely to survive. Therefore antenatal screening places greater emphasis on detecting this trisomy. 1 in 600 live births are affected by trisomy 21. Down's syndrome patients often have decreased muscle tone, shortened limbs, single palmar creases, digit abnormalities, crowded facial features, upward slanting eyes, Brushfield's spots on the iris, a short head which has a flat back with loose folds of skin behind the neck, small ears, mental retardation and other congenital anomalies such as cardiac abnormalities. The mortality rate for people with Down's syndrome is normal until the age of 40, when an increased rate of aging leads to increased mortality for Down's syndrome sufferers.

Sex chromosome abnormalities

Sex chromosome abnormalities occur when there is an abnormal number of sex chromosomes [11, 25, 27, 28]. There is no routine screening for sex chromosome abnormalities and in some cases affected individuals can go throughout their lives undiagnosed.

Turner syndrome (monosomy X or 45,XO) occurs when the offspring has only one sex chromosome (an X chromosome) and children affected by it are infertile females of normal intellect characterised by their small stature. Affected patients can have renal or cardiac abnormalities or hearing impairment. Kleinefelter syndrome (47,XXY) is a result of the offspring having three sex chromosomes (two X chromosomes and one Y chromosome). People with Kleinefelter syndrome are infertile males with a reduced IQ, testicular abnormalities and tall stature.

Other possible sex chromosome abnormalities include the super female (47,XXX), the tall male (47,XYY) and the fragile X syndrome. However, there is too much genetic information missing for the development of a human foetus if there is no X chromosome so it is not feasible to have only a Y chromosome (45,YO).

Other chromosomal anomalies

Other errors that can occur during meiosis include translocation where part of one chromosome is added to another and isochromosome formation where the separation of the chromosome pair is in the transverse axis rather than the longitudinal [11]. In some cases short segments of the chromosome can be deleted. The results of the loss or addition of smaller parts of the chromosome are less clinically obvious than with too few or too many complete chromosomes.

Triploidy is where there is an extra haploid set of chromosomes, i.e. there are 69 chromosomes in total [27]. This chromosomal anomaly occurs as a result of either the fertilisation of the ovum by two spermatozoa or the fertilisation of a diploid egg and is fatal.

Abnormalities of chromosomes can also occur after conception in a process called mosaicism [11]. In mosaicism some cells have chromosomal defects whilst others have none [29]. The earlier mosaicism occurs in the development of the unborn child, the greater the mutation.

1.2.2 Non-chromosomal Anomalies

Non-chromosomal abnormalities are the most common form of congenital malformation and result from the unsuccessful development of part of the foetus. There are many different types of non-chromosomal abnormalities varying in location and severity.

Neural tube defects

Neural tube defects (NTDs) occur when there is an error in the formation of the neural tube [11, 17, 24, 26-28, 30-33]. Major central nervous system (CNS) defects occur between 3-6 weeks after fertilisation and minor defects between weeks 6-38. Little is known about why NTDs occur, although, taking folic acid supplements reduces the risk. Some NTDs cannot

support life, for example anencephaly, where most of the brain and spinal cord are absent and the bones of the cranial vault are not present leaving any existing part of the brain exposed. Other NTDs can have a better prognosis if the brain and spinal cord are present but have not correctly developed, for example, if the skin and other tissues fail to close over the nervous tissue. Encephalocoele occurs when there is a gap in the skull through which the brain protrudes. The prognosis for births affected by encephalocoele depends on the size of the hole. Spina bifida is where the spinal column is imperfectly closed as the vertebral arches do not fuse and the posteria dura mater fails to develop correctly thus the spinal column protrudes through the bony gap. The consequences of spina bifida include paralysis of the legs, incontinence and impaired intellect.

Congenital heart defects

Congenial heart disease is the most common single system malformation and originates in weeks 3 to 8 when the heart tube separates to form the two atria and two ventricles and these are subsequently connected [11, 28]. Ventricular septal defect occurs when the left and right ventricles fail to separate completely [27]. Congenital heart lesions occur when the separating walls do not fuse which can lead to respiratory difficulty, cyanosis, rapidly developing cardiovascular shock and heart murmurs [24].

Defects of the digestive system

There are a number of defects of the digestive system for example it may be blocked at some point, such as the oesophagus or intestines, which can result in vomiting or abdominal distension [11, 17, 24, 26-29, 31]. More serious defects occur as a consequence of errors in the development of the digestive system. Oesophageal atresia is where a short section at the top of the oesophagus has a 'blind end' blocking food from reaching the stomach. Often oesophageal atresia is coupled with a tracheo-oesophageal fistula which is where the oesophagus is also connected to the trachea. Exomphalos occurs when the gut fails to re-enter the abdominal cavity. Here the abdominal wall is incompletely sealed resulting in the protrusion of the gut from the umbilicus in a sac. Gastroschisis is similar only there is no sac to protect the gut. Less serious defects include an imperforate anus preventing the excretion of faeces. Duodenal atresia is where the beginning portion of the small intestine is missing.

Defects of the respiratory system

A diaphragmatic hernia is where the diaphragm, which separates the abdominal and respiratory cavities, fails to fuse or muscularize [17, 26-29, 31]. The gut protrudes into the respiratory cavity, often preventing the complete development of the lungs and displacing the heart to the right. If the lungs fail to develop completely then the baby cannot survive.

Defects of the urinary system

Abnormal kidney growth can lead to Potter syndrome, the consequences of which include low set ears, compression abnormalities with flexion limbs, hypoplastic lungs and renal failure [17]. The kidneys can form multiple cysts and this is referred to as polycystic disease [26, 27]. Ectopia vesicae is where the bladder is herniated and its mucous membrane is exposed. Hypospadias is where the urethra fails to open at the tip of the penis in males. The urethra can also be obstructed leading to a distended bladder. The kidneys and the bladder can be absent in what is referred to as renal agenesis [29, 31].

Skeletal defects

There are a number of skeletal defects. Osteogenesis imperfecta occurs when the bones are poorly mineralised and are, therefore, susceptible to multiple fractures [17, 27, 28]. Short-limbed dwarfism is characterised by shortened limbs, a large head, prominent forehead and chest underdevelopment. Limb reduction deformities involve either absent limbs or shortened limbs with rudimentary hands or feet.

Other structural anomalies

A cleft lip is a linear defect extending from the lip to the nostril and can be surgically corrected [25, 26]. A cleft palate often occurs with a cleft lip when the primary and secondary palates do not fuse [27]. Club foot is where the forefoot is supinated resulting in the outer sole bearing the weight of the body and the flexing of the ankle i.e. the feet are twisted inwards and downwards [27]. When a baby has rocker-bottom feet the soles of the feet are convex. A baby can be born with extra digits which can be either fully formed or only tags of flesh, this is referred to as super-numerous digits [24].

1.2.3 Detection and Diagnosis of Congenital Malformations during Pregnancy

Many attempts are made during a mother's pregnancy to assess the wellbeing of the foetus and, in particular, to ascertain whether or not the unborn child has a congenital malformation. The diagnosis of such conditions before birth enables the parents to prepare for the birth of a child that has something wrong with it. If it is determined early enough there is also an opportunity for the parents to make an informed choice about whether they would like to continue with the pregnancy or whether they want to terminate. Termination is legal before the 24th week of gestation and for any pregnancies that exceed that the obstetrician must be certain that the child would either die in infancy or have considerable pain and suffering to allow a termination [34]. On the other side of the coin if it is shown that the foetus shows no sign of a defect then it can help to reassure the parents. There are two aspects to assessing the wellbeing of the foetus: screening and diagnostic testing.

Screening

Screening tests, along with other risk factors such as maternal age, social class or race, are used to assess the risk of a congenital malformation. All pregnant women are invited to take part in the screening program. All the way throughout the program the risks of specified congenital malformations are reassessed as more information is collected and if the risk is high the mother is referred for diagnostic testing as discussed below. A risk that is perceived to be high is determined by follow-up assessment procedures, the number of which is dictated by the available resources. The screening program offered to pregnant women varies between health authorities but a typical program as discussed by Sullivan is summarised here [34]. The screening process cannot be used to assess the risk of all congenital malformations as some internal defects cannot be seen on an ultrasound nor do they have an impact on the levels of certain biomarkers.

Ultrasound is used to search for obvious signs of congenital anomalies [11, 17, 25, 34]. The first trimester ultrasound scan is offered to mothers between the 10th and 14th week of gestation [29, 34]. This initial assessment of the foetus can detect serious structural anomalies, for example, anencephaly. A detailed foetal anomaly scan is offered to mothers at 18 to 20 weeks' gestation and informed consent is required before proceeding. In this second trimester

assessment the spine, head shape and internal structures, abdominal shape and content, pelvis, thorax, arms, legs, face, lips and cardiac outflow tracts are examined for obvious structural anomalies. In addition to the signs of structural defects, certain properties of the foetus are assessed to determine the risk of chromosomal anomalies. There are numerous different measurements and observations that can be made to achieve this, examples include the nuchal translucency (NT) and the triple test.

The NT is measured after 12 to 14 weeks' gestation [24, 25, 29, 34]. The NT measurement is the determination of the thickness of the subcutaneous fluid that collects at the back of the neck of the foetus. A high NT (greater than 3mm at 12 weeks' gestation) is associated with chromosomal anomalies, particularly Down's syndrome, as well as some structural and genetic disorders.

Alpha fetoprotein (AFP), human chorionic gonadotropin (HCG) and unconjugated oestriol (uE_3) levels in the amniotic fluid are measured as part of the triple test for Down's syndrome [11, 24, 25, 33, 34]. The levels of these three markers in a sample of maternal serum are measured at 15 to 18 weeks' gestation. If the unborn child has Down's syndrome then the HCG level is higher than expected and the AFP and uE_3 levels are lower than expected. Consequently the three measures can be used in the risk assessment for Down's syndrome.

AFP levels can also be used to test for NTDs [11, 17, 24, 33, 34]. The AFP levels vary by gestational age and if an accurate gestational age is known then the expected AFP level can be determined for the foetus at any given time. If the neural tube fails to close then AFP escapes in increased amounts and therefore if the AFP is measured and found to be higher than expected then it is likely that there is a NTD.

Invasive diagnostic testing

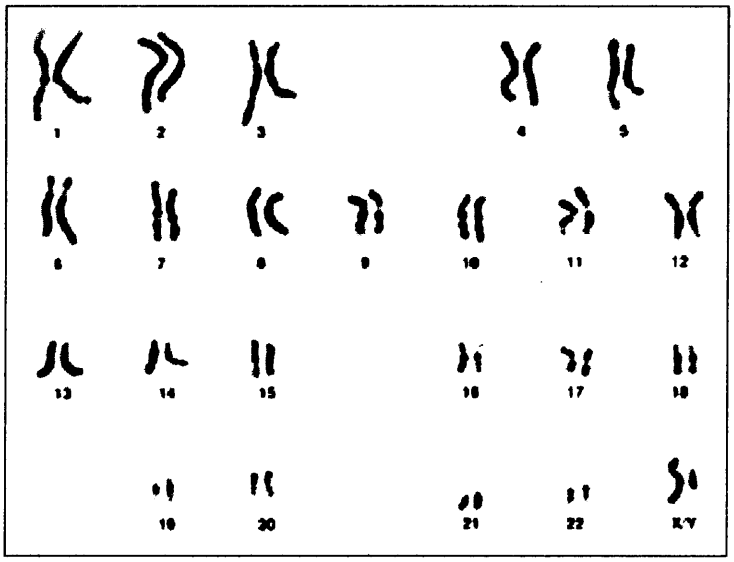
The mothers that are asked to return after screening undergo diagnostic testing which is used to confirm or rule out the presence of a given congenital malformation [34]. However, there are limitations to diagnostic testing. The severity of the condition cannot be ascertained, for example, it may be possible to diagnose Down's syndrome but not to determine the level of mental disability that the child will have so they could be severely mentally retarded or only mildly affected intellectually. This uncertainty makes decisions regarding termination very

difficult. In some cases a false positive result for congenital defects can be obtained, particularly from ultrasound [11]. Additionally, not all defects can be detected during pregnancy and some abnormalities do not manifest until as late as adulthood.

If there is a concern that a structural defect may be present then third trimester scans are performed to reassess this [34]. If chromosomal abnormalities are suspected then invasive diagnostic testing is offered to the mother. Invasive diagnostic testing carries with it an approximate 1% risk of spontaneous abortion and can be traumatic for parents. If the parents choose to undergo invasive diagnostic testing then this provides the means to determine the foetal karyotype or perform DNA analysis for gene mutations.

A karyotype is an arrangement of all the chromosomes in a cell in a standardised way revealing the characterisation of the chromosomal complement of that cell indicating the size, number and form of these chromosomes [8, 25, 35]. An example of the karyotype for a human male is given in Figure 1.2, in which, the chromosomes are arranged according to the standard configuration of ordering the pairs of chromosomes by size. The normal human karyotype is 46,XX for females and 46,XY for males, deviations from this either result in a congenital defect or are fatal, as are structural changes that can occur in the chromosomes. Therefore, the karyotype can be used to diagnose a chromosomal anomaly through the arrangement of the chromosomes.

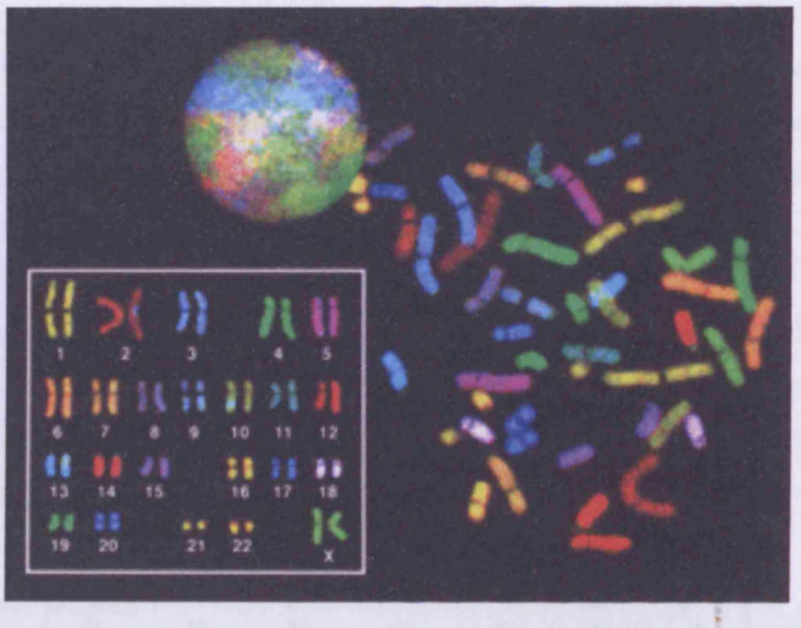
Figure 1.2: the karyotype for the human male



The karyotype of a foetus is obtained by initially taking a sample of amniotic fluid, placental tissue, foetal blood or foetal bone marrow [8, 35]. The cells in the sample are cultured until sufficient cells have been generated through mitosis to make a diagnosis. The cells are then harvested and stained with a dye that enables the bands to be visible under a microscope. The banding patterns on the chromosomes are different for each chromosome pair, enabling them to be identified. The difficulty with this process is that an abnormal result may have originated after the cells were abstracted, therefore, the results should be double checked.

In spectral karyotyping different probes specific to one chromosome pair carry different quantities of a set of fluorescent dyes that are hybridized to the chromosome [8]. The relative amount of each of the resultant fluorochromes gives each chromosome pair distinct spectral characteristics. The chromosomes can now be identified in fluorescence microscopy and spectral imaging. An example of a spectral karyotype is given in Figure 1.3.

Figure 1.3: a spectral karyotype



There are three types of invasive diagnostic testing: chorionic villus sampling (CVS), amniocentesis and cardiocentesis. Research into the use of maternal blood samples is currently underway in the hope that these techniques no longer need to be considered [34]. However, cells from previous, and sometimes unknown, pregnancies can remain in the mother's blood, making it difficult to determine which are required for analysis.

CVS is the acquisition of placental tissue using ultrasound guidance either trans-abdominally or trans-cervically [11, 24, 25, 34]. The advantage to CVS is that it can be performed as early as 10 weeks after fertilisation enabling an easier termination if requested. A chromosome count can be achieved in 24 to 48 hours and then cells are cultured for 14 to 21 days to obtain more detailed information on the structure of the chromosomes. In addition, genetic and metabolic diagnoses can be made from CVS. However, CVS carries a higher risk of miscarriage, as a result of infection or bleeding, at between 0.5 and 2% depending on the clinician performing the procedure.

Amniocentesis involves the trans-abdominal insertion of a fine needle into the amniotic cavity under ultrasound guidance to abstract amniotic fluid [11, 24, 25, 29, 34]. This procedure cannot be performed until after 15 weeks of gestation. However the risk of miscarriage, caused by infection or spontaneous rupture of the membranes at the location of insertion, is lower at 1%. Cytogenetic and metabolic analyses are possible alongside the analysis of the amniocytes, which comprise cells from multiple foetal sites including the lungs, skin and renal tract. An initial set of results is available especially for Down's syndrome and then a full culture is complete within 2 or 3 weeks.

Cardiocentesis is rarely performed as it requires considerable skill to perform it compared to the other two invasive diagnostic testing procedures [11, 24, 25, 34]. In this procedure, blood is sampled from the umbilical cord or intrahepatic umbilical vein and then analysed. Cardiocentesis enables rapid determination of the karyotype along with the diagnosis of immune deficiencies and blood and platelet disorders. The spontaneous abortion rate is 1% after 20 weeks.

1.2.4 Detection and Diagnosis of Congenital Malformations after Birth

Once the baby is born, an inspection is made to check for any congenital anomalies that may have been missed in the screening process [24]. The face is studied for signs of trisomies and for a cleft lip or palate. The abdominal region is checked for herniation as seen in gastroschisis. Additionally any distension of the abdomen or vomiting may suggest an internal obstruction in part of the digestive or urinary systems. Cyanosis or respiratory difficulties could indicate a herniated diaphragm. Examination of the external genitalia may reveal abnormalities and lead to difficulties in determining the gender of the child. The anus is

checked to ensure that it is perforated and that the opening is in the correct place. The vertebrae are scrutinised for possible signs of a NTD. The ears are checked for presence and multiple lobes. The hands and feet are checked for webbing, super-numerous digits and other defects. The hip is studied to ensure that it is not dislocated.

1.2.5 Aetiology of Congenital Malformations

The main concern for epidemiologists studying congenital malformations is to establish the relevant risk factors. This continuing endeavour to determine the aetiology of congenital malformations has provided the motivation for the research presented in chapters 2 to 4. The aetiologies of the different types of congenital malformations are largely unknown and vary between each different defect. Very few defects have simple explanations and a range of risk factors is common. The multifactorial aetiology exhibited for the majority of defects makes the identification of individual risk factors difficult.

There are two types of risk factors: mutagenic risk factors, affecting the DNA, and teratogenic risk factors, affecting the embryo or foetus during development. Mutations are inherited and some people are genetically more susceptible to the effects of teratogens. The mechanism involved in teratogenesis or mutations is rarely known, hence, most risk factors cannot be linked to the condition with a high degree of certainty.

It is fairly common knowledge that maternal age can be a major factor for a small number of congenital malformations. Chromosomal aneuploidies, specifically Down's syndrome, are more common amongst mothers giving birth at an older age [23, 36, 37]. In contrast, young maternal age is associated with an increased risk of gastroschisis [23, 37]. In a study of the impact of paternal age, where maternal age was adjusted for, the risk of gastroschisis increased with decreasing paternal age as did the risk of Down's syndrome [38]. However, it is not clear why maternal and paternal age influence the risk of these birth defects in this way. Rates of some congenital anomalies can also vary between ethnic groups [6]. Higher social status is related to a high risk of Down's syndrome but this is likely to be a response to increased maternal age in the higher socioeconomic groups [37]. Other possible risk factors for congenital anomalies include inheritance, maternal drug use, viral infections, maternal health disorders, environmental agents and radiation, which are discussed in more detail below.

Inheritance

Genes are composed of DNA and each one is concerned with the transmission of one hereditary factor [28]. A mutation occurs when there is a change in the DNA. A mutation can easily happen as each time a cell divides DNA is replicated and errors can occur in this process [39]. Chromosomal abnormalities are a form of mutation but often the mutation only affects parts of chromosomes. In addition to natural errors, radiation and chemical agents can damage DNA. Any damage that occurs can be detected and corrected by a DNA repair system but some can escape detection and lead to problems particularly if in the germ cells. If a mutation occurs in a germ cell then that mutation can be inherited by offspring. Maternal age can often be a risk factor in genetic disorders, as mutations are more likely to occur in older mothers. This phenomenon may be a result of the longer suspension of the second phase of meiosis, resulting in a higher risk of an error.

One of the first considerations, when investigating the aetiology of a specific congenital malformation, is whether or not it is genetic. This initial step can be achieved by considering the family history of babies that were born with the given condition and using DNA testing. If the congenital malformation is a monogenic disorder, i.e. the result of a mutation in a single gene, with no environmental factors then the aetiology is more easily discernable using Mendelian principles. An example of a monogenic disorder is the Meckel syndrome which involves encephalocoele, polydactyl and polycystic kidneys [39].

Mendelian principles include autosomal dominance, autosomal recession and sex-linked inheritance and the way in which a disorder manifests under each situation can help deductions on the nature of the inheritance [11, 24, 25, 28, 32, 39]. If the defect is autosomal dominant then it manifests in heterozygotes, thus, there is a 50% chance of passing it on to the child regardless of gender. Therefore, under this situation if a child exhibits the disorder then one parent will also have it. Sickle cell anaemia is an autosomal dominant disorder, where the affected individual has sickle-shaped blood cells that lead to episodic pain in joints, fever, leg ulcers and jaundice. Autosomal recessive conditions manifest only when two copies of the mutated gene are present. In this case there may be no family history but both parents will be carriers leading to a 25% chance of offspring being affected and a 50% chance of them being a carrier. An example of an autosomal recessive disorder is cystic fibrosis, a mutation of

chromosome 7 that affects the endocrine glands, leading to chronic respiratory infections and impaired pancreatic function. In sex-linked inheritance the mutation is carried on the X chromosome, for example, haemophilia which is characterised by the impaired ability to control bleeding. Male offspring have a 50% chance of having the defect and females never have it but have a 50% chance of being a carrier and consequently a 50% chance of passing it on to their own sons. Some inherited disorders are controlled by multiple genes, like tuberous sclerosis, in which case the probabilities of having the defect cannot be calculated simply.

Unfortunately the majority of disorders do not have a simple aetiology [11, 32, 39]. Multifactorial disorders can show signs of familial clustering but not exhibit Mendelian inheritance. In this situation the genetic affect does not occur unless certain conditions are met. The conditions required are determined by a combination of genotype and environmental factors. The aetiology is difficult to determine when it is multifactorial and, for many disorders, little is known of the genetic or environmental factors that lead to manifestation. An example of congenital malformations that exhibit familial clustering include spina bifida and cleft lip and palate but it is known that inheritance only explains the occurrence of a proportion of the cases.

Maternal drug use

The most notorious instance of maternal drug use that led to congenital malformations is that of thalidomide [6, 39, 40]. Thalidomide was thought to be a wonder drug mainly used in the treatment of leprosy and has been found recently to help control HIV. In 1961 the use of thalidomide in the treatment of morning sickness revealed the catastrophic effect of administering thalidomide to pregnant women. Thalidomide was found to increase the risk of rare congenital anomalies particularly phocomelia where the long bones of the limbs are absent or severely deficient but also heart defects, absence of external ears and malformed intestines, 7000 children were affected. A number of hypotheses exist surrounding the causation of these defects, one of which is the prevention of angiogenesis of the limbs between 20 to 36 days after conception. This hypothesis is developed from the knowledge that thalidomide can treat cancer by interfering with angiogenesis. Thalidomide was withdrawn from the market in response to these adverse affects but is gradually being brought back into use with the condition that it should not be used if the patient intends to get pregnant.

There are many other medicinal drugs that can lead to congenital malformations [6, 39, 41]. Anticonvulsants, like valprioc acid, used in the treatment of epilepsy increase the risk of craniofacial, cardiovascular and central nervous system abnormalities. Anticoagulants, for example warfarin, used to thin the blood increase the risk of nasal hypoplasia, bone dysplasia, choanal atresia, microcephaly and hydrocephaly. The risk of heart defects is increased in the treatment of bipolar disorder using lithium. Antihypertensive agents raise the risk of renal tubular dysplasia, oligohydramnios and defects of ossification. A high dose of vitamin A in the form of retinoic acid, which is often used in skin care, heightens the risk of craniofacial, cardiovascular and central nervous system defects along with mental retardation. Other problematic drugs include diethylstilboestrol, phenytoin and aminoptenin. Some drugs can be beneficial, for example, taking folic acid supplements is known to reduce the risk of NTDs and cleft lips and palates [6, 42]. Additionally, taking prenatal vitamins has been shown to decrease the risk of heart defects and limb deformities [43].

In some cases mothers use recreational drugs whilst pregnant, which is known to cause harm to the child [6, 8, 39, 41, 44]. Foetal alcohol syndrome is a range of adverse affects of maternal consumption of alcohol including craniofacial anomalies, heart defects, neural abnormalities and developmental delays. The impact of the use of cocaine and amphetamines is unknown but is theorized to include agenesis of the corpus callosum, brain clefts and limb reduction defects. Furthermore, cocaine passes directly into the foetus from the mother and the baby cannot metabolize it. Cocaine constricts the blood vessels decreasing the oxygen supply to the organs which leads to growth retardation.

Nicotine and marijuana pass directly from the mother to the unborn child, activating an adrenergic discharge that results in vasoconstriction and decreases the placental intervillous blood flow [44, 45]. This phenomenon is known to reduce oxygenation for the foetus and in turn increase the amount of carboxyhaemoglobin leading to prematurity, low birth weight and spontaneous abortions. Although there is not as much evidence of a link between smoking and congenital malformations, it is thought that smoking slows down the rate of cell replication, reducing the rate of DNA replication and protein synthesis, which could cause malformations. Tobacco also contains oestrogen which could have an impact on the formation of male genitalia. Although this is not proven there is evidence that smoking reduces the fertility of parents and also leads to more genetic mutations.

Viral infections and maternal health disorders

Viral infections can lead to congenital malformations [6, 25, 39, 46]. Rubella leads to increased risk of congenital malformations including cataracts, heart defects, hepatomegaly, thrombocytopenia and mental retardation. In the USA, between 1964 and 1965, 30,000 children were born with rubella-associated defects. Little is known about the teratogenesis other than the fact that the rubella infects the placenta and then the foetus, which is unable to get rid of the virus. The first 16 weeks of gestation is when the foetus is susceptible to developing disabilities as a consequence of infection.

Cytomegalovirus is a herpes virus that attacks and enlarges epithelial cells, increasing the risk of an unborn child developing microcephaly, ventriculomegaly, cerebral calcification, heart defects, hepatomegaly, thrombocytopenia, growth restrictions and mental retardation. Parvovirus is a term used for a group of viruses that contain DNA in an icosahedral protein shell and it results in aplastic anaemia and hydrops. Limb hypoplasia, microcephaly and chorioretinitis can result from maternal varicella or chicken pox. Finally toxoplasmosis is a single-celled parasitic infection leading to flu-like symptoms and maternal infection can lead to hydrocephalus, microcephaly, cerebral calcification and neurological problems.

Maternal diabetes, if not well controlled, provides an unfavourable environment for a developing embryo and, consequently, can result in heart defects, NTDs, abnormalities of the skeletal system, sacral agenesis and microcolon [6, 39, 41, 47]. Maternal epilepsy may also lead to congenital malformations however it is not possible to discriminate between the impact of the disorder itself and the drugs taken in its treatment.

Environmental agents and radiation

Little is known of the true results of exposure to specified environmental pollutants with the majority of links between environmental contaminants and congenital anomalies remaining unsubstantiated with little strong evidence [6, 37]. In fact, only exposure to mercury is shown to influence birth defects with clear evidence, studies of other environmental exposures display inconsistencies or inadequacies in the procedures of investigation adopted and suggest that there is a need for further investigation.

Mercury is used in electrical equipment, dentistry, paints, fungicides, pesticides, fluorescent lamps and batteries [39]. The release of methyl mercury from the Chisso Corporation acetaldehyde plant into the Minamata Bay in Japan led to many cases of a severe neurological disorder in 1955 [48]. The methyl mercury was ingested by fish which were then consumed by pregnant women leading to foetal toxicity. Births to exposed mothers developed cerebellar dysfunction which is characteristic of what is now referred to as congenital Minamata disease. Mercury poisoning as seen in Japan causes cerebral palsy, ataxia, disturbed psychomotor development, mental retardation, microcephaly and NTDs [37, 39]. In a study of the impact of arsenic, cadmium, lead and mercury on the risk of NTDs using urine biomarkers no affect was found for any of these metals but they do suggest that further work is required especially with mercury [49].

There is some evidence that exposure to radioactive materials can increase the risk of birth defects. In high doses, ionising radiation can lead to microcephaly, mental retardation and retarded growth [39]. Smaller levels of exposure to radiation do not have such conclusive evidence of influencing the rate of these congenital anomalies from studies of x-rays on pregnant women, exposure to radiation after atomic bombs dropped in Hiroshima and Nagasaki, residential proximity to Sellafield and Cardiff nuclear reprocessing plants and the Chernobyl disaster [6].

Endocrine disruptors are contaminants that influence hormone levels in the body. The changes in the parental hormone levels may affect the development of a foetus and could consequently lead to abnormalities. In addition to affecting the developmental stages of the foetus, it has also been hypothesised that endocrine disruptors could influence the gender of the child. In response to the hypothesis concerning the risk of congenital anomalies numerous studies of exposure to known endocrine disruptors have been carried out for dioxins such as 2,3,7,8-tetrachlorodibenzene-p-dioxin (TCDD), polychlorinated biphenyls (PCBs) and trihalomethane (THM).

A study of an explosion in Seveso, Italy, that led to the exposure of the population to TCDD, did not find an increase in the rate of congenital anomalies but there was a problem of underreporting [6]. TCDD is also found in a compound called Agent Orange used in Vietnam however studies of veterans and Vietnamese nationals have provided conflicting results [6].

Other studies of TCDD suggest there may be a relationship with spina bifida but an insufficient sample size prevented the gathering of conclusive evidence [50].

In Japan exposure to PCBs through contaminated cooking oil lead to the congenital 'Yusho' epidemic with gingival hyperplasia, natal teeth, spotty calcification and irregularities of the skull amongst some of the associated conditions [6, 37]. All affected births had a deep brown pigmentation of the skin and nails which led to the term cola-coloured babies.

Exposure to THM through municipal water supplies was investigated in relation to the risk of NTDs using two datasets for California [51]. One dataset revealed an inverse relationship between exposure to THM and the risk of congenital anomalies and the other indicated no affect. The explanation given in the paper was that high THM exposure could possibly lead to more spontaneous abortions of NTDs which would result in an apparent decreased risk of NTDs amongst live births.

Many studies focus on putative point sources of pollution and one of the more common subjects for these types of study is incinerators or crematoriums since these are major sources of dioxins and atmospheric metals. There is some evidence that increased risk of facial clefts, renal dysplasia, heart defects, anencephalus and NTDs is associated with exposure to incinerators [52, 53] but in other investigations no affect was detected [54]. However, although some studies incorporated meteorological factors and stack height into their exposure measures, the typical approach was to treat distance from the site as a proxy for exposure. Hence, there may be a high degree of exposure misclassification. Additionally, one study looked at the rate of congenital malformations before and after the opening of the site and found no change [53].

In a study of ambient air quality in Texas in 1997 to 2000, associations were found between carbon dioxide and conotruncal heart defects; sulphur dioxide and isolated ventricular septal defects; and PM_{10} and isolated atrial septal defects [55]. Additionally, the risk of cardiac anomalies, obstructive uropathies and skin anomalies increased with road traffic density linearly [52]. However, deficiencies in the way exposure was measured in both studies mean that further investigation is required.

Organic solvents are used in adhesives, cleaning materials, pesticides, the photographic industry, metal cleaning, dry cleaning, paint, anaesthetics and petrol [39]. Exposure to organic solvents is thought to increase the risk of oral clefts, central nervous system defects, transposition of the great arteries and a hypoplastic left heart. Single site studies also suggest that further results of such exposure could include gastroschisis, intestinal agenesis and urinary anomalies. Skin, nail, teeth and facial anomalies along with small penises in boys are the possible effects of exposure to polychlorinated biphenyls, which are used in hydraulic fluids, plasticizers, transformers, capacitors and carbonless copy paper. Trichloroethylene is an industrial solvent which could be linked to facial, central nervous system, chromosome and cardiac defects when the drinking water supply is contaminated by it. If drinking water is contaminated with nitrates then the risk of anencephaly, musculo-skeletal defects and central nervous system anomalies could be heightened.

At the beginning of this thesis the concern about exposure to landfill sites and the risk of congenital malformations was highlighted. Waste sites are the subject of many studies into the environmental risk factors of birth defects and these studies are discussed in the next section.

1.2.6 Landfill Sites and Congenital Anomalies

A detailed account of studies on the health impacts of living in the vicinity of a landfill site can be found in a paper by Vrijheid [56]. Some of the studies on landfill sites and congenital anomalies are discussed in detail below.

American and Asian studies

In a case-control study of landfill sites and industrial facilities in Texas, births to mothers living within 1km of the centre of landfill sites and industrial facilities were compared to those to mothers residing beyond 1km of any of the sites [57]. The odds ratio of congenital malformations for mothers living within 1km of any of the sites was 0.84 (95% CI: 0.54-1.3) with adjustment for maternal ethnicity, education and smoking status. This odds ratio provides evidence that there is no increased risk of congenital malformations associated with living in close proximity to any of the landfill sites.

A thorium waste disposal site in Wayne Township, Passaic County in northern New Jersey was located within 100 feet of buildings and a drainage ditch on the site's border runs into a brook that flows through built up areas [58]. A single site study was performed using 76 households within three blocks of the site as exposed and 76 households between nine and ten blocks away as controls. The risk ratio was 2.1 (95% CI: 0.62-7.25) for birth defects which indicates that although the risk of congenital malformations is elevated for the exposed blocks statistical significance is not achieved. However, statistical significance is difficult to achieve with such small numbers as reflected in the wide confidence interval.

Alaska Native villages are remote and the solid waste management is deficient [59]. The open landfill sites, containing all the waste generated by the villages, are not maintained, unlined, contain uncovered waste and generally have no marked perimeter and are open access. Gilbreath et al. rank the landfill sites for each of the 197 villages using hazard point factors. The study found that there was no statistically significant association between exposure and the rate of congenital malformations with a relative risk of 1.37 (95% CI: 0.92-2.02).

The national industrial hazardous waste site in Southern Israel is on an industrial park and contains 35,000 tonnes of waste [60]. The study population was those births that took place in the Soroka University Medical Centre between 1995 and 2000. The exposed population comprised those births within 20km of the industrial park. The relative risk for major congenital anomalies amongst the Bedouin population living in traditional settlements was 1.63 (95% CI: 1.39-1.80) and for the Bedouin population living in permanent localities it was 1.79 (95% CI: 0.96-3.35). Therefore, there was an increased risk of major congenital anomalies within 20km of the site for the Bedouin population however this was not seen for the Jewish population where no effect was found. The explanation given by Bentov et al. for the disparity between ethnic groups is that Bedouin women are more confined to their place of residence than Jewish women and consequently there is a smaller degree of exposure misclassification.

European studies

In an ecological study of 48 out of the 52 landfill sites in Denmark, data from the Danish Birth Defect Register were considered for 1997-2001 [61]. Three zones of exposure were considered based on distance from the site and no relationship between the risk of congenital

anomalies and exposure was found. However, no adjustment for confounders was made and the denominator was determined by building density which is not an accurate measure of the residence of mothers at the time of birth.

Boyle et al. performed a small area study of landfill sites in the Eastern Region of Ireland in 1986-1989 [62]. 83 landfill sites were considered and district electoral divisions were used to determine the number of cases and randomly selected controls within a certain radius of the centre of the landfill sites. The odds of congenital malformations born to mothers residing within 3km of any site at the time of birth were compared with those living from 3km to 7km from the sites. No adjustment for confounding was made in the calculations of the odds ratio. The odds ratio for all landfill sites was 0.90 (95% CI: 0.78-1.04) and for the five major landfill sites only it was 0.94 (95% CI: 0.74-1.19) suggesting that there is no relationship between exposure to landfill and congenital malformations.

EUROCAT is the European network of registers for the epidemiologic surveillance of congenital anomalies which contains information on 1.2 million births [63]. It feeds off 36 registries in 18 countries [36]. The EUROHAZCON study is based on information from EUROCAT. 21 hazardous landfill sites in 5 European countries are considered in this case-control study [64]. Two controls are randomly selected from all births without congenital anomalies for each case. Distance is used as a proxy for exposure with births to mothers residing within 3km of the centre of the sites compared to those living between 3km and 7km of the sites at the time of birth. The odds ratio for non-chromosomal anomalies amongst mothers living within 3km of these hazardous landfill sites is 1.33 (95% CI: 1.11-1.59) with adjustment for maternal age and socioeconomic status. Additionally, it was also reported that there was a consistent decrease in risk with increased distance from the sites in six distance bands. In a similar study, also carried out as part of the EUROHAZCON study, on 23 landfill sites the equivalent odds ratio of chromosomal anomalies was 1.41 (95% CI: 1.00-1.99) after adjustment for maternal age [65].

Vrijheid et al. also carried out an extension of the EUROHAZCON study where the landfill sites were categorized by an expert panel based on documented data into three hazard levels [66]. There was a reported increased odds ratio with increasing water hazard ranking for the landfill sites. However, when air hazard and overall hazard are considered there was no such pattern. The results are similar for chromosomal anomalies. Harrison suggests that the lack of

relationship found in the ranking study as opposed to the main study may be either a result of no causality with a chance result in the first study or the inability of the experts chosen to rank landfill sites correctly [67].

British studies

Elliott et al. performed a study on 19,196 landfill sites in Great Britain using data from registers held by the Small Area Health Statistics Unit (SAHSU) [68]. Births to mothers within 2km of the centre of any of the sites were compared to those births to all other mothers. The relative risk was 1.01 (99% CI: 1.005-1.023) with adjustment for deprivation, year of birth and gender. This suggests that there is only a very slight yet statistically significant increase in risk associated with living within 2km of the centre of the site which is a huge contrast to the EUROHAZCON study where there is a 33% increase in risk associated with living within 3km of their identified sites. The disparity could be down to the choice of landfill site as many of the landfill sites used by Elliott et al. will be benign [69]. Another important difference between the studies is the choice of background regions with Elliott et al. selecting the entire remaining population as unexposed whereas in the EUROHAZCON study the unexposed region was close to the landfill site but not within 3km.

In a similar study 61 Scottish special waste sites were investigated [70]. The adjusted relative risk for congenital malformations within 3km of the centre of the sites was 0.96 (99% CI: 0.89-1.02). This result suggests that there is no increased risk of congenital malformations surrounding Scottish sites but there is around the remaining sites in Great Britain. The reasoning given by Morris et al. for this disparity is the socio-demography of the areas surrounding the landfill site and the type of management of waste in Scottish special waste sites. However, Irvine highlights some further possible explanations including better quality of data collection in Scotland, the landfill sites in England are larger and receive more waste, statistical significance is not the same as aetiological significance and inadequate adjustment for socio-economic deprivation for England and Wales [69].

Fielder et al. performed a study on the Nant y Gwyddon landfill site in Wales from 1983 to 1997 [9]. Cases were taken from the National Congenital Anomaly System (NCAS) held by the Office of National Statistics (ONS). The exposed region comprised five wards within 3km of the centre of the site, from which complaints had been received regarding the site, and the

unexposed region included 22 wards with similar socioeconomic level. The investigation revealed that the incidence of all congenital anomalies was raised in the exposed wards both before and after the opening of the site. In contrast, in a similar study of the Trecatti landfill site northeast of Merthyr Tydfil in Wales, the incidence of all congenital anomalies was raised in the 3 exposed wards compared to the 18 unexposed wards after opening but not before [71]. The relative risk of birth defects in the exposed region after opening was 1.9 (95% CI: 1.3-2.9). Dolk criticises these studies indicating that the exposure assessment needs improving, the reporting rates to the ONS have been poor and single site studies lack power [72].

A multiple site study has been performed on Welsh landfill sites from 1983 to 1997 using the ONS data and denominator data from the Child Health System [73]. 20 sites were identified in Wales that opened in that period and the distances of maternal residence from the centre of the nearest site were calculated. The relative risk of congenital anomalies within 2km of the centre of the sites was 0.87 (95% CI: 0.75-1.00) before opening and 1.21 (95% CI: 1.04-1.40) after opening with adjustment for maternal age, year of birth, hospital of birth, gender and socioeconomic deprivation. The standardized risk ratio was 1.39 (95% CI: 1.12-1.72) which suggests a statistically significant increase in the risk of congenital anomalies within 2km of the centre of the sites after their opening. The results were similar when considering births within 3km as the exposed region. However, data for 1998 to 2000 do not indicate a significant effect with a relative risk of 1.04 (95% CI: 0.88-1.21).

In a paper by James et al. contours for the relative risk of congenital anomalies in 1983 to 1997 were created in the 20km square region centred over the Nant y Gwyddon landfill site [10]. The relative risk surface was adjusted for maternal age, year of birth, gender, hospital of birth and socioeconomic status. There were two regions over which the relative risk surface was elevated above 2: one to the north east of the landfill site within 3km and a second to the northwest at about 7km from the site. The existence of elevations in the relative risk near to the Nant y Gwyddon landfill site suggests that exposure to substances released from the landfill site may influence the risk of congenital anomalies. The concluding remarks in this paper highlight the need for new methodologies for assessing the global significance of these regions of heightened relative risk and to compare spatial contours of risk to those of exposure to a given risk factor.

1.3 Gender and Sex Ratios

Congenital malformations are not the only birth outcome that is under constant debate in the literature. The sex ratio is conventionally quoted as either the proportion of males born or the ratio of boys to girls. The primary sex ratio is the sex ratio at conception, the secondary sex ratio that at birth and the tertiary sex ratio that at sexual maturity. The sex ratio at birth has been under discussion for years as a public health concern since, as Chahnazarian states, an imbalance between the sexes may "...constitute a threat to social norms and values" and alter the population growth through the availability of mates, the age at marriage and the age of childbearing [74]. Hesketh et al. outline specific consequences caused by an excess of males including an expansion of the sex industry and increased violence and anti-social behaviour threatening the stability and security of society [75]. Evidence of this exists as a result of a preference for males in some cultures which has lead to an excess of male births, it is not clear what the result of an excess in females would be.

The social problems that populations can face when an imbalance in the sex ratio is encountered have triggered an interest into determining the aetiology of the sex ratio in the hope that such imbalances can be avoided. However, the sex ratio, like most congenital malformations, has a relatively unknown aetiology motivating ongoing investigations and there is growing concern that environmental factors may be implicated. Consequently, the sex ratio provides another application which could benefit from the spatial methodology presented in this thesis.

The use of the investigation of sex ratios and congenital malformations as applications of the new spatial statistical methodology has some practical advantages. Information on both sex ratios and congenital anomalies is collected by a number of registers and surveys that also collect information on some possible risk factors in addition to the geographical location of the residence of the mother. The collection of information on all births by some of the registers and the ability to link registers means that investigations into either birth outcome have the advantage of good quality denominator data. Investigators looking into other health outcomes, such as cancer, have to use patients with another health outcome as a replacement for more appropriate denominator data. Both the sex ratio and congenital malformations have relatively short latency periods, which reduce the impact of migration, a major source of bias, affecting any spatial analysis.

It is clear that there are adverse social implications of an imbalanced sex ratio. In order to avoid imbalances, through successful prediction of when an imbalance may occur, it is necessary to understand what factors influence sex ratios. The relatively unknown aetiology for sex ratios led to the investigation into sex ratios presented in Chapter 5.

The gender of a human is dependent on the sex chromosomes of the conceptus, which are XX in females and XY in males. The mechanism by which the gender is determined has Mendelian properties which should result in a 50% chance of having offspring of either sex. Therefore, it is expected that the sex ratio for a population should be equal to 0.5. However, there is substantial evidence that the sex ratio varies spatially and temporally. The most prominent of which is the general decrease in sex ratios through time in post-war industrialized countries [76-78]. These heterogeneous patterns in the sex ratio are not understood but the general consensus is that son and daughter production can be, in some way, altered in response to certain environmental conditions.

Edlund's study into sex ratios in countries where infanticide and prenatal sex selection is practised reveals some of the reasons why it is important to try to understand imbalanced sex ratios [79]. In China in 1990 the sex ratio was 0.533 compared to the sex ratios for the world population of around 0.515 suggesting that China has significantly fewer females than the world population. This relative lack of females means that there is a backlog of unmarried men which is increasing the age gap between partners. Women, as the minority sex, will be able to choose their partners and are more likely to select a partner of a higher social standing than her own increasing the risk of celibacy for males of a low social status resulting in sex-determined social status. However, these problems in finding partners, that the male population are exposed to, will increase the paternal stress levels which many researchers believe increases the likelihood of having female children. Thus, the situation may eventually correct itself and tend towards a more balanced sex ratio. This underlines the importance of further understanding of the sex ratios because if that is not the case China's population may be heading towards some serious social problems in the future.

Many theories regarding the reasoning, mechanisms and conditions behind these variations have been proposed but gathering conclusive proof is difficult. Some of these theories, but by

no means all, are presented here along with some insight into the difficulties of testing these conjectures.

1.3.1 Evolutionary Theory

In 1741 Sussmilch took a religious approach to sex ratios and theorised that our creator ensured there were more boys than girls born to adjust for the higher losses of boys, through recklessness, exhaustion and dangerous tasks, and to ensure that everyone was able to get married at the appropriate time [74]. Subsequent theories take a more scientific approach and are discussed below.

Beginning with Darwin and his Theory of Natural Selection in 1871 many scientists have provided suggestions as to why human sex ratios are on occasion biased towards a certain sex, some of which are more plausible than others. All these theories on the variation of sex ratios hinge on the idea that the aim of an individual is to promote the passage of their genes through time by maximising the number of descendents they have [80]. The most famous of these theories is the Trivers-Willard hypothesis which states that under certain well-defined conditions natural selection favours systematic deviations from a sex ratio of 0.5.

The Trivers-Willard hypothesis is discussed in depth in a large number of books and papers, the main points of which are given here [81-89]. The hypothesis makes three assumptions. Firstly, it is assumed that a female in good condition is more able to bear and nurse their offspring so that by the end of the parental investment period their children are stronger and healthier than those of a mother in a poor condition. Secondly, this disparity in the condition of offspring is assumed to be maintained into adulthood. Thirdly, the difference in conditions of adults has a greater impact on the reproductive success for males than it does for females. If these three conditions are met then a female in good condition will maximise the number of grandchildren by having male offspring, whereas a female in poor condition will maximise the number of grandchildren by having females. Therefore, in females a deviation from the mean condition will bias the sex ratio of their offspring towards a specific sex.

The assumptions made are more applicable to animals than to humans. The condition of the mother typically influences their ability to invest into their children and those children will maintain their condition into adulthood whether their condition is good or poor. Animal males

that are in a good condition will compete with other males and successfully prevent males of a poor condition from breeding with more females, whereas females only influence their reproductive success through their ability to invest in their young. Therefore, the condition of the male influences their reproductive success more than it does in females.

In humans, offspring benefit from not only maternal but also paternal investment. Hence the model only works if at the upper end of the condition scale male reproductive success exceeds that of his sisters and at the lower end the reproductive success of a female exceeds that of her brothers. This phenomenon is true to a certain extent with the propensity for females to marry above their social standing. The Trivers-Willard hypothesis has its supporters and its critics with studies both corroborating and nullifying the theory. The difficulty behind assessing the validity of this hypothesis in humans is discussed in depth by Koziel et al. [90].

Supposing the Trivers-Willard hypothesis was correct then it would be expected that in populations where the condition of females does not vary hugely the sex ratio would be extreme. However, in the majority of studies conducted, the sex ratio, although not equal to 0.5, does not hugely deviate from this value. The lack of extreme sex ratios can be explained by a theory laid down by Fisher in 1930, however, 40 years prior to that an analogous theorem was provided by Düsing which is little known and this is given here.

Düsing explains that deviations from a balanced sex ratio have a tendency to be self-correcting where an excess of one gender stimulates the production of another [86]. Suppose there is a lack of females with x females and nx males. The population produces z offspring so each female produces z/x offspring and each male produces z/nx offspring. If a female produced more female offspring then the daughters will collectively produce a larger number of offspring. This concept is shown mathematically below.

Suppose mother 1 has A sons and a daughters and mother 2 has a sons and A daughters then mother 1 will have

$$A\frac{z}{nx}+a\frac{z}{x}$$

grandchildren whereas mother 2 will have

$$a\frac{z}{nx}+A\frac{z}{x}$$

grandchildren. Assume that A > a such that A = ba (b > 1) then mother 1 will contribute

$$\frac{az}{x}\left[\frac{b}{n}+1\right]$$

individuals to the second generation whereas mother 2 will contribute

$$\frac{az}{x}\left[\frac{1}{n}+b\right]$$

individuals. Therefore mother 2 contributes

$$\frac{1+bn}{b+n}$$

times as many individuals as mother 1 to the second generation.

If the sex ratio is 0.5 then n is equal to 1 and the two mothers produce the same number of descendents in the second generation regardless of the gender combination. However, this is not true for a sex ratio that deviates in anyway from 0.5. Suppose there are twice as many males as females then n is 2 and mother 2 will have

$$\frac{1+2b}{b+2}$$

times as many offspring as mother 1 in the second generation. In this case a mother producing a threefold female excess will have 7/5 times as many grandchildren as a mother with a threefold excess of males. Therefore, the mother can maximise the number of grandchildren she has by biasing her sex ratio in the opposite direction towards which the population sex ratio is biased.

Here, the relative number of grandchildren is used as a measure of fitness, so fitness is a function of the parent's progeny sex ratio, b, given the population sex ratio, n. Fitness is unaffected by progeny sex ratios if and only if the population sex ratio is 0.5. However, in a population with an unbalanced sex ratio it may be advantageous for individuals to adjust their sex ratio to be biased towards the opposite gender to the population sex ratio, because this maximises the number of grandchildren they have or their fitness. The benefits of adjusting the progeny sex ratio means that the sex ratio is self correcting and this explains why sex ratios do not deviate too far away from 0.5.

There are a number of other theorems that attempt to convince the scientific community that if the population sex ratio moves away from 0.5 in one direction then progeny sex ratio

adjustment in the opposite direction can optimise the number of descendents in the second generation for those parents. Therefore, the sex ratio is self correcting for imbalances away from a value of 0.5. These hypotheses are based on different concepts with Düsing focusing on fitness, Fisher on expenditure and the Shaw-Mohler equation on genetic contribution.

Combining the theorems involving adjustments in progeny sex ratios with the Trivers-Willard hypothesis suggests that there could plausibly be a cycle in the sex ratio brought on by the need of humans to promote the passage of their genes into future generations. The Trivers-Willard hypothesis indicates that depending on the condition of the mother the sex ratio may move away from that of 0.5 to increase the number of grandchildren she has. However, when the population sex ratio is unbalanced it becomes advantageous for individuals to adjust the progeny sex ratio in the opposite direction to optimise the number of people in their future generations. Therefore, the sex ratio is self correcting when it becomes unbalanced. Thus this cycle could explain why the sex ratio can deviate from 0.5 and also why that deviation never becomes too large.

Many researchers in the field of sex ratios find sex allocation theory, particularly the Trivers-Willard hypothesis, attractive even though all the theorems discussed have yet to be validated. There is a wide range of literature on sex allocation theory incorporating increasingly complicated models that have been devised over the years, only the simplest of which have been discussed here. There are many alternative theorems too, for example, the theory of attractiveness as outlined by Burley [91]. Sex allocation theory provides a good philosophical insight into the possible reasons for variations in the sex ratio but it does not reveal which factors trigger changes in the sex ratio and the mechanism through which the adjustment manifests itself. These factors and mechanisms are arguably of greater interest to the scientific world.

1.3.2 Mechanisms that may adjust the Sex Ratio

Meiosis produces equal numbers of X- and Y-bearing spermatozoa so fertilization should lead to equal numbers of males and females. However, in the majority of countries there are more boys than girls, for example, in the United States the sex ratio is 0.515 [92]. There has to date been no mechanism that fully explains this deviation of the sex ratio from 0.5 but in this section some of the theories are discussed.

In some countries, like India, China and South Korea, there is a preference towards male offspring as they can work and bring more wealth back into the family, whereas women leave the family to marry another man. In extreme cases this want of males over females has led to female births going unreported and subsequent infanticide or, with the increased availability of ultrasound, sex selective foeticide [75, 93]. At a less drastic level prospective parents can alter their sexual behaviour in a way that they believe increases the likelihood of having a son, for example, altering the timing of intercourse, deeper penetration, maternal diet changes or enhancement of the male spermatozoa count [94].

However, these practices are not common in places, like the United States and Europe, where there is also a larger number of male offspring. There are three possible sources of bias in the sex ratio: the proportion of X- and Y-bearing spermatozoa in ejaculate, the relative success of X- and Y-bearing spermatozoa in fertilization and the proportion of XX and XY embryos that are spontaneously aborted [86]. James claims that all these plausible mechanisms are controlled by the hormonal levels of the parents, with high concentrations of testosterone or oestrogen increasing the probability of a son and gonadotrophins or progesterone increasing the chance of having a daughter [89, 95-98]. There is evidence that this hypothesis is true, particularly in a study on phenotypes that are associated with low, neutral and high testosterone levels that revealed a statistically significant link between this and the sex ratio [99]. The sex ratio for the low testosterone phenotype was 0.340 and that of the high testosterone phenotype was 0.561. The main focus of James' papers is the testosterone/gonadotrophin ratio, a low value of which is associated with a greater probability of having a daughter.

The initial thought is that perhaps the bias in the sex ratio is a product of the number of spontaneous abortions. This assumption is certainly true to some extent particularly in developing countries. There is an excess of males in spontaneously aborted foetuses, therefore, in India as the overall foetal wastage rate fell, as a response to improvements in female well-being, this had a greater impact on the number of males that survived to term. Thus, the sex ratio increased in response to fewer foetal deaths, a phenomenon that Jayaraj et al. claim had a greater impact on the sex ratio in India than the sex selective foeticide and infanticide [94]. Krakow believes that the survival of a foetus can be affected by hormonal levels and the two genders will respond differently to this leading to a biased sex ratio

amongst foetal deaths [100]. However, the larger number of male foetal deaths suggests that for the secondary sex ratio to be biased towards males in the way that it is, the bias in the primary sex ratio must be even larger [92]. Therefore, there must be more to the sex ratio than the rate of spontaneous abortions.

Pergament et al. reveal that X- and Y-bearing spermatozoa present in equal numbers to the female tract, therefore, the female tract could be in some way favourable to the Y-bearing spermatozoa, which is corroborated by Krakow [92, 100]. The ability of the female tract to select between X- and Y-bearing spermatozoa may manifest itself through the different physiological activities of the two genotypes. The Y-bearing spermatozoa have smaller heads, length, perimeter and area than the X-bearing spermatozoa and this translates into a greater level of motility. The difference in the two genotypes in terms of motility means that the viscosity of the cervical mucus determines the penetrability and can consequently influence the primary sex ratio, with the impact on the motility of the Y-bearing spermatozoa being less pronounced [86]. The viscosity of the cervical mucus is controlled by maternal oestrogen levels and, thus, this proposed mechanism relates to James' theory that the sex ratio is controlled by hormone levels. Another suggestion is that the pH in the vaginal tract may also influence the X- and Y-bearing spermatozoa differently and influence their success in fertilization [100].

The discussion on mechanisms takes a full circle as Boklage indicates that studies either show that there are equal numbers of X and Y fertilized eggs or that there is a slight, but not statistically significant, excess in the number of XX embryos [101]. Therefore, the excess of males is not a result of a bias in spermatogenesis or in fertilization but in a preferential loss of females during the embryonic period, thus, bringing the focus back onto spontaneous abortions. An element of the genomic imprinting of the respective paternal sex chromosome is that male embryos progress through the early stages of development more rapidly and, therefore, a viable pregnancy is established more efficiently with males. The result of this is a sex ratio biased towards females amongst embryo losses in the first trimester. Fewer than 25% of births reach full term with 2/3 of the losses occurring before embryonic recognition. Therefore, this larger proportion of pregnancy wastage in the embryonic period outweighs the consistent excess of males in spontaneous abortions during the foetal period. This alleged excess in female foetal loss cannot be identified because the parents are typically unaware of the pregnancy let alone the gender of the embryo at this time.

Krakow backs this concept up with a discussion of how the sex ratio may be influenced during the pre-implantation period when the number of X and Y fertilizations is equal [100]. The ova may be able to select the sex by failing to undergo the second phase of meiotic division when fertilized by an undesirable genotype, thus, failing to form a zygote of a specific sex. Another difference between males and females is that male blastocysts develop more rapidly than female ones, therefore, the length of time it takes the uterus to undergo the sufficient level of preparation for successful implantation can determine the survival of the blastocyst in a sex specific way. It is difficult to investigate how the sex ratio can be influenced before fertilization, during fertilization and during the implantation period without resorting to controlled experiments in animals. Moreover, it may not be possible to generalize the results of these experiments to human sex ratios. However, regardless of the mechanisms involved in the biasing of the sex ratio, it is important to try and understand which factors influence the sex ratio.

1.3.3 Conditions that influence the Sex Ratio

Similar to congenital malformations, despite extensive literature on possible risk factors, little is known of the true aetiology of sex ratios. Many of the papers have conflicting results and no definitive answers have been reached regarding causal links, which may be a result of a complex multifactorial aetiology. The uncertainty surrounding the aetiology of the sex ratio makes it a good birth outcome to investigate. Some of the investigations into possible risk factors for the sex ratio are discussed below.

Environmental contaminants

The decrease in sex ratios in post World War II industrialised countries led to the investigation into environmental contaminants as a possible sex determinant, especially those with endocrine disrupting properties. For example, the decrease in sex ratio in the Aamjiwnaang First Nation in Canada is thought to be linked with exposure to dioxins through the consumption of fish from the Great Lakes [102]. However, as Jongbloet et al. reveal the 18th and 19th centuries were characterised by an increasing sex ratio and environmental pollution was almost non-existent during this period [103]. Therefore there must be other

factors at play, although, this is not to say that environmental factors can be ruled out when discussing sex ratios.

In their study of sex ratios in Italy, Figa-Talamanca et al. revealed that for two Italian cities the sex ratio was significantly lower in the urban areas compared to the rural areas [104]. For example, Bari had a sex ratio of 0.507 in urban regions and 0.524 in rural locations. However this did not generalize to the whole of Italy and, therefore, this effect could merely be a random event. On the other hand, using whether a location is urban or rural as a measure of exposure and the ecological nature of this study may mean that important true effects were missed.

Milham took a more refined approach to studying the impact of environmental contaminants on the sex ratio by considering aluminium workers [105]. The sex ratio amongst all aluminium workers was 0.512 but when considering only those who were classed as carbon setters, anode setters or carbon changers the sex ratio was only 0.381. However, no explanation for this effect is provided.

In a study by Lichtenfels et al., concentrations of particulate matter in the air were measured from 9 selected air monitoring sites in São Paulo, Brazil [106]. The sex ratios amongst the births to mothers who resided within 2km for each of these 9 monitoring sites were calculated and plotted against the concentration of particulate matter measured from the corresponding site. The plot revealed that the sex ratio fell with increasing concentrations of particulate matter.

In Minamata there was widespread methylmercury pollution in the 1950s which lead to a reduced sex ratio particularly amongst the offspring of fishermen where the sex ratio was 0.382 compared to the Japanese national sex ratio of 0.515 [107]. The problem lay in maternal not paternal exposure and the investigators revealed that the sex ratio amongst stillborns for those exposed was significantly higher than that of the nation as a whole, suggesting that the problem was an excess in male foetal losses.

In Canada's 'Chemical Valley', the location of numerous petrochemical and chemical plants and one of Canada's largest hazardous waste dumps, the sex ratio was as low as 0.348 which has a p-value of 0.01 [108].

Simonsen et al. found that inorganic lead lowers the odds ratio for having male offspring with an odds ratio of male births equal to 0.75 associated with an increase of 20μg/dl in the paternal blood lead concentration [109]. However, in contrast Jarrell et al. report no association between the sex ratio and maternal blood lead suggesting that if exposure to lead is a risk factor for the sex ratio then the effect is a paternal one [110].

In 1987 Saddam Hussian ordered a chemical weapons attack on a small Kurdish city called Sardasht. Subsequently in 1988 there was a statistically significant increase in the sex ratio [111]. The explanation given by Saadat for the increase in the sex ratio is that mustard gas influences the levels of testosterone in males.

The greatest focus of environmental based studies of sex ratios is on dioxins that are thought to have endocrine disrupting properties [98], for example, polychlorinated biphenyls (PCBs). In Yu-Cheng, Taiwan 2000 people were exposed to PCBs through the use of contaminated cooking oils. There was no change in the sex ratio for exposed mothers but the sex ratio amongst exposed fathers was only 0.46 compared to the control of 0.54 [112]. In another study of the same event the sex ratio for the exposed fathers was given as 0.49 but this was still significantly lower than that of the control population [113].

The Great Lakes were contaminated with PCBs which in turn led to the bioaccumulation of PCBs in fish consumed by locals. Weisskopf et al.'s study into this problem revealed that the odds of a male child decreased by 46% with every unit increase in the natural log of the maternal PCB serum concentration [114]. In another study of the Great Lakes, the results for maternal exposure to PCBs was similar, with a sex ratio of 0.546 when the PCB serum level was below 8.1 and when the serum level was higher the sex ratio was 0.494, although, this difference was not statistically significant [115]. However, rather than revealing no relationship between the sex ratio and paternal exposure the sex ratio was significantly higher when the PCB serum level was greater than 8.1 at 0.571 compared to 0.451 when the PCB serum level was below 8.1. The difference between the two studies on the Great Lakes may relate to the adjustment for the partner's exposure level that was applied in the latter study. Therefore, this was considered to be a more accurate study and in contrast to Yu-Cheng higher paternal exposure increases the sex ratio according to the Great Lakes data.

In a study of 600 Russian pesticide workers who were exposed to tetrachloro-dibenzo-p-dioxin (TCDD) and pentachlorodibenzo-p-dioxin (PnCDD) investigators found that the sex ratio amongst exposed fathers was significantly lower at 0.38 compared to the control of 0.512 [116]. In Seveso, Italy, exposure to TCDD in a 1976 explosion also had a significant impact on the sex ratio [117, 118]. If both parents were exposed the sex ratio was 0.442 compared to 0.608 when neither of them had been exposed. The effect seems to be mainly a paternal one, where the sex ratio amongst exposed fathers whose partners were not exposed was 0.436 whereas when the mother was exposed and the father was not the sex ratio was higher at 0.545.

Jongbloet et al. hypothesize that exposure in fathers to dioxin-like compounds alters spermatozoa transit times delaying the fertilization of the oocytes so that the oocyte is older when it is fertilised, this is referred to as post-ovulatory overripeness ovopathy (PoOO) [103, 117]. PoOO can lead to defective implantation, retardation in the rate of zygote development and an increase in prenatal loss. These detrimental consequences are more common occurrences after fertilization with a Y-bearing spermatozoon because the Y chromosome cannot compensate for some of the sub-lethal X-linked genes from the ovum whereas X chromosomes can dominate these recessive genes.

Exposed mothers do not always see such a dramatic drop in sex ratios because dioxins can alter the production of oestrogen and, therefore, may result in suboptimal vaginal conditions that promote an increased proportion of Y fertilizations. Therefore, the sex ratio of exposed mothers is a balance between a larger proportion of Y fertilizations and an increase in the proportion of XY conceptus wastage [117].

The contrast between the rice oil and fish consumption incidents may be down to the relative level of exposure, with the rice oil the exposure was more severe. However, dioxins can result in hormone imbalances and, depending on the type of dioxin, that hormone imbalance can be of an oestrogenic, anti-oestrogenic or anti-androgenic nature [95, 113]. In the rice oil incident there was a greater concentration of anti-oestrogenic coplanar PCBs which may have resulted in reduced motility in the spermatozoa and consequently PoOO, thus, the sex ratio fell. However, in the Great Lakes situation the PCBs may have been oestrogenic thus increasing the sex ratio. An example of how the different types of dioxin influences the sex ratio is given

by Taylor et al. where there was an elevation in the sex ratio with increased maternal blood serum levels of oestrogenic PCBs and a decreased sex ratio for the highest maternal blood serum levels of anti-oestrogenic PCBs [119]. However, this study did not achieve statistical significance which may be the result of the small number of mothers included in the study (n = 50).

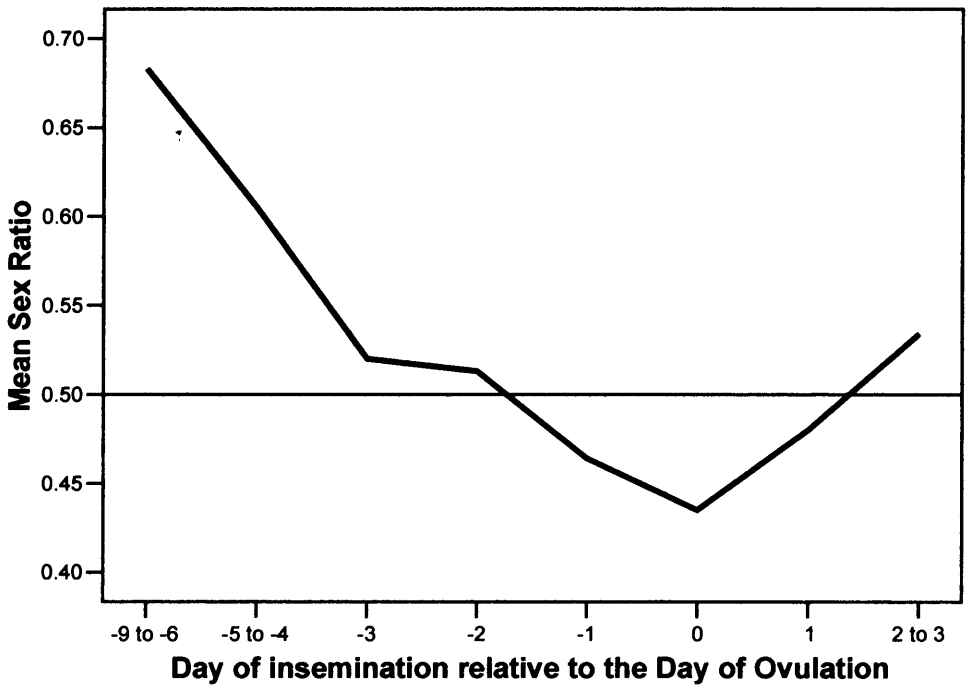
The impact of environmental exposures on the progeny sex ratio may depend on the parents' genotype. Glutathiane S-transferases (GST's) are involved in cellular detoxification of several toxins, including compounds that are present in petrol. Some genotypes have an absence of the enzyme activity associated with GST's meaning that they would be more susceptible to exposures to such toxins relative to those with genotypes that do have active enzymes. Anasari-Lari et al. investigated the impact of exposure to petrol amongst gasoline workers [120]. The sex ratio for the gasoline workers was 0.409 and that of the controls was 0.516 suggesting that when exposed to petrol the sex ratio is significantly lower. Further to this they considered the genotypes and discovered that gasoline workers with active GSTM1 and GSTT1 genes have a sex ratio that was not statistically significantly different to that of controls with active GSTM1 and GSTT1 genes (OR: 0.66; 95% CI: 0.34-1.28). However, amongst people with an active GSTM1 and an inactive GSTT1 the sex ratio for the gasoline workers was significantly lower than that of controls (OR: 0.45; 95% CI: 0.21-0.96). Therefore, the impact of the environmental contaminants on the sex ratio could depend on the composition of the population's genotypes and, hence, may not be easy to identify.

Timing of insemination

The time of insemination influences the gender of the child. Guerrero refers to the day of ovulation as day 0 and all other days are numbered relative to that [121]. The results of his study are presented in Figure 1.4. The sex ratio is as high as 0.683 for between day -9 and -6 indicating that early inseminations hold a greater probability of having a son. The sex ratio then falls as the time of insemination tends towards the day of ovulation to 0.435. Subsequent inseminations have gradually increasing sex ratios again with a sex ratio of 0.534 between days 2 and 3. Therefore, insemination around the time of ovulation carries a greater chance of having a daughter and the probability of a son increases as the time between insemination and ovulation increases. The estimate of the time of insemination relative to the time of ovulation is based on changes in body temperature and may not be accurate but this would only act to

reduce the strength of the association seen here. Although the figures differ, possibly as a function of the method of assessing the time variable, Harlap revealed consistent evidence of this U-shaped relationship [122].

Figure 1.4: the sex ratio by the day of natural insemination relative to the day of ovulation (day 0) [121]



The impact of the time of insemination relative to the time of ovulation is thought to influence the sex ratio by either a variation in the number of X- and Y-bearing spermatozoa at the site of fertilization or the ability of the oocytes to select a specific genotype of spermatozoa [123]. The adjustment of sex ratios in this way may be possible because the mucus viscosity and pH levels in the vaginal tract follow a similar U-shaped curve. James also claims that the hormone levels vary with a similar pattern and that these too can influence the sex of a child [95]. Guerrero found that the reverse pattern was seen for artificial insemination which supports this hypothesis since inseminations of an artificial nature are not subject to the conditions in the vaginal tract.

The relationship between the sex ratio and the time of insemination relative to the time of ovulation means that the coital rate has an indirect effect on the sex ratio too. Increasing the coital rate, increases the chance of having a successful insemination earlier on in the cycle and, therefore, increases the probability of having a boy [124]. In addition to this indirect link,

there is a more direct alternative. Debris from previous inseminations can reduce the penetrability of the mucus in the vaginal tract for subsequent inseminations. Therefore, increased frequency of coitus leads to a lower penetrability, which increases the likelihood of the success of a Y-bearing spermatozoon over an X-bearing one.

Natural conceptions that take longer to achieve have a higher sex ratio. Those conceptions that take over 20 months after the parent began trying had a sex ratio of 0.576 compared to 0.511 for those that took less time [125]. In fact, each extra year of trying to conceive is associated with a 4% higher probability of having a son. There are a number of reasons why this may occur. If a mother has problems with fertility then they are more likely to have more viscose mucus increasing the likelihood of having a boy. In addition to mucus viscosity, mothers can experience hormonal problems and poor follicular development which could influence the sex ratio. Finally, after multiple failed attempts at conception, a couple may increase the coital rate to improve their chances and in doing so increase the likelihood of having a son.

The time of year in which the baby is born can influence the sex ratio too, with a low sex ratio in the three months surrounding July at 0.473 and a significantly higher sex ratio in the three months surrounding October at 0.545 [126]. A similar result was obtained by Nonaka et al. [127]. In addition to the baby's time of birth influencing the sex ratio, the mother's time of birth has an impact. The sex ratio amongst mothers born between February and April is 0.503, which was significantly lower than that for the whole population at 0.512 [127]. However, no explanation has been provided for either of these phenomena.

Environmental Conditions

Lerchl found that there was a correlation between the sex ratio and the temperature in Germany between 1946 and 1995 with a 10 month lag indicating that the effect of the temperature takes place one month before conception [128]. Higher temperatures are associated with higher sex ratios. The 10 month lag suggests that it could be the temperature of the testes that is important and that affects the dissociation of the X- and Y-bearing spermatozoa. However, this may not be such an important issue with improvements in thermal clothing. Additionally, the relationship may be indirect since it may relate to changes in the frequency of coitus based on the temperature which in turn affects the sex ratio.

Stress inducing events

Psychological stress is thought to lead to an increased likelihood of having children of the opposite sex [97]. Stress leads to secretion of the adrenocorticotrophic hormone (ACTH), which in men lowers the testosterone levels and in women increases testosterone giving these opposing effects. However, in the examples discussed below the sex ratio is not considered for males and females separately so it is difficult to assess this claim.

Natural or man-made disasters can influence the sex ratio when they disrupt the regular function of a community. The typical effect is that the sex ratio falls however in some cases the reverse has been noted. The reasoning behind the effect of stress induced by such events on the sex ratio is as yet unclear, although, many theories have been presented. These include increased spontaneous abortions, reduced spermatozoa motility, changes in spermatozoa concentration, altered spermatozoa morphology and lower coital rates. The decline in sex ratio with stress induced by adverse events coincides with the Trivers-Willard hypothesis.

The smog of Greater London that occurred between the 5th and the 9th of December 1952 was followed by a fall in the sex ratio to 0.431 for 5 days exactly 320 days after the event, whereas the sex ratio for the days surrounding this trough was 0.541 [129]. In the same study, a similar pattern was seen for a flood in Brisbane in 1966 that lowered the pH of the water. In this case the sex ratio dropped to 0.351 320 days later compared to the sex ratio of 0.540 in the days surrounding the time of this dip. The Kobe earthquake in 1995 also led to a subsequent decline in sex ratio from 0.516 to 0.501, however, this drop was seen 280 days later and not 320 as seen in the case of the other two hazards [130]. This disparity in the lag of the decline in sex ratio may relate to the relatively short period over which the stress would have occurred in the earthquake compared to that of the smog and the flood.

Hansen et al. looked into the impact of the death or the admission to hospital for myocardial infarction or cancer of either a partner or an older child on the sex ratio of offspring [131]. Those exposed to one of these devastating events had a sex ratio of 0.490 which was significantly lower than that of the unexposed group at 0.512. They also found that the closer the event was to conception the lower the sex ratio was.

On September the 11th, 2001 there was a terrorist attack on New York inducing anxiety and bereavement across the whole of the United States [132]. The result was that the sex ratio fell in states as far away as California [132] and in New York the sex ratio dropped to 'it's lowest level' in the January following 9/11 [133].

Financial stress induced by the collapse of an economy has been found to lead to reduced sex ratios. The East German collapse in 1991 coincided with a sex ratio of 0.511 in East Germany, which is lower than the expected sex ratio of 0.514 calculated from historical figures and that of the more economically stable West Germany [134]. In Sweden a study on 129 years of economic and gender data suggested that a 1% increase in the annual private consumption was associated with the birth of approximately 25 more male births in Sweden than expected from history [135]. The collapse of an economy may mean that fewer goods and services than are needed or desired are available and this in turn may induce stress similar to that experienced in natural and man-made disasters.

In times of war the population sex ratio typically increases, for example, in Belgium, France, the UK and Germany the sex ratios were higher around the World Wars [136]. The suggestion that this phenomenon is merely an adjustment to rebalance the sex ratio after the increased male mortality during wartime is not a sufficient explanation. More plausible explanations are that the rises in the sex ratio occur because the stresses of wartime are counteracted by the promise of relief from adversity, the population is united or the adverse event is more anticipated, therefore, the stress may not be as prominent. Additionally, the coital rate is increased when soldiers come home on leave which can increase the probability of having male offspring. In contrast, the sex ratio in Slovenia after a 10 day war decreased from 0.518 to 0.504 [137, 138]. This conflicting example could be the result of the war being too short to influence the coital rate.

Nutrition

Nutrition is an important factor in the health and well-being of the offspring and the mother since it influences the energy available to the mother. Therefore, it may also be a factor in explaining changes in sex ratios. Cagnacci et al. found that women with a low pregnancy weight had a sex ratio of 0.497 which was significantly lower than the overall sex ratio of 0.515 [139].

Other studies based on the relationship between the sex ratio and nutrition focus on proxy measures, like BMI or caloric availability, which may not be good indicators of nutrition and as a result some struggle to get consistent results [140, 141]. However, Gibson et al. conclude that there is a relationship between the gender of offspring and the mid-upper arm muscle area (AMA) and found that those with a high AMA had a sex ratio of 0.600 whereas those with a low AMA had a sex ratio of 0.375, which was significantly lower [142].

Evidence that the nutritional intake of a mother is related to the sex ratio supports the Trivers-Willard hypothesis, with women of a poorer condition or lower nutrition having a preference towards female offspring. James discusses a study that revealed that changes in fat are associated with the oestrogen levels and that in mice a low fat diet produced more daughters, this could be true of humans [95].

However, contrasting results have been found that suggest that nutrition does not influence sex ratios. In 1944-45 the Dutch Hunger Winter lead to acute famine in the Netherlands and Stein et al. believed that the extra effort required to raise a male meant that there was selective fertilization of the ovum or selective attrition of the conceptus to bias the sex ratio towards girls [143]. However, there was no significant reduction in the sex ratio for mothers who were exposed to the famine during conception, therefore, this hypothesis was not supported.

In addition to considering the weight of the mother, Cagnacci et al. investigated the impact of the weight gain of the mother during pregnancy [139]. The mothers who gained the most weight during their pregnancy had a sex ratio of 0.493. This lower sex ratio suggests that either high weight gain increases the frequency of male foetal loss or when a mother is pregnant with a female child she has a metabolic tendency to store more energy rather than using it for reproductive processes. The reverse causality at play in this situation makes it difficult to make inferences.

Diet may well influence the sex ratio but the complexity of assessing the nutritional intake of an individual means that it is very difficult to investigate this as a causal link. Although people have attempted to do so, the variables used to represent nutrition are not necessarily good indicators of who has a better diet and who does not. Further, poor nutrition can prevent a woman from ovulating so it may not be possible to identify an effect [97].

Parental drug use

Smoking is a commonly discussed factor in the deviation of sex ratios from their expected values. In Fukuda et al.'s study of sex ratios in relation to parental smoking habits the sex ratio was as low as 0.451 when both parents smoked which is significantly lower than that amongst offspring where neither parents smoke which was 0.548 [144]. However, the main focus is on paternal smoking since maternal smoking does not seem to influence the sex ratio [145, 146]. Viloria et al. reveal that the sex ratio for males who smoke more than 20 cigarettes a day is 0.270 compared to 0.583 for non-smokers, although, this is based on relatively small numbers [147]. There are reports that paternal smoking reduces spermatozoa quality, density, count and motility and that it also alters the morphology of the spermatozoa. The speculation is that these changes are more prominent in the Y-bearing spermatozoa enabling more X-bearing spermatozoa to reach and fertilize the egg, thus, decreasing the sex ratio amongst smoking fathers.

The use of some contraceptive methods when the contraceptive methods are unsuccessful could alter the sex ratio [148]. Oral contraceptives can alter the pH and other characteristics of the vaginal tract, thus, influencing the sex ratio and, although, not statistically significant there does appear to be an increase in the sex ratio with 0.543 in the oral contraceptive group when the national average was 0.515. The parents that used spermicides had a sex ratio of 0.507 which was lower than the national level, although again, this was not statistically significant. When using the rhythm method the sex ratio increased to 0.567 which could relate to the increased proportions of conceptions that occurred late in the cycle, which is thought to increase the probability of having a son. The use of the fertility drugs clomiphene, nematocide, dibromochloropropane and anaesthetic gases have also been linked to changes in the sex ratio [92]. Mothers who underwent hormonally induced ovulation had a higher proportion of daughters with a sex ratio of 0.463 compared to the expected sex ratio of 0.514 [97]. On the other hand, mothers who used IVF, where the selection of the most advanced blastocysts for embryonic implantation, have an increased sex ratio of 0.637 amongst singleton births. This increased sex ratio occurs because in general male embryos have a more rapid level of cell development up to the blastocyst stage so the selection process is biased towards males [149].

Family status and parental dominance

Norberg states that the decline in sex ratio that has occurred in the USA, Great Britain and Canada coincides with a rise in single-parent births [87]. This concept led to the formulation of the partnership hypothesis, which claims that fathers who share a shelter with a female make a larger parental investment therefore the parents are more able to support a male child, which requires a greater amount of energy. This hypothesis relates to some degree to the Trivers-Willard hypothesis. There is a significant difference in the sex ratios for different marital statuses of parents. The sex ratio for parents who are married is 0.514, for parents living together but not married it is 0.522 and the sex ratio is 0.499 for parents who are not living together. However, there is reverse causality at play here as the gender of the child may influence the stability of a relationship. Therefore, the focus should really be on the marital status at the time of conception and not at the time of the birth.

Whiting found that mothers in a polygamous marriage have a greater probability of having female offspring than monogamous mothers [150]. The sex ratio for polygamous mothers was 0.466 whereas it was 0.533 for monogamous mothers. The suggestion as to why this occurred was that polygamous mothers have a lower frequency of coitus and are more likely to have sex when their sexual desires are greatest which occurs at the time of ovulation. Insemination at the time of ovulation is associated with a greater probability of having a daughter. Monogamous females on the other hand are more likely to have sex at the demand of the husband.

The probability of having a child of a specific gender varies from parent to parent with some parents having a propensity to have boys and others for girls. For example, the sex ratio for the 5^{th} child when the 4 prior children are all boys is 0.542 whereas when the 4 prior children are all girls it is 0.478 [151]. In addition, the risk of having a son increases with increasing parity, for example, families of size 2 have a sex ratio of 0.514 compared to 0.519 for families of size 12 [97]. However, this may be a result of parents with one child of each sex being less likely to want any more children whereas other parents may continue to have children until they get the gender that they most desire. However, if this relationship between parity and sex ratio is true then the decline in sex ratio may be the product of the trend of smaller family sizes.

The dominance of the parents is also thought to be important in determining the sex ratio. The maternal dominance hypothesis suggests that dominant women are more likely to have a male child. Dominance is a personality trait related to high testosterone levels which has also been associated with a higher probability of having a male child. In addition, this maternal dominance hypothesis relates to the Trivers-Willard hypothesis since dominant women are thought to be in a good condition. Grant et al. found that the sex ratio for women entered into the Dictionary of New Zealand Biography (DNZB) was 0.586 which was significantly higher than the New Zealand national average of 0.514 [152]. The women entered into the DNZB were considered to be dominant women.

Astolfi et al. found that there was a significant relationship between the gender of the child and the age gap between parents amongst Italian births [153]. The sex ratio for parents where the father was 16 to 25 years older than the mother was 0.549 compared to the overall sex ratio of 0.515. They believe that the larger age gap between the parents, where the father is older than the mother, is indicative of a higher social status for the father. The high social status of the father is suggestive that they will have higher levels of testosterone and, therefore, will be more likely to have sons than other fathers. In corroboration with this, the sex ratio amongst European royal families is 0.578 compared to 0.514 for the rest of the population [97]. However, this may be a result of the increased desire for Royalty to have sons rather than daughters so they are more likely to continue having children until they get one.

Parental ethnicity

The paternal and not maternal race seems to be an influencing factor in explaining sex ratios. Even after adjustment for parental age, education, birth order and maternal marital status the relative sex ratio for White fathers to Black fathers is 1.026 [154]. In the same paper they also gave a relative sex ratio for White to Native American fathers as 1.022. The sex ratio amongst Black fathers is quoted as 0.507 compared to White fathers which is 0.514 [97]. Therefore, White fathers are more likely to have boys than Black or Native American fathers. The nature of this relationship between paternal race and the sex ratio is not clear.

1.4 Existing Spatial Statistical Methodology

The above discussions of the aetiology of congenital malformations and of the sex ratio highlights the absence of any concrete evidence of causal links for a number of possible risk factors especially that of environmental exposure to toxicants. Spatial statistical methods provide useful tools for identifying regions of higher risk of a given health outcome and exploring the aetiology.

One of the first examples of the use of spatial statistics in discovering risk factors for a disease was in John Snow's study of the spread of Cholera [155]. In 1849 London suffered the worst outbreak of the disease in its history. In a 250 yard radius of the Broad Street pump over 500 people were killed in a short space of time. Snow mapped the location of all deaths from Cholera in this outbreak. The map acted as an exploratory tool indicating that if the local population distribution was considered there were more cases closer to the pump with only scattered cases further away. Snow ascertained that all of the victims had consumed water from the Broad Street pump. The pump handle was removed in response and the number of cases dropped subsequently. Hence in 1849 Snow had become the first person to use mapping and statistics to successfully test the link between exposure to a contaminant and the incidence of a disease [156].

There are three main steps in spatial statistics that are considered as part of this thesis and these are:

1. mapping the risk of the health outcome in the study region to identify the location of regions of excess risk;
2. testing for spatial homogeneity of the risk of the outcome over the mapped region;
3. testing for the spatial association between the mapped risk of the health outcome and that of exposure to possible risk factors.

The literature on the methodologies that have been developed for each of these steps is extensive with numerous books and papers entirely devoted to spatial statistics [1-4, 156-173]. A brief and by no means exhaustive description of some of the main procedures used in each of the steps is outlined below.

1.4.1 Mapping Risk of the Health Outcome

Health outcomes are available in two formats: area level aggregated count data or as individual level point data. The most commonly obtained format is count data where the number of cases is given for each administrative region in the study population. Point data is such that each case has been assigned an exact geographical location. Both types of data have to be dealt with in different ways when mapping the health outcome of interest.

Count data

In maps displaying count data the typical approach is to present summary measures for each region, the boundaries of which must be defined *a priori*. The most common summary measure is the standardised event ratio (SER) [157]. The SER is calculated by dividing the observed number of cases by the expected number of cases. The SER's for areas with small population sizes have high variances and consequently spuriously high SER's can occur by chance. This problem can be overcome by Bayesian smoothing techniques.

There are several approaches to obtaining smoothed estimates of the SER's for each administrative unit these include a minimum mean squared argument [174] and hierarchical smoothing from a fully or empirical Bayesian perspective [157, 175]. In a fully Bayesian approach the prior model is selected based on previous believes whereas in an empirical Bayesian approach the prior model is determined from the data.

The use of administrative boundaries provides a natural hierarchical framework. Many spatial models have been proposed for the levels of this hierarchy [175]. The reduction in the variance of the estimates for the SER's is achieved using shrinkage estimators. A hierarchical model is adopted *a priori*, either empirically or from previous knowledge of the situation, and information from this hierarchical model is used to shrink the Bayes point estimates towards a value that is related to the distribution of all units contained within the hierarchical structure.

This Bayesian approach to disease mapping using count data is utilised in many papers and examples of crude and smoothed maps can be seen for cancer incidence [176-179], heart disease mortality [180], perinatal mortality [181] and neoplasms of the brain and central nervous system [182]. These smoothed aggregated maps of SER's have a number of problems

including ecological bias caused by aggregation and that the result of the smoothing depends on the choice of the prior hierarchical model which may be incorrectly selected [179]. The problem of ecological bias can be overcome by using point data as opposed to count data, although it should be noted that point data is not always available or totally reliable because of crude post coding for earlier years.

Point data

Kernel density estimation methods can be used to map the relative risk surface for a given health outcome in the study area [183, 184]. The relative risk contours are determined by dividing the density of observed cases by that of the expected cases with adjustment for known confounders. The advantages of mapping risk in this way compared to using Bayesian smoothing methods include the removal of ecological bias through the use of individual level data, no boundaries of regions need to be selected and no assumptions are made regarding the distribution of the relative risk values.

The kernel density estimation procedure is not adopted as frequently as Bayesian smoothing techniques, which is likely to be a result of the unavailability of the point data required to apply kernel density methodology. However, contours for mapping regions of excess relative risk for congenital malformations [10], cancer [185] and female births [186] have been mapped using kernel density methodology.

1.4.2 Testing for Spatial Homogeneity of the Health Outcome

After mapping the risk of the health outcome over the study area of interest, if there does seem to be regions of elevated risk then it is important to test for spatial homogeneity of risk. There are many different approaches for testing for spatial homogeneity of risk and these are briefly described in this subsection.

Traditional approaches

Pearson's chi-squared statistic can be considered using the expected and observed counts in each administrative boundary [187]. A high value for the test statistic suggests that there is spatial heterogeneity and the p-value can be determined using Monte Carlo simulation. The

Potthoff-Whittinghill test is more commonly used and is based on the premise that when testing for spatial homogeneity the problems can be treated as having k binomial samples of differing size and that if there is spatial homogeneity then the parameter p for all these binomial distributions is equal [188, 189]. Extensions of this approach consider multinomial and Poisson distributions. The Potthoff-Whittinghill test has been used in application to the geographical distribution of cancer [177, 178, 190, 191], neoplasms of the brain and central nervous system [182] and perinatal mortality [181].

Distance and adjacency methods

Spatial autocorrelation quantifies the linear relationship between values at points in space that are of a fixed distance apart, for example, a common situation in which autocorrelation is useful is when considering the correlation between the SER's in administrative boundaries that are close to each other. If the SER's are similar in neighbouring regions then the spatial autocorrelation of the SER's is high. There are a number of autocorrelation coefficients that have been suggested [165]. Moran's I measures the similarity between regions that are close together [192]. If I is close to zero then there is spatial independence and if it is close to 1 then there is clustering of cases of the health outcome. In contrast, if Geary's c is close to zero then there is clustering and if it is close to 1 then there is spatial independence [193]. There are many other discussions on spatial autocorrelation and numerous other test statistics proposed [194-199]. Autocorrelation coefficients have been considered in application to the distribution of LaCrosse encephalitis [200], perinatal mortality [181], neoplasms of the brain and central nervous system [182] and cancer mortality [201].

A distance based test for detecting spatial dependence is given by Whittemore et al. [202]. The mean distance between all possible pairs of cases is used as the test statistic and Monte Carlo simulations are used to perform the test. However, this test does not adequately adjust for the population density and improved, more powerful approaches are presented by Tango [203] and Song and Kulldorff [204].

Moving window tests

Tests for detecting clusters of cases of a health outcome are used as part of the surveillance of a study region for evidence of 'hotspots' of elevated risk of a given condition or disease. Areas over which the risk is thought to be elevated are then flagged for further investigation and these regions are referred to as clusters. Reviews of cluster analysis discussing the pros and cons of investigating clusters can be found in a number of papers [205-207].

A 'geographical analysis machine' method was proposed by Openshaw et al. [208]. The geographical analysis machine approach involves superimposing a regular grid across the study region and centring circles of different radii on the intersections of the grid lines. Specific circles are flagged if the p-value for that circle, based on the number of cases and population size within the administrative regions whose centroids lie within the circle, is sufficiently small. The chosen threshold for the p-value must be very small, typically 0.005, to overcome the problem of multiple comparisons. In addition to problems of multiple testing, each of the tests are dependent [157]. Many evolutions of the geographical analysis machine method are reviewed by Openshaw [209].

In Besag and Newell's method a circle is drawn centred on each case with a radius that is sufficiently large in size to include the k^{th} nearest neighbouring case [210]. The number of administrative regions within the circle required to include the k nearest cases is used to calculate the p-value. Cuzick and Edwards' approach involves randomly selecting controls from the population at risk [211]. The number of cases amongst the k^{th} nearest neighbours of both cases and controls for each case is used to calculate the p-value. In both of these tests the value of k must be chosen *a priori* and considering different values of k introduces problems of multiple comparisons. An evolution of Cuzick and Edwards' method is based on the number of cases within a specified distance of cases [212].

Spatial scan statistics attempt to overcome the problem of multiple comparisons by producing only one test statistic. These scan statistics are designed to scan across the study region to find the maximum number of events within a window defined *a priori* by some criteria. In Turnbull et al. the windows are set to contain a constant population and are centred on each area centroid [213]. The test statistic is then the maximum number of cases that can be contained within one of these windows. Monte Carlo simulation under the assumption that

there is a random distribution of cases is used to determine the p-value. Kulldorff et al. define each of the windows to be of sufficient size to contain fewer people than a pre-specified fraction of the population [214]. The test statistic is based on a maximum likelihood ratio, the distribution of which is determined using Monte Carlo simulation. Other spatial scan statistics have been considered [215, 216]. There are some examples of the use of these likelihood ratio spatial scan statistics that explore the spatial distribution of cancer cases [176, 217, 218].

Risk surface estimation

The above tests for clusters of cases assume the shape of the cluster *a priori* and there are numerous situations for which this pre-specified cluster shape will be problematic. For example, in studies of air pollution cases may cluster around a main road and would therefore form a line. Cluster analysis that assumed a circular cluster would not identify this as a cluster with a low p-value. A likelihood ratio based method of testing for arbitrarily shaped disease clusters has been proposed but it was found to overestimate the true cluster by a large amount [219]. Consequently an alternative method is required.

Kernel density methodology, as used for mapping the risk of a health outcome, enables the identification of a region over which the relative risk is high, suggesting that there is a cluster of cases in that region, without assuming the shape of the cluster. However, unlike the tests specifically designed for detecting clusters, current methodology does not enable the calculation of a global p-value for each cluster [10].

1.4.3 Testing for the Spatial Association between Exposure and the Health Outcome

If spatial heterogeneity has been detected in the risk of a given health outcome then the next step is to search for spatially heterogeneous risk factors that could explain the non-uniform distribution of the health outcome. A causal link can be explored by testing for a spatial relationship between the risk of the health outcome and mapped exposure to a risk factor. A number of testing procedures have been utilised for this purpose and they are discussed in this section.

Ecological studies

Ecological studies are broad scale geographical studies that are carried out at either an international or national level. Examples of ecological studies can be found in numerous papers dealing with comparisons between cancer and per capita food consumption [220]; cancer and environmental factors in different countries [221]; leukaemia and social class and proximity of residence to estuaries [222]; and antioxidant status and cancer mortality [223].

The main problem in ecological studies is that of ecological bias where exposure to the risk factor varies within the geographical units selected as do the confounding factors. This ecological fallacy can prevent the detection of true causal links. Consequently, there has been a shift from the use of ecological studies to those on a smaller geographical scale, referred to as small area studies [157]. Small area studies can be performed more frequently than before with the increasing availability of accurate high resolution data and the development of new methodologies for statistical analysis at the small area level. The need for small area studies is motivated by the need to consider localised environmental pollution effects. The remaining methodologies are based on working at the small area level or, in some cases, the individual level.

Traditional approach

The most basic approach is to define an unexposed and exposed region and either the odds ratio (case-control studies) or the relative risk (population-based studies) is calculated for the health outcome in the exposed group compared to the unexposed group. The odds ratios or relative risks are often adjusted for known confounding.

A frequently discussed problem is that of environmental exposure to a putative point source. If this problem is being considered then typically the exposed region is defined as the area within a specified distance of the source and the unexposed region includes either the remainder of the study region or is a region that is outside the exposed region but still within a given distance of the site [224]. This approach has been applied to the risk of congenital anomalies in relation to residential proximity to landfill sites where Elliott et al. compared those within 2km of the sites (exposed) to the remaining study population (unexposed) [68].

In contrast Dolk et al. compared those within 3km of the landfill sites (exposed) to those between 3 and 7km from the sites (unexposed) [64].

Other studies that use this method include studies of congenital malformations and landfill sites [65, 70]; cancer and high voltage power lines [225]; cancer and landfill sites [226]; cancer and nuclear installations [227]; renal disease and cadmium exposure [228]; mortality and cokeworks [229]; stroke mortality and main roads [230]; respiratory problems and factories [180]; respiratory problems and main roads [231-233]; and mesothelioma and dockyards [224]. More than two exposure regions can be considered, for example, 4 levels of exposure to background radiation are considered in a study of childhood cancers [234]. Other studies with more than two exposure regions include those exploring cancer in the vicinity of Sellafield [235].

Stone's test

Stone's test was developed specifically for assessing elevated risk that is a result of exposure to a point source [157]. The study region is divided up and the sub-regions ranked in order of increasing exposure usually defined by distance from a point source but can be determined by any exposure measure [236]. It is assumed that the number of cases in each sub-region follows a Poisson distribution with mean $\lambda_i E_i$ where λ_i is the relative risk for the health outcome and E_i is the expected number of cases for sub-region i for $i = 1,...,n$ under the null hypothesis. The unconditional version of the test has the following hypotheses: $H_0 : \lambda_1 = \lambda_2 = ... = \lambda_n = 1$ and $H_1 : \lambda_1 \geq \lambda_2 \geq ... \geq \lambda_n$. One proposed test statistic is the maximum likelihood ratio where the maximum likelihood estimator is given by a minimax formula.

A couple of extensions of Stone's test have been presented. In the unconditional test the null hypothesis can be rejected when the study region has elevated or lowered risk and not necessarily when there is increased risk towards the centre of the site. Conditional hypotheses, where $H_0 : \lambda_1 = \lambda_2 = ... = \lambda_n = \rho$, overcome this problem [237]. Stone's method has also been adapted to enable adjustment for known confounders [238].

Stone's test has been applied to numerous situations including investigations into the relationships between cokeworks and mortality [229]; cokeworks and respiratory or cardiovascular disease [239]; cokeworks and birth outcomes [240]; coke ovens and cancer [241]; sewerage plants and cancer [242]; pesticide factories and cancer [243]; oil refineries and cancer [244]; petrochemical plants and cancer [245]; incinerators and cancer [246, 247]; radio transmitters or radio stations and cancer [190, 191, 248]; nuclear installations and cancer [249]; nuclear waste reprocessing and leukaemia [250]; and dockyards and mesothelioma [224]. The apparent popularity of Stone's test compared to other approaches presented in this section motivates the use of Stone's test as the standard test to which all new methods presented in this thesis for spatially comparing exposure to a risk factor and a health outcome are compared.

Score test

Score tests search for a trend in Poisson data based on a set of areas each with an associated exposure [251]. These tests can be used to test the null hypothesis of constant risk in all sub-regions against the alternative that the risk of the adverse health outcome is monotonically increasing with rising exposure to the risk factor. The test statistic is calculated using a weight for the associated exposure, the number of observed cases and the number of expected cases for each of the subgroups. This test statistic follows a chi-squared distribution under the null hypothesis.

Score tests have been utilised in studies of leukaemia and high powered radio stations [191]; leukaemia and hazardous waste sites [252]; and bronchitis mortality and reprocessing plants [253]. The tests were performed by weighting the difference between observed and expected cases with five different functions of distance. These functions include inverse distance, inverse square distance, inverse square root of distance, exponential decay and exponential decay with a threshold.

Linear risk scores

Linear risk score tests involve assigning each case a score relating to some measure of risk determined by some pre-specified alternative hypothesis [254]. The risk scores are summed for all the cases in the study region to give a total score. In a study of the risk of childhood

leukaemia and non-Hodgkin's lymphoma near nuclear installations the two risk scores considered are the reciprocal of the distance and the reciprocal of the distance rank of the ward [249]. Hence, in the latter situation a case in the ward closest to the nuclear installation will have a score of 1 whereas a case in the second closest ward would have a score of 0.5 and so on. The p-value is determined by simulation under the null hypothesis of no association between the health outcome and exposure.

Spatial point process modelling

A spatial point process is a stochastic process where the random locations of points in space are the main focus [255]. Diggle considers a Poisson point process model for spatial variation in the intensity of a relatively rare phenomenon to assess whether there is an association between that phenomenon and proximity of residence to a pre-specified point source [256]. The data is considered in the form of a set of events $\mathbf{x}_i$, representing the locations of all occurrences of the phenomenon of interest, in some region A. It is assumed that the data $\{\mathbf{x}_i \in A : i = 1,...,n\}$ are a partial realisation of an inhomogeneous spatial Poisson point process with intensity function $\lambda(\mathbf{x})$ which represents the mean number of events per unit area in the vicinity of $\mathbf{x}$. It is known that the intensity is subject to substantial spatial variation and it is suspected that the intensity is higher in the vicinity of the pre-specified point source, $\mathbf{x}_0$. A multiplicative form is assumed for $\lambda(\mathbf{x})$ such that $\lambda(\mathbf{x}) = \rho\lambda_0(\mathbf{x})f(\mathbf{x}-\mathbf{x}_0;\theta)$ where ρ is the number of events per unit area, $\lambda_0(\cdot)$ describes the spatial distribution of the intensity in the absence of association with proximity to the pre-specified point source and $f(\cdot)$ represents the change in intensity with distance from $\mathbf{x}_0$.

Diggle uses a second set of data $\{\mathbf{y}_i \in A : i = 1,...,n\}$, which is assumed to be a partial realisation of an inhomogeneous Poisson point process with intensity function $\lambda_0(\mathbf{x})$, to estimate $\lambda_0(\mathbf{x})$ via kernel density estimation constructed with a Gaussian kernel. The function $f(\cdot)$ is assumed, here, to be unimodal, with a maximum at $\mathbf{x} = \mathbf{0}$ and an exponential decay towards a constant value as $\mathbf{x}$ increases in distance from $\mathbf{0}$, although $f(\cdot)$ can be generalised. A likelihood ratio test statistic and Monte Carlo simulations are then used to test for an association between the phenomenon of interest and proximity to the pre-specified source.

The main problem with this methodology is that $f(\cdot)$ must be assigned a mathematical form which is unlikely to be realistic with topographical and meteorological factors influencing the concentrations of contaminants.

Poisson regression models

When dealing with area level or aggregated data a natural distribution for the observed number of cases is the Poisson distribution with mean equal to the relative risk multiplied by the expected number of cases for each area [157]. In this situation a generalised linear model given by

$$\log\theta_i = \alpha + \beta w_i + \mathbf{z}_i^T \boldsymbol{\gamma}$$

where θ_i is the relative risk, w_i is the measure of exposure and $\mathbf{z}_i$ is a vector of known risk factors for area i. The parameters α, β and $\boldsymbol{\gamma}$ can be estimated using either a maximum likelihood or Bayesian approach. Poisson regression models have been adopted in the analysis of Magnesium in drinking water supplies and mortality from acute myocardial infarction [257]; and of natural radiation and childhood leukaemia [258].

Logistic regression models

Case-control data can be analysed using logistic regression models of the form

$$\log\left(\frac{\pi_i}{1-\pi_i}\right) = \alpha + \beta w_i + \mathbf{z}_i^T \boldsymbol{\gamma}$$

where π_i is the probability of a case [157]. When dealing with matched case-control data a nuisance parameter must be used for each matching variable and in a situation where there are many matching variables conditional logistic regression methods should be adopted. Logistic regression has been used to relate cancer and smelter towns [259, 260]; cancer and air pollution [261, 262]; sudden infant death syndrome and air pollution [263]; cancer and power lines [225]; respiratory problems and air pollution [264, 265]; mortality and housing [266]; and birth outcomes and trihalomethane concentrations in water supplies [267].

1.5 Outline of this Thesis

The initial focus of this project was the increased concern about the risk of birth defects surrounding the Nant y Gwyddon landfill site. There are two commonly discussed types of parametric methodology utilised to investigate this type of problem. The first is the traditional approach of calculating the relative risk for an exposed region in relation to that of an unexposed region. The second is a parametric testing procedure for comparing exposure to potential emissions from the site to the rate of the health outcome as outlined by Stone. A non-parametric alternative that is commonly discussed is kernel density methodology where contours of relative risk are generated for the study region.

In Chapter 2 parametric methodology is discussed using the risk of congenital anomalies in the region surrounding the Nant y Gwyddon landfill site as an aid for describing, developing and assessing the procedures. Chapter 2 starts with methods of mapping the risk of the health outcome over the study region to identify regions of elevated risk. A method of assessing the homogeneity of risk over the mapped region has been developed using likelihood ratios. If spatially heterogeneous risk is detected then the traditional approach or Stone's test can be used to compare the risk of the health outcome to different levels of exposure to the risk factor, these tests are described in detail. The main test described by Stone is for use with data aggregated into groups determined by the level of exposure to the risk factor of interest. However, in addition a test for use with individual level data is presented but this individual level approach is not very powerful and consequently a new method of comparing the risk of the health outcome and exposure to a risk factor at an individual level is developed in Chapter 2 from likelihood ratios. Receiver operating characteristic (ROC) curves for all the testing procedures are examined to assess the relative merits of each of these parametric approaches.

In contrast Chapter 3 considers non-parametric approaches, again using the Nant y Gwyddon landfill site for illustrative purposes, beginning with methods of mapping relative risk contours for the health outcome using kernel density estimation. A global significance test is established for assessing the statistical significance of regions over which the relative risk surface of the health outcome is elevated. Finally, a procedure for testing the spatial association between the relative risk surface of the health outcome and mapped exposure is generated. The pros and cons of the parametric and non-parametric approaches are discussed at the end of Chapter 3 using ROC curves to assess the performance of the tests.

The methodologies described in Chapters 2 and 3 can be used to investigate any health outcome along with any possible risk factors. The lack of adequate exposure data for the Nant y Gwyddon landfill site does not mean that the methods cannot be utilised. Two alternative applications have been considered in Chapters 4 and 5 to illustrate the usefulness of the procedures developed. The first application of this new methodology is given in Chapter 4 and involves the examination of the risk of congenital anomalies in areas near the Nant y Gwyddon landfill site. There are three different aspects of this application that were considered and these are discussed below.

The first aspect of the investigation into congenital malformations concerns the method of adjusting for confounding factors when mapping the risk of congenital malformations. A logistic regression model was used to obtain predicted probabilities of a specific mother having a congenitally malformed offspring. These predicted probabilities were used as weights in the adjustment for possible confounding factors [10]. Until now, this approach to confounder adjustment has not been questioned. The methodology presented in this thesis was used to assess whether the method adopted for confounder adjustment adjusts for each of the confounding risk factors adequately.

The second aspect focuses on demonstrating how the new methodology can be used to compare the spatial distributions of two health outcomes. There are numerous different types of congenital anomalies and, if all congenital malformations are grouped together for analysis, the variation between the aetiologies of different birth defects can hinder investigations into causal links. Congenital anomalies are surprisingly rare events given the complexity of the development of an unborn child with approximately 2-3% of births affected [268]. Moreover, one specific type of congenital malformation can be extremely rare, for example, 1 in 1,000 births have a NTD and gastroschisis occurs in only 1 in 10,000 births [23]. In order to achieve higher statistical power most investigators have to group different types of birth defects for their study [37]. However, if the congenital malformations that are grouped together have contrasting aetiologies then the grouping can prevent real effects related to exposure to certain risk factors from being identified. Hence, the spatial pattern of chromosomal anomalies was compared to that of all other congenital malformations using the methodology developed in this thesis, in order to investigate the extent to which the aetiologies of congenital anomalies can vary.

The final aspect relates to the age of the data used to locate regions of excess relative risk. The regions of excess risk of congenital anomalies were identified for the period between 1983 and 1997 thus it is not clear whether elevated risk is still an issue near the Nant y Gwyddon landfill site. Therefore, the risk of congenital malformations was mapped using more recent data from 1998 to 2004 to determine whether elevated risk in the Nant y Gwyddon region is a current problem. However, changes in the reporting and collection of information on cases of congenital anomalies prevent direct comparisons between the two time periods.

The second application of the novel methodology concerns the secondary sex ratio and is presented in Chapter 5. The secondary sex ratio in Westernized countries is decreasing through time. The temporal pattern of the sex ratio in England and Wales is examined to see if this decline in sex ratio is apparent. Subsequently the temporal pattern in Cardiff and the Vale is also considered to assess whether this pattern is still evident at a local level. The reported decline in the sex ratio has prompted some investigators to hypothesise that gender is influenced by exposure to environmental contaminants.

If exposure to environmental contaminants does influence gender then it is anticipated that the sex ratio would exhibit a heterogeneous spatial variation because the spatial distribution of environmental contaminants is not uniform. One study has already been presented where regions over which there is an excess of females in the Falkirk District of Scotland have been mapped [186]. Similarly the spatial variation in sex ratios in Cardiff and the Vale of Glamorgan is considered in this thesis. However, the spatial analysis of sex ratios presented in Chapter 5 is taken further with an exploration of possible risk factors to try to explain any spatial variation.

Chapter Two: Parametric Methodology for Spatial Analysis

There are many different spatial statistical methods that can be used to investigate potentially spatially dependent health outcomes and risk factors. The key steps in the process are mapping risk, testing for spatial homogeneity and testing for spatial relationships between the health outcome and risk factors. In this Chapter the focus lies with parametric approaches to spatial analysis. Numerous parametric spatial statistical methodologies have been presented in the literature and some of these are discussed in Section 1.4. There are so many different methods available that only a small proportion could be considered in detail so only the most commonly discussed approaches are considered here. These numerous existing methodologies do not take full advantage of increasingly available individual level exposure and health outcome data and this deficiency motivated the development of the new methodologies presented alongside the existing methodology in this thesis.

The initial step in the spatial analysis, mapping risk, can be achieved parametrically by partitioning the study region into predefined sub-regions. This division of the study region is typically performed using administrative boundaries. A summary statistic for the health outcome is then calculated for each of the sub-regions and presented using a choropleth map. If there are differences in the summary statistic between sub-regions then there is a possibility that the health outcome is spatially heterogeneous. However, the patterns in the health outcome may have occurred as a result of random error which leads to the next step in the spatial analysis, testing for spatial homogeneity.

There are numerous parametric approaches to testing for spatial homogeneity including Pearson's chi-squared test [187], Potthoff-Whittinghill test [188, 189], distance and adjacency methods [165, 192-199, 202, 203, 269] and moving window tests [208-216, 270]. These tests for spatial homogeneity are discussed in slightly more detail in Section 1.4. Any of these testing procedures could be considered. However, the test considered in this Chapter is based on a likelihood ratio approach for assessing the spatial homogeneity of risk for when the study region is partitioned into k sub-regions. The reason for this choice is that the study region was already divided into sub-regions for the mapping stage of the spatial analysis. Additionally, this procedure is in keeping with the methodologies considered in this Chapter for exploring

the spatial association between exposure to a risk factor and the health outcome, which are all based on a likelihood ratio approach.

In both sections for mapping risk and testing for homogeneity, the risk of congenital anomalies around the Nant y Gwyddon landfill site is used for illustrative purposes. Birth outcomes, such as congenital malformations, provide good health outcomes for use in spatial analysis because of the shorter latency period relative to other health outcomes thus reducing the impact of population movements.

After testing for homogeneity, procedures are considered that investigate the possible spatial association between risk factors and the health outcome, the final step in the spatial analysis. There are many parametric approaches to testing for a spatial association between a health outcome and a risk factor. The approaches discussed in Section 1.4 include ecological studies [157], a traditional approach [157], Stone's test [236], score tests [251, 271], linear risk score tests [254], Poisson regression models [157] and logistic regression models [157]. The two most commonly used existing methodologies are considered in this Chapter, the traditional approach and Stone's test.

The traditional approach is based on partitioning the study region into two sub-regions: one that is referred to as the exposed region and the other as the unexposed region. In this Chapter the division of the study region is based on concentric circles specified by distance from a pre-specified source of environmental exposure and consequently is referred to as the concentric circle approach. The main test proposed by Stone involves partitioning the study region into more than two sub-regions. If the data is provided as area-level count data then there is no option but to use methodologies based on sub-regions such as the concentric circle approach or Stone's test. However, if the data is provided as individual level point data it may be that aggregating the data so that these procedures can be implemented does not allow efficient use of the individual level information that is provided.

Stone suggests a procedure for testing for a spatial relationship between exposure to risk factors and the health outcome that uses individual level point data. However, this method is not very powerful and consequently a new likelihood ratio method is developed that is a more powerful alternative. In the absence of relevant exposure data, simulations are performed to assess the efficiency of the proposed likelihood ratio test relative to that of Stone's test.

2.1 The Dataset

The risk of congenital malformations in the vicinity of landfill sites is used for illustrative purposes in this Chapter. The health outcome data was taken from a dataset that was created by carefully matching the Child Health System (CHS) and the National Congenital Anomaly Register (NCAR) at the Office of National Statistics between 1983 and 1997 for Wales [10, 272]. This amalgamation of data provides information on the location of the residence at birth (based on the postcode),.the date of birth, the gender of the birth, birth weight, gestational age, the age of the mother, the hospital of birth, the Townsend score for the father, whether or not the birth had at least one congenital anomaly and if so whether they had a chromosomal anomaly. The postcode for each birth was converted to east and north grid coordinates using the British National Grid reference system.

The CHS holds a register of all children born in Wales containing demographic, birth and administration details in addition to some maternal, immunisation and child health surveillance data. The CHS provided data on all births registered in Wales between 1983 and 1997, which acted as our denominator data in the development of the methodology presented in Chapters 2 and 3. The cases of congenital malformations or numerator data were identified by a match of the CHS information with that of NCAR.

NCAR contains data on live and still births with at least one congenital anomaly. The more serious congenital anomalies chosen for investigation by the Small Area Health Statistics Unit (SAHSU) in their papers include neural tube defects (ICD-9 740.0-740.2, 741.0-741.9, 742.0; ICD-10 Q00.0-Q00.2, Q05.-Q05.9, Q01.0-Q01.9); cardiovascular defects (ICD-9 745.0-747.9; ICD-10 Q20.0-Q28.9); abdominal wall defects (ICD-9 756.7; ICD-10 Q79.2-Q79.4); hypospadias and epispadias (ICD-9 752.6; ICD-10 Q54.0-Q54.9, Q64.0); surgical correction of hypospadias and epispadias (M731, M732); and surgical correction of gastroschisis and exomphalus (T281) [68]. Therefore, these are the congenital anomalies considered in this thesis.

In investigations of congenital anomalies one of the common risk factors considered is exposure to contaminants released from landfill sites. There are 32 landfill sites in Wales that have been identified by the Environment Agency that contained large volumes of waste, were in close proximity to residential populations and were opened or had a significant change in

use during the study period [272]. The location of each birth is shown in Figure 2.1 and is colour coded according to the distance from the nearest of these 32 specified landfill sites.

Figure 2.1: the location of all births in Wales colour coded according to the distance from the nearest of the 32 specified landfill sites 1983-1997

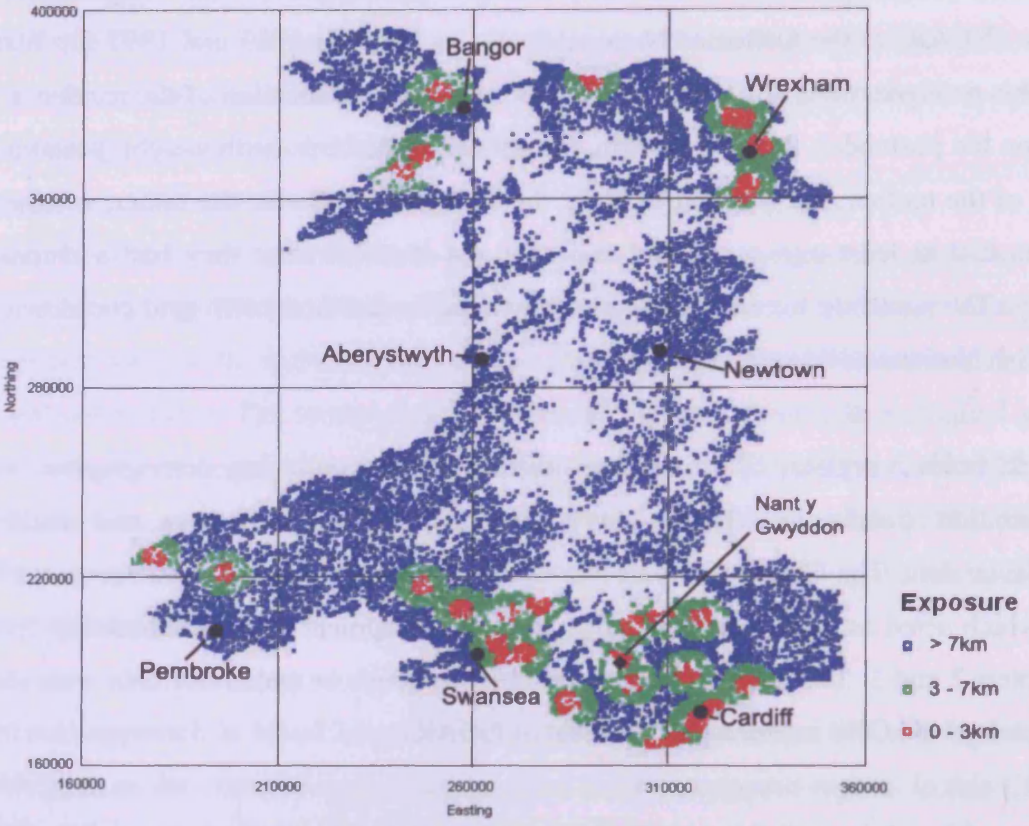
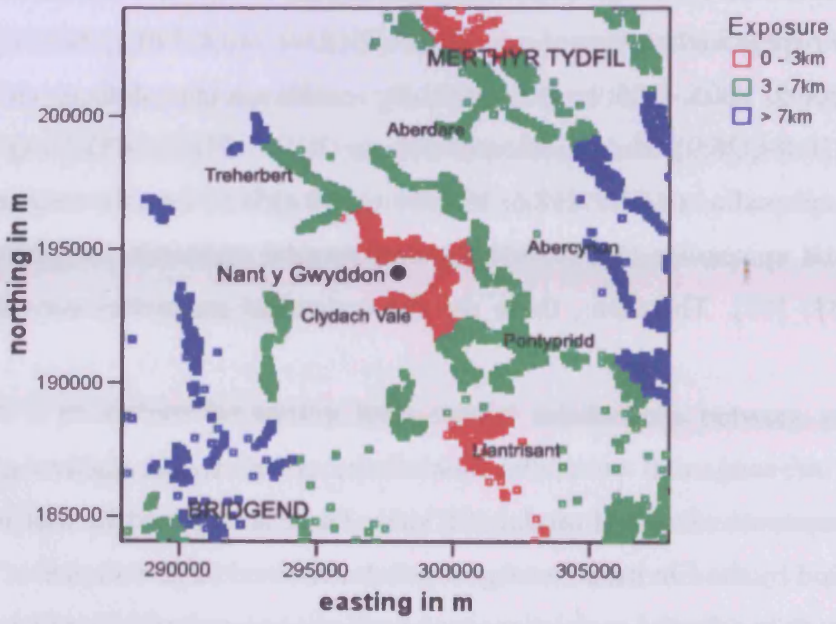


Figure 2.2: the location of all births in the 20km square region centred over Nant y Gwyddon colour coded according to the distance from the nearest landfill site 1983-1997



The study region selected for illustrative purposes in this Chapter is the 20km square region centred over the Nant y Gwyddon landfill site in the Rhondda valley as shown in Figure 2.2. The first consideration when investigating the risk of congenital anomalies in the 20km square region centred over the Nant y Gwyddon landfill site is the spatial distribution of risk. Mapping the risk of birth defects in the study region provides a visual approach to determining the spatial distribution of congenital anomalies and is discussed in the next section.

2.2 Mapping Risk

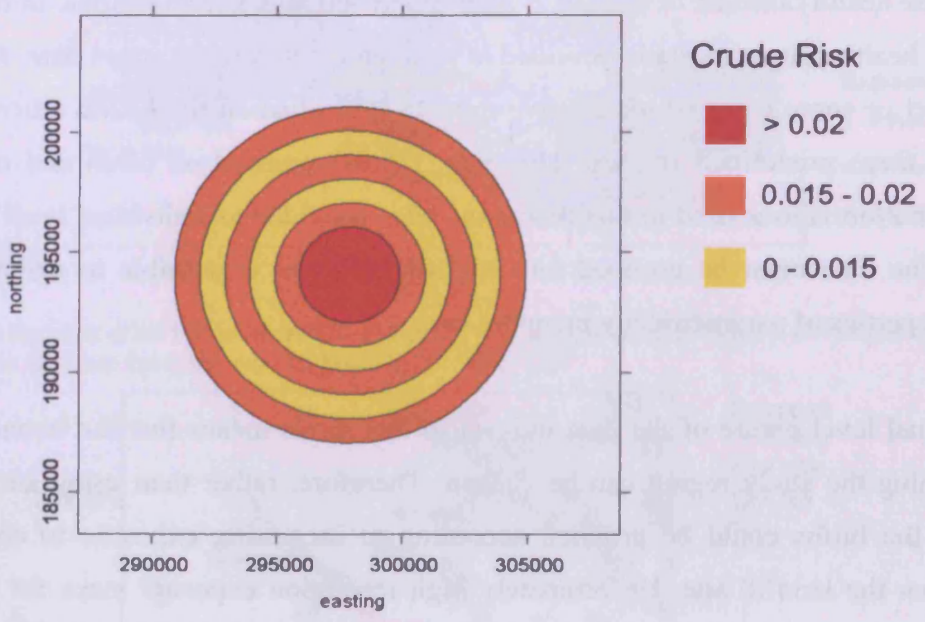
When mapping the risk of a health outcome, preserving the anonymity of individuals is very important in public health and as a result points representing cases of a medical condition should never be presented in fine detail so that it is possible to identify individuals. Therefore, the risk of the health outcome of interest is aggregated and mapped by regions. In the majority of cases the health outcome data is provided in an aggregated form as count data, for example at ward level or some other administrative unit. In this situation the health outcome can be mapped by these predefined regions. However, the information on cases and controls for congenital malformations used in this thesis has been provided as individual level point data. Therefore, the data must be grouped into regions before it is possible to demonstrate the different properties of parametrically mapping risk.

The individual level nature of the data utilised in this thesis means that the boundaries used for partitioning the study region can be chosen. Therefore, rather than using administrative boundaries the births could be grouped according to increasing exposure to contaminants released from the landfill site. Unfortunately high resolution exposure maps for the Nant y Gwyddon landfill site are not currently available and consequently a proxy for exposure must be used. In numerous studies of landfill sites, in the absence of any additional information, distance from the centre of the site has been used as a proxy for exposure [64, 65, 68, 70]. Therefore, for illustrative purposes exposure to the Nant y Gwyddon landfill site is assumed to decrease with increasing distance from the site and consequently concentric circles are used as the boundaries to partition the study region into sub-regions.

2.2.1 Crude Risk

The crude risk of congenital malformations is mapped in Figure 2.3, using regions of increasing distance from the centre of the site defined by concentric circles drawn at 2, 3, 4, 5, 6 and 7km radii centred over the Nant y Gwyddon landfill site. The crude risk of a congenital anomaly amongst the births within each region is calculated by dividing the number of congenital malformations by the total number of births in that region. The inner circle represents the region of greatest exposure to the landfill site and the level of exposure decreases with increasing distance. The east and north grid coordinates for each of the births are used with those for Nant y Gwyddon to calculate the distance of each birth from the centre of the landfill site. These distances can then be used to allocate each birth to the corresponding concentric circular sub-region.

Figure 2.3: map of the crude risk of congenital malformations around Nant y Gwyddon

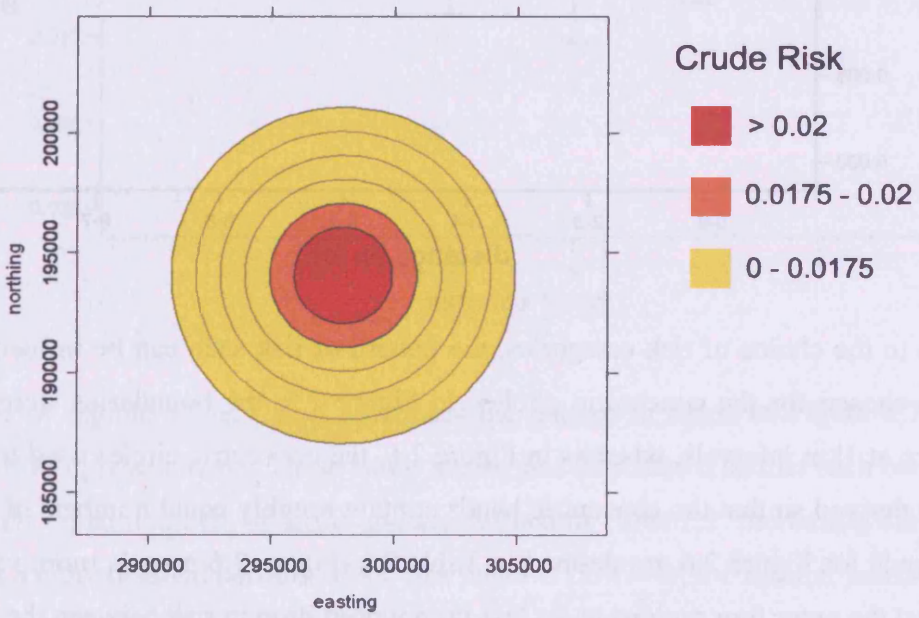


In Figure 2.3, easting is the east grid coordinate and northing is the north grid coordinate, this terminology is used throughout the thesis. The risk of a congenital anomaly seems to be highest within 2km of the centre of the site. However, those regions further than 2km from the centre of the site do not have a monotonically decreasing risk, as would be expected if proximity of residence to the landfill site influenced the risk of congenital malformations over such distances. The absence of an obvious trend of decreasing risk with increasing distance

from Nant y Gwyddon may be down to the chosen boundaries. Dolk et al. state that experts advise that the proximate zone within which the majority of the exposure to contaminants occurs is within 3km of a landfill site [64]. Therefore, it may have been better to have a larger number of concentric bands within 3km of the site. However, the difficulty of dividing the study region into more, smaller areas is that the population size within each region will become smaller and, thus, the crude risk will be more susceptible to sampling variation, which may conceal true patterns of risk.

In addition, the alternate orange and yellow bands that appear beyond 3km of the centre of the site occur as a result of only small fluctuations in the crude risks above and below 0.015, the chosen cut off point for the moderate and low risk categories. The selected risk values used to determine the risk categories are arbitrary and, in fact, replacing 0.015 with 0.0175 reveals a different risk pattern, see Figure 2.4, where the risk does appear to decrease monotonically with increasing distance from the centre of the site. Therefore, one of the problems with mapping risk in this way is that the chosen categories for risk can influence the picture obtained.

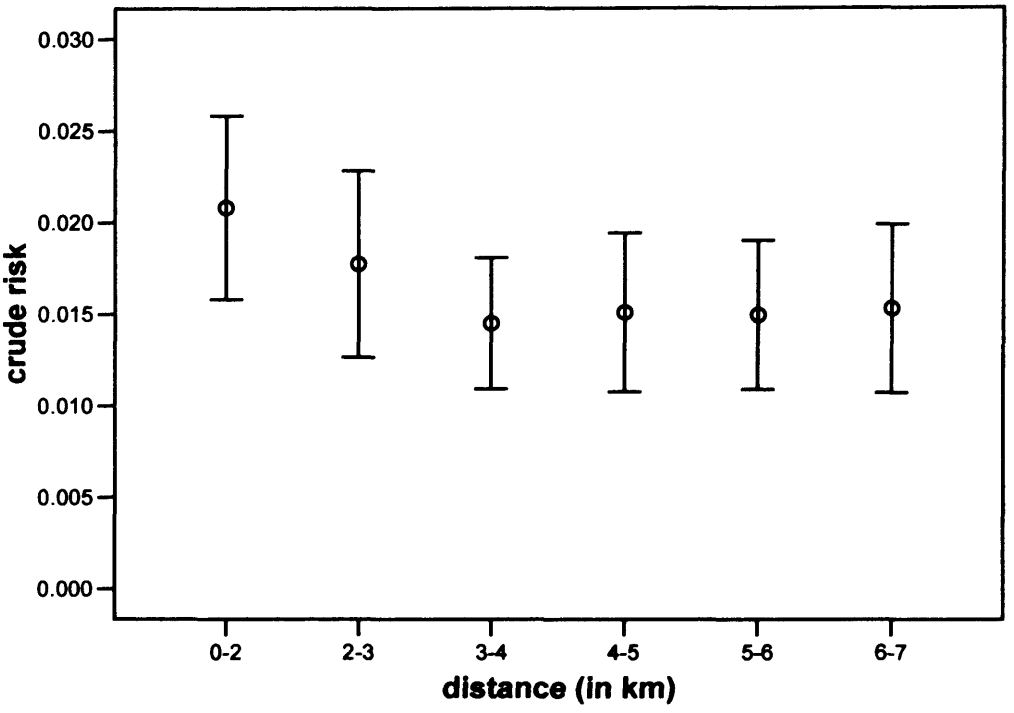
Figure 2.4: map of the crude risk of congenital malformations around Nant y Gwyddon



The problem of choosing appropriate risk categories for colouring can be overcome by simply plotting the risk of the health outcome against the distance from the source as risk categories

are no longer required. Figure 2.5 presents the risk of congenital malformations against the distance from the centre of the Nant y Gwyddon landfill site with 95% confidence intervals. This plot gives precise values for the crude risk of congenital anomalies in each concentric region. There are only relatively small changes in risk for the bands further than 3km from the site with all confidence intervals overlapping. These small changes are indicative of the landfill site having very little influence on these regions. However, the crude risk does increase towards the centre of the site within 3km and this may be the result of close proximity of residence to the landfill site increasing the risk of birth defects.

Figure 2.5: the crude risk of congenital malformations by distance from Nant y Gwyddon with 95% confidence intervals



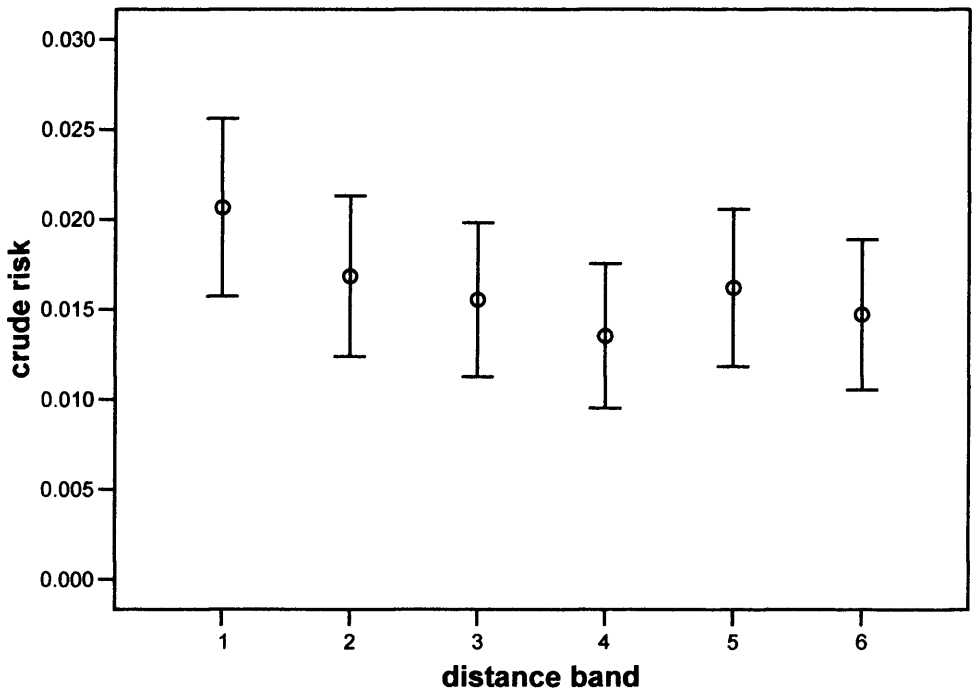
In addition to the choice of risk categories, the pattern of risk seen can be influenced by the boundaries chosen for the concentric circles. In Figure 2.5, the boundaries were drawn for convenience at 1km intervals, whereas in Figure 2.6, the concentric circles used to define the regions are derived so that the concentric bands contain roughly equal numbers of births. The distance bands for Figure 2.6 are defined in Table 2.1. Figure 2.6 reveals more variability in risk amongst the outer four regions and a less pronounced drop in risk between the second and third band out from the centre of the site. Therefore, the cut off points chosen for the boundaries of the concentric circles influence the pattern revealed in the risk of congenital malformations. It should be noted that the distance bands generated by forcing equal

population sizes do not hugely differ from those defined at 1km intervals and still differences in the results are apparent.

Table 2.1: definition of the distance bands in Figure 2.6

distance band	distance in metres
1	0 - 2024
2	2024 - 3178
3	3178 - 3836
4	3836 - 4830
5	4830 - 5890
6	5890 - 7000

Figure 2.6: the crude risk of congenital malformations by distance band from Nant y Gwyddon with 95% confidence intervals



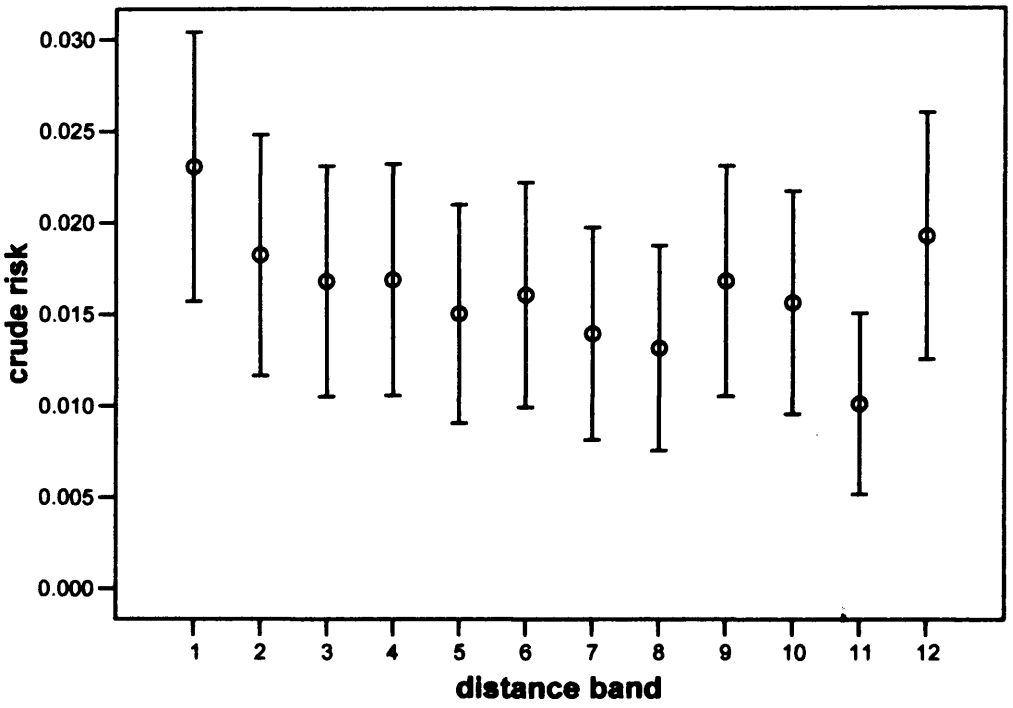
The number of regions chosen also influences the pattern in risks. In Figure 2.7, there are twelve bands as opposed to six. These bands are, again, determined so that each band contains approximately equal numbers of births and are defined in Table 2.2. Increasing the number of bands gives a more detailed account of the risk pattern. However, the smaller number of births in each of the regions increases the sampling variation in the crude risk values, which is reflected in the wider 95% confidence intervals shown in Figure 2.7 compared to those displayed in Figure 2.6. Therefore, the fluctuations in the risk between adjacent concentric bands may be the result of random variation and not of true effects. Hence, care has to be

taken when choosing an appropriate number of bands, unfortunately there is no standard way of calculating an optimal number of bands.

Table 2.2: definition of the distance bands in Figure 2.7

distance band	distance in metres
1	0-1339
2	1339 - 2024
3	2024 - 2584
4	2584 - 3178
5	3178 - 3479
6	3479 - 3836
7	3836 - 4477
8	4477 - 4830
9	4830 - 5598
10	5598 - 5890
11	5890 - 6437
12	6437 - 7000

Figure 2.7: the crude risk of congenital malformations by distance band from Nant y Gwyddon with 95% confidence intervals



Despite problems with selecting the number and location of concentric circles, all the maps and graphs above suggest that the risk of congenital malformations is higher for births to mothers who reside a shorter distance from the Nant y Gwyddon landfill site. However, by using these crude risks it is not possible to determine whether this increased risk of birth

defects is a result of the landfill site or some other factor that also varies spatially. For example, the increased risk may be a result of a higher deprivation level amongst those mothers who reside in close proximity to the Nant y Gwyddon site. Therefore, to rule out the influence of known confounders another measure of risk needs to be considered.

2.2.2 Adjusted Risk

The standardised event ratio (SER) is defined as

$$SER = \frac{y}{e}$$

where y is the observed and e the expected counts of cases [158]. The expected counts can be calculated using predicted probabilities from fitting a logistic regression model with explanatory variables defined by the known confounders so that the derived SER's are adjusted appropriately. Care must be taken when using SER's as it is possible to get large changes in SER estimates with only relatively small differences in expected counts if the expected count is small. This variation can to some extent be overcome by Bayesian smoothing methods which involve taking information from neighbouring regions to reduce the variation in the SER [157, 175]. The problem with this approach is that the smoothing process can result in several regions close to each other having similar SER's making it difficult to differentiate between the effects of smoothing and true spatial dependence in the health outcome.

All births in Wales within 3 to 7km of the centre of any of the 32 landfill sites identified by James et al. were used to fit a logistic regression model in SPSS with the following confounding factors: maternal age, socioeconomic class, year of birth, hospital of birth and gender [272]. Mothers within 3km of the centre of the site could have additional risk due to the close proximity of their residence in relation to the site therefore they were omitted from the logistic regression procedure. Those further than 7km from the site could be exposed to another site or they may not be representative of a control region, thus they were omitted from the model too. This exclusion of births from the fitted logistic regression model was also adopted by James et al.

The logistic model generated was then used to predict the probability of having a congenital anomaly for each birth within the 20km square area studied, based on the information on the

five risk factors. The sum of the estimated probabilities of all births in each region determined the expected number of congenital malformations in that area. If the only risk factors for congenital anomalies are those that have been considered in the adjustment then the SER should be equal to or very close to 1 in every region.

Figure 2.8: the SER of congenital malformations around Nant y Gwyddon by distance band defined in Table 2.1

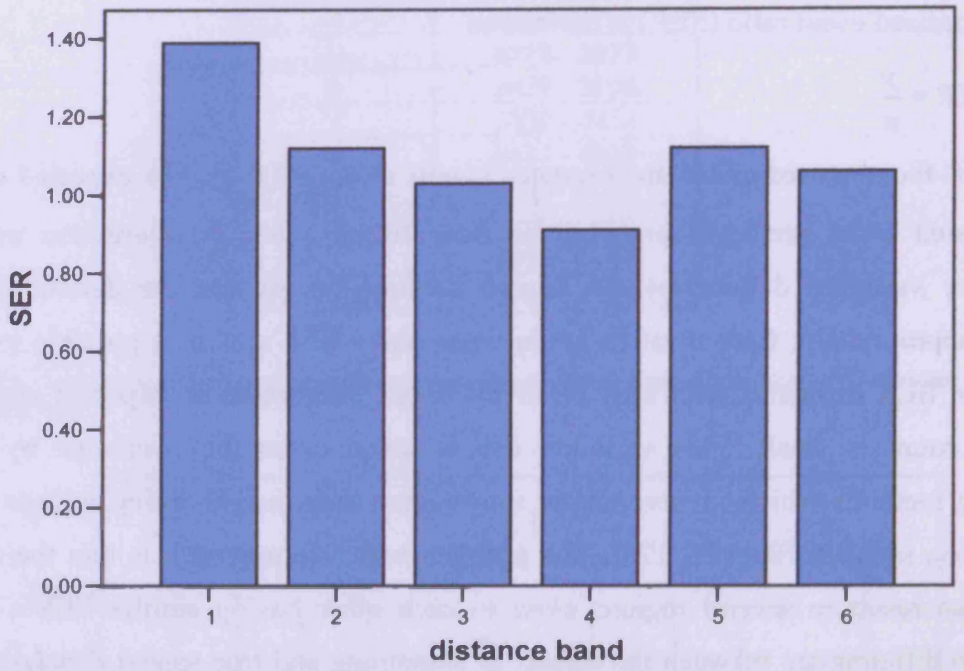


Table 2.3: the SER of congenital anomalies with 95% confidence intervals for the 6 distance bands around Nant y Gwyddon defined in Table 2.1

Distance band	SER (95% confidence interval)
1	1.39 (1.09, 1.77)
2	1.12 (0.86, 1.46)
3	1.03 (0.79, 1.37)
4	0.91 (0.67, 1.23)
5	1.12 (0.85, 1.46)
6	1.06 (0.79, 1.41)

Figure 2.8 is a plot of the SER's for congenital malformations around the Nant y Gwyddon landfill site, the 95% confidence intervals of which can be seen in Table 2.3. The pattern in the risk of a birth defect remains the same after adjustment for confounding with increased risk towards the centre of the site. This result suggests that the landfill site may influence the risk of birth defects, although the spatial patterns that have been seen here may be the result of random variation in the risk of birth defects, a situation investigated shortly.

Mapping the risk of congenital malformations around the Nant y Gwyddon landfill site revealed that the risk of birth defects varies spatially and identifies that the risk is elevated in regions of the closest proximity to the landfill site even after adjusting for known confounders. The next step in the spatial analysis is to formally test for homogeneous spatial patterns of risk.

2.3 Testing for Spatial Homogeneity of Risk

When considering the risk of congenital anomalies around the Nant y Gwyddon landfill site, after mapping the risk, it is necessary to determine whether or not the risk is spatially homogeneous. In this section the focus is on a situation where the study region has been partitioned into more than two exposure regions, say k areas. A test for spatial homogeneity of the risk of a health outcome assesses whether there is a statistically significant difference in the risk between the sub-regions without imposing ordering on the sub-regions. The test is based on a likelihood ratio approach and so a brief description of likelihood ratio methods is given below and, subsequently, the test for homogeneity is explained and an illustration of its application is given.

2.3.1 The Likelihood Ratio

The likelihood of a sample is the probability of that sample being observed [255]. Suppose $x_1, x_2, ..., x_n$ is a sample of observations of a random variable X then given independence the likelihood is the product of the probabilities of the individual values, i.e.

$$L = \prod_{i=1}^{n} P(X = x_i).$$

The probabilities of each of these individual values are determined by the distribution of the random variable X. The distribution selected is usually the most suitable standard statistical distribution and is specified by the use of an appropriate parameter set. The parameters are estimated by maximising the likelihood and are therefore referred to as the maximum likelihood estimators.

In a likelihood ratio test the aim is to determine which of the hypotheses, initially stated, is most likely to produce the sample observed. Therefore, the likelihood for the sample is calculated under both the null and the alternative hypotheses and compared.

The test statistic for the likelihood ratio test is

$$-2\ln\left(\frac{L_0}{L_1}\right)$$

where L_0 and L_1 are the maximised likelihood functions under the null and the alternative hypotheses, respectively. If the null hypothesis is true then the test statistic approximately follows a chi-squared distribution with m degrees of freedom, where m is equal to the number of parameters estimated in L_1 minus the number of parameters estimated in L_0.

2.3.2 The Likelihood based Spatial Homogeneity Test

A likelihood ratio test can be used to assess the homogeneity of risk. In order to achieve this, the region is split into k groups and suppose

$$X_{ij} = \begin{cases} 1 & \text{if individual j in group i is a case} \\ 0 & \text{otherwise.} \end{cases}$$

Suppose the baseline risk of the condition of interest for individual j in group i is b_{ij} and the additional risk attributable to being in group i is γ_i for $i = 1,...,k$.

The hypotheses are

$H_0 : \gamma_i = 0$ for all groups, i.e. the risk of the health outcome is spatially homogeneous

and

$H_1 : \gamma_i \neq 0$ for some of the groups, i.e. the risk of the health outcome is spatially heterogeneous.

The contribution to the likelihood from individual j in group i is determined by a Bernoulli distribution and is given by

$$(b_{ij} + \gamma_i)^{X_{ij}} (1 - (b_{ij} + \gamma_i))^{1 - X_{ij}}.$$

The contribution to the likelihood from all individuals in group i is then

$$\prod_{j=1}^{n_i} (b_{ij} + \gamma_i)^{X_{ij}} (1 - (b_{ij} + \gamma_i))^{1-X_{ij}} .$$

Hence, the full likelihood is

$$L = \prod_{i=1}^{k} \prod_{j=1}^{n_i} (b_{ij} + \gamma_i)^{X_{ij}} (1 - (b_{ij} + \gamma_i))^{1-X_{ij}} .$$

Under the null hypothesis this becomes

$$L_0 = \prod_{i=1}^{k} \prod_{j=1}^{n_i} b_{ij}^{X_{ij}} (1 - b_{ij})^{1-X_{ij}} .$$

However, under the alternative hypothesis estimates of the values for γ_i are required. The maximum likelihood estimator of γ_i can be found by differentiating the likelihood function with respect to γ_i and equating this to zero. Alternatively, this complex approach can be made easier by maximising the log likelihood instead, the solution of which is also the maximum likelihood estimator of γ_i since log is a monotonic function.

Therefore, the log likelihood is

$$l = \ln(L) = \sum_{i=1}^{k} \sum_{j=1}^{n_i} \{X_{ij} \ln(b_{ij} + \gamma_i) + (1 - X_{ij}) \ln(1 - (b_{ij} + \gamma_i))\}.$$

Differentiating with respect to γ_i and equating to zero gives

$$\frac{\partial l}{\partial \gamma_i} = \sum_{j=1}^{n_i} \left\{ \frac{X_{ij}}{b_{ij} + \hat{\gamma}_i} - \frac{1 - X_{ij}}{1 - (b_{ij} + \hat{\gamma}_i)} \right\} = 0 \qquad (2.1)$$

where $-\min(b_{ij}) < \hat{\gamma}_i < \min(1 - b_{ij})$ for all i and j.

The solution to the equation above is the maximum likelihood estimator of γ_i, denoted by $\hat{\gamma}_i$. However, it cannot be obtained in closed form algebraically hence numerical methods such as the bisection method must be used to determine the roots.

Under the alternative hypothesis the estimated likelihood is

$$L_1 = \prod_{i=1}^{k}\prod_{j=1}^{n_i}\left(b_{ij}+\hat{\gamma}_i\right)^{X_{ij}}\left(1-\left(b_{ij}+\hat{\gamma}_i\right)\right)^{1-X_{ij}}.$$

Hence,

$$\frac{L_0}{L_1} = \prod_{i=1}^{k}\prod_{j=1}^{n_i}\left(\frac{b_{ij}}{b_{ij}+\hat{\gamma}_i}\right)^{X_{ij}}\left(\frac{1-b_{ij}}{1-\left(b_{ij}+\hat{\gamma}_i\right)}\right)^{1-X_{ij}}.$$

It follows that the test statistic is

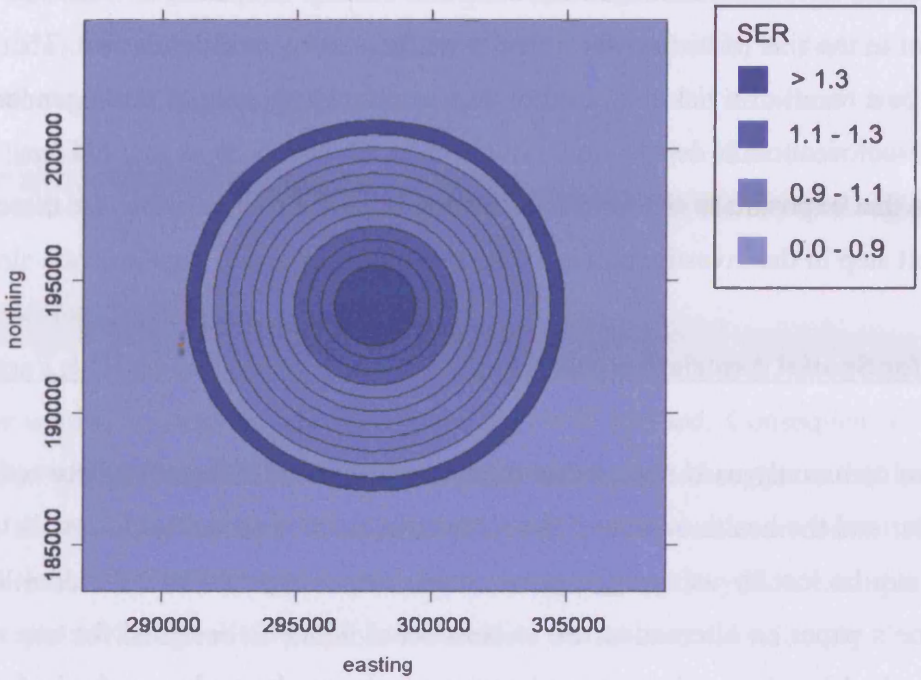
$$-2\ln\left(\frac{L_0}{L_1}\right) = -2\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left\{X_{ij}\ln\left(\frac{b_{ij}}{b_{ij}+\hat{\gamma}_i}\right)+\left(1-X_{ij}\right)\ln\left(\frac{1-b_{ij}}{1-\left(b_{ij}+\hat{\gamma}_i\right)}\right)\right\}$$

If the null hypothesis of spatial homogeneity is true then this test statistic follows a chi square distribution with k degrees of freedom. It should be noted that in contrast to the other likelihood ratio tests that are discussed in this Chapter the groups do not have to be in any specific order based on exposure to the risk factor of interest.

2.3.3 Applying the Likelihood based Spatial Homogeneity Test to Nant y Gwyddon

In order to demonstrate this test for spatial homogeneity of risk the Nant y Gwyddon landfill site was considered as before. The study region for Nant y Gwyddon was partitioned into 13 regions, as shown in Figure 2.9, where the pattern of SER's for congenital malformations is mapped. The concentric circles are drawn at 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5, 5.5, 6, 6.5 and 7km from the centre of the site. The spatial pattern of SER's seems to be heterogeneous around the Nant y Gwyddon landfill site. Surprisingly, the SER is high between 6.5 and 7km. This is unlikely to be related to the Nant y Gwyddon landfill site but could be the result of another landfill site or some other risk factor that has not yet been considered. If the distance band between 6.5 and 7km is discounted there seems to be a monotonically decreasing risk of birth defects with increased distance from the landfill site. Clearly, there is a spatial pattern for the risk of congenital malformations but it is unclear whether this pattern is real or just a result of random variation.

Figure 2.9: map of SER's for congenital malformations around Nant y Gwyddon



In order to determine whether the spatial pattern seen is random the likelihood based test for homogeneity was applied to these 13 regions. The baseline risk for birth j in group i (b_{ij}) was the predicted probability from the fitted logistic regression model for that birth. The log likelihood ratio was 26.17 which gave a p-value of 0.018 when compared to the chi-squared distribution with 13 degrees of freedom. This small p-value provides statistical evidence that the heterogeneous pattern of risk of congenital anomalies is unlikely to be the result of random error. Therefore, the pattern may be caused by a spatially heterogeneous risk factor that has not yet been considered.

One of the limitations of this methodology is that in some instances there may be no solution to equation 2.1, although, this has not occurred for the Nant y Gwyddon example. In this situation the likelihood function can be plotted against different values of γ_i to determine where the maximum is, thus, giving the appropriate value of $\hat{\gamma}_i$. In this situation the likelihood is usually maximised on a boundary or these groups may be merged with other groups so that the root is not less than the minimum baseline risk and, supposing this leaves m groups, then the log likelihood ratio is compared to the chi-squared distribution with m degrees of freedom to obtain the p-value.

Applying the test for spatial homogeneity of risk to Nant y Gwyddon reveals that there is a spatial pattern in the risk of birth defects that is unlikely to be a random event. Therefore, the pattern may be a result of a risk factor other than maternal age, year of birth, gender, hospital of birth and socioeconomic deprivation, which have already been accounted for. The main hypothesis is that exposure to the landfill site may, in part, be responsible for these patterns. Thus, the next step in the investigation is to test this hypothesis.

2.4 Testing for Spatial Association between the Health Outcome and a Risk Factor

There are two commonly used approaches to testing for spatial association between exposure to a risk factor and the health outcome: the concentric circle approach and Stone's test. Some information can be lost by using aggregated count data as opposed to individual level point data. In Stone's paper an alternative test is considered which is designed for use with point data. This method has its weaknesses and consequently an alternative method of testing for spatial association between exposure to a risk factor and the health outcome is developed in this Section that makes better use of the extra information that point data provides.

2.4.1 Existing Methodology

Concentric circle approach

The concentric circle approach involves dividing the study region into two areas: an exposed region and an unexposed reference region. These two regions are defined by distance from the assumed point source. The relative risk is then calculated for congenital malformations amongst mothers who reside in the exposed region relative to those residing in the unexposed region. A relative risk that is greater than one suggests that the risk of birth defects is higher in the exposed region. If, additionally, the confidence interval does not include one then the elevation of risk in the exposed region is statistically significant at the appropriate level.

Elliott et al. defined the exposed region as within 2km of any of the 19,196 landfill sites considered and the unexposed region as any region that was further than 2km from all of these sites [68]. The difficulty with this approach is that some parts of the unexposed region could be a long distance from any of the exposed regions and, therefore, are unlikely to provide a

sensible comparison. On the other hand, Dolk et al. defined the exposed region as within 3km of the 21 landfill sites identified and the unexposed as between 3 and 7km from any of these sites [64]. This definition of exposed and unexposed areas is better, as the choice of the 3km radius for the exposed region ensures that any areas that may be influenced by contaminants from the landfill site are included. In addition, the use of the 7km limit for the reference region ensures that the unexposed area is not too far away from the exposed region to prevent a reasonable comparison.

The definitions of exposed and unexposed regions used by Dolk et al. have been adopted for use in this section to demonstrate the concentric circle method. Consequently, the exposed region is that within 3km of Nant y Gwyddon and the unexposed region is that between 3 and 7km from the centre of the site. However, Dolk et al.'s study is a case-control study so odds ratios had to be used instead of relative risks. In the Nant y Gwyddon example, data on all births is available therefore it is possible to calculate the relative risks directly.

The relative risk for congenital malformations for mothers residing within 3km of the centre of the Nant y Gwyddon landfill site is 1.303 with a 95% confidence interval of (1.031, 1.646). This result suggests that there is 95% confidence that those mothers who reside within 3km of the landfill site are between 3.1% and 64.6% more likely to have a congenitally malformed birth than those who reside beyond 3km from the site. Therefore, there is statistically significant evidence at the 5% level that those living in close proximity to the Nant y Gwyddon landfill site are at a greater risk than those who live further away. However, it should be noted that the lower limit of the confidence interval is close to 1 therefore caution must be exercised.

One of the main problems with the concentric circle approach is that the choice of exposure and reference region influences the result. For example, if the exposed region is considered to be within 2km rather than 3km of the centre of the site then the relative risk is in fact 1.355 (95% CI: 1.028-1.785) which is slightly different to the relative risk determined for 3km. The disparity is even more noticeable when the reference region is altered. For example, if the reference region is all mothers in the 20km square region centred over the site who reside over 3km from the Nant y Gwyddon landfill site then the relative risk for the exposed living within 3km of the site is 1.774 (95% CI: 1.434-2.194). Therefore, it is clear that the choice of the two regions influences the resultant relative risk however there is no right or wrong answer

regarding the region definition used thus highlighting the fundamental problem of working with aggregated data defined by arbitrary sub-regions.

The difficulty with using the crude relative risk is that there could be a large degree of confounding. For example, it may be that mothers living close to the landfill site are in fact subject to a greater degree of deprivation than the remainder of the study population which could cause an increased risk of congenital malformations in the exposed region. Thus it is not clear whether it is deprivation or exposure to contaminants released from the landfill site that has caused the elevation in risk of birth defects in the exposed region. Therefore, although, it is not possible to rule out all possible confounding effects the risk factors that are known can be adjusted for when determining the relative risk. This adjustment rules out any possible effects each risk factor has on the relative risk and consequently increasing the likelihood that the effects seen are a result of exposure to contaminants released from the landfill site.

When adjusting for confounding factors the expected number of cases is determined by summing the predicted probabilities from the logistic regression model fitted with all known risk factors rather than the total number of births multiplied by the overall risk of congenital malformations. The relative risk is then calculated as before by dividing the observed number of cases by the expected number of cases which is equivalent to calculating the SER. The SER of congenital malformations for births to mothers living within 3km of the Nant y Gwyddon landfill site is 1.299 with a 95% confidence interval of (1.069, 1.564). The confidence intervals were calculated using the methodology described in Palmer et al.'s paper [73]. Therefore, the risk of congenital anomalies is still statistically significantly higher within 3km of the centre of the site after adjustment has been made for known confounders.

If there is no putative point source then the study region can be defined by another method other than using distance. For example, if there was a detailed exposure map for a given contaminant then a threshold can be selected such that regions with concentrations above that threshold are considered to be exposed and the remainder of the study region unexposed. However, if a detailed map of exposure is available then a lot of information is lost by merely partitioning the study region into two exposure groups.

Moreover, one of the main problems with the concentric circle approach is the arbitrary nature of defining the exposed and unexposed regions. In addition to this hurdle, the concentric circle

approach is overly simplistic. The use of only two regions means that spatial variation of risk on a smaller scale may be overlooked therefore the pattern may be thought to be uniform when in fact the risk is spatially heterogeneous. Furthermore, the calculation of the risk of congenital malformations in the exposed region may give some indication of the risk associated with proximity of residence to the landfill site but it does not give a detailed dose-response effect.

These difficulties that arise from the simplicity of the concentric circle approach can be overcome to some extent by adopting a more sophisticated method where the study region is partitioned into more sub-regions of varying degrees of exposure or by investigating the health outcome and the risk factor at an individual level. When using more than two regions the problem of arbitrary selection of boundaries is still present but it is eradicated by methods that use individual level data. However, individual level data is not always available, in which case sub-regions must be utilised. In the subsequent sub-sections methodology is described for dealing with more than two exposure regions and, ultimately, with individual level data.

Stone's test

Stone's test was developed to assess the influence of exposure to a putative point source on the risk of a given health outcome [236, 237]. In this test the study region is partitioned into k groups of increasing exposure to the risk factor. A likelihood ratio approach is then considered to test whether the risk of the health outcome increases with increased exposure. This differs from the likelihood based test for spatial homogeneity because the test for spatial homogeneity is searching for statistically significant differences in risk between regions regardless of exposure to the risk factor.

Suppose that

$$X_{ij} = \begin{cases} 1 & \text{if individual j in group i is a case} \\ 0 & \text{otherwise.} \end{cases}$$

Although Morton-Jones et al. describe a version of Stone's test with confounder adjustment, where the baseline risk of the health outcome is considered to be variable, the baseline risk is assumed to be constant here for simplicity [238]. Hence, the baseline risk for all patients j is

denoted by b. The additional risk attributable to being in group i is γ_i for $i = 1,...,k$. Thus the modelled risk is $b + \gamma_i$ for individual j in group i.

Consequently the hypotheses are

$H_0 : \gamma_i = \gamma_j$ $\forall$ i and j, i.e. the risk of the health outcome is the same in all groups

$H_1 : \gamma_1 > \gamma_2 > ... > \gamma_k > 0$, i.e. the risk of the health outcome rises with increased exposure.

The hypotheses can be simplified by the following substitution

$$\begin{aligned}
\delta_k &= \gamma_k \\
\delta_{k-1} &= \gamma_{k-1} - \gamma_k \\
&\vdots \quad \vdots \\
\delta_1 &= \gamma_1 - \gamma_2 .
\end{aligned}$$

so

$$\begin{aligned}
\gamma_k &= \delta_k \\
\gamma_{k-1} &= \delta_k + \delta_{k-1} \\
&\vdots \quad \vdots \\
\gamma_1 &= \delta_k + \delta_{k-1} + ... + \delta_1 .
\end{aligned}$$

This means that the hypotheses are equivalent to

$H_0 : \delta_i = 0$ $\forall$ i

and

$H_1 : \delta_i > 0$ $\forall$ i .

This is far more convenient mathematically as the domain for each δ_i is independent of all others.

The contribution to the likelihood from each patient is determined by the Bernoulli distribution as it was in the test for spatial homogeneity, therefore, it follows that the full likelihood in this situation is

$$L = \prod_{i=1}^{k} \prod_{j=1}^{n_i} (b + \gamma_i)^{X_{ij}} (1 - (b + \gamma_i))^{1 - X_{ij}} .$$

Hence the log likelihood function is

$$l = \ln L = \sum_{i=1}^{k} \{X_{i.} \ln(b+\gamma_i) + (n_i - X_{i.}) \ln(1-(b+\gamma_i))\}$$

where $X_{i.} = \sum_{\forall j} X_{ij}$.

Substituting in δ_i gives

$$\begin{aligned} l = {} & X_{k.} \ln(b+\delta_k) + (n_k - X_{k.}) \ln(1-(b+\delta_k)) \\ & + X_{(k-1).} \ln\left(b + \sum_{i=k-1}^{k} \delta_i\right) + (n_{(k-1)} - X_{(k-1).}) \ln\left(1 - \left(b + \sum_{i=k-1}^{k} \delta_i\right)\right) \\ & + \ldots + X_{1.} \ln\left(b + \sum_{i=1}^{k} \delta_i\right) + (n_1 - X_{1.}) \ln\left(1 - \left(b + \sum_{i=1}^{k} \delta_i\right)\right) \end{aligned}$$

The maximum likelihood estimator for γ_i can be obtained by differentiating with respect to δ_i, equating to zero and solving in terms of either δ_i or γ_i so

$$\frac{\partial l}{\partial \delta_1} = \frac{X_{1.}}{b + \sum_{i=1}^{k} \delta_i} + \frac{n_1 - X_{1.}}{1 - \left(b + \sum_{i=1}^{k} \delta_i\right)} = 0 \qquad (2.2)$$

$$\begin{aligned} \frac{\partial l}{\partial \delta_2} = {} & \frac{X_{2.}}{b + \sum_{i=2}^{k} \delta_i} + \frac{n_2 - X_{2.}}{1 - \left(b + \sum_{i=2}^{k} \delta_i\right)} \\ & + \frac{X_{1.}}{b + \sum_{i=1}^{k} \delta_i} + \frac{n_1 - X_{1.}}{1 - \left(b + \sum_{i=1}^{k} \delta_i\right)} \\ = {} & \frac{X_{2.}}{b + \sum_{i=2}^{k} \delta_i} + \frac{n_2 - X_{2.}}{1 - \left(b + \sum_{i=2}^{k} \delta_i\right)} = 0 \quad \text{by equation 2.2.} \end{aligned}$$

Thus the equations to solve become

$$\frac{X_{1\cdot}}{b+\sum_{i=1}^{k}\delta_i} - \frac{n_1 - X_{1\cdot}}{1-\left(b+\sum_{i=1}^{k}\delta_i\right)} = 0$$
$$\frac{X_{2\cdot}}{b+\sum_{i=2}^{k}\delta_i} - \frac{n_2 - X_{2\cdot}}{1-\left(b+\sum_{i=2}^{k}\delta_i\right)} = 0$$
$$\vdots \qquad\qquad \vdots$$
$$\frac{X_{k\cdot}}{b+\delta_k} - \frac{n_k - X_{k\cdot}}{1-(b+\delta_k)} = 0.$$

In terms of γ_i these are

$$\frac{X_{1\cdot}}{b+\gamma_1} - \frac{n_1 - X_{1\cdot}}{1-(b+\gamma_1)} = 0$$
$$\frac{X_{2\cdot}}{b+\gamma_2} - \frac{n_2 - X_{2\cdot}}{1-(b+\gamma_2)} = 0$$
$$\vdots \qquad\qquad \vdots$$
$$\frac{X_{k\cdot}}{b+\gamma_k} - \frac{n_k - X_{k\cdot}}{1-(b+\gamma_k)} = 0.$$

Solving the above equations gives $b+\hat{\gamma}_i = \overline{X}_i$ and $\hat{\delta}_j = \hat{\gamma}_j - \hat{\gamma}_{j+1} = \overline{X}_j - \overline{X}_{(j+1)}$ where $\overline{X}_i$ is the mean of X_{ij} in group i.

The maximum likelihood estimator of δ_i cannot take negative values, therefore, if the data to which this likelihood method is applied results in negative values for any of the δ_i then the i^{th} group must be combined with a neighbouring group to recalculate δ_i. This process of recalculating the δ_i's is repeated until there are no negative maximum likelihood estimators for δ_i.

Hence, if there are k groups and the δ_i have been recalculated if necessary to ensure that $\delta_i > 0$ for all i, then the test statistic is

$$T_b = -2\ln\left(\frac{L_0}{L_1}\right) = -2\sum_{i=1}^{k}\left\{X_{i\cdot}\ln\left(\frac{b}{b+\hat{\gamma}_i}\right) + (n_i - X_{i\cdot})\ln\left(\frac{1-b}{1-(b+\hat{\gamma}_i)}\right)\right\}.$$

The binomial assumption employed in the testing procedure above means that the calculation of the likelihood function is computationally difficult particularly when there are non-constant baseline risks. In contrast in Stone's paper a Poisson approximation for the binomial distribution is adopted. The binomial distribution is a more reasonable assumption to make as the population size is fixed, therefore, there can only be a certain number of cases, whereas a Poisson distribution has an infinite tail, suggesting that any number of cases can occur. However, when dealing with rare health outcomes and a large population size it is reasonable to make a Poisson assumption.

Therefore, for rare health outcomes and a large population size the hypotheses are equivalent to

$$H_0 : O_i \sim Poisson(E_i)$$

$$H_1 : O_i \sim Poisson(\lambda_i E_i)$$

where O_i and E_i are the observed and expected counts of cases in group i, respectively, and λ_i is the relative risk for group i, which are such that $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_k$. It should be noted that Stone expresses the additional risk in a multiplicative form $\lambda_i E_i$, a form that eases the algebra.

The likelihood is the probability of observing $O_1, O_2, ..., O_k$ hence under the null hypothesis the likelihood is

$$L_0 = \exp\left\{-\sum_{i=1}^{k} E_i\right\} \prod_{i=1}^{k} \frac{E_i^{O_i}}{O_i!}.$$

Under the alternative hypothesis the likelihood function is

$$L_1 = \exp\left\{-\sum_{i=1}^{k} \lambda_i E_i\right\} \prod_{i=1}^{k} \frac{(\lambda_i E_i)^{O_i}}{O_i!}.$$

Taking logs gives

$$l_0 = \ln L_0 = -\sum_{i=1}^{k} E_i + \sum_{i=1}^{k} O_i \ln(E_i) - \sum_{i=1}^{k} \ln(O_i!)$$

$$l_1 = \ln L_1 = -\sum_{i=1}^{k} \lambda_i E_i + \sum_{i=1}^{k} O_i \ln(\lambda_i E_i) - \sum_{i=1}^{k} \ln(O_i!)$$

If the constraint that $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_k$ is ignored then the maximum likelihood estimator for λ_i is the solution to

$$\frac{\partial l_1}{\partial \lambda_i} = -E_i + \frac{O_i}{\lambda_i} = 0.$$

Consequently the maximum likelihood estimator for λ_i is

$$\hat{\lambda}_i = \frac{O_i}{E_i}.$$

In order to ensure that the condition $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_k$ is met Stone provides the following minimax formula for the maximum likelihood estimator

$$\hat{\lambda}_i = \min_{s \leq i} \max_{t \geq i} \frac{\sum_{r=s}^{t} O_r}{\sum_{r=s}^{t} E_r} \qquad i = 1,2,...,k.$$

The application of this formula is equivalent to calculating the relative risks in each of the groups and if $\lambda_i < \lambda_{i+1}$ the groups i and i + 1 are both assigned the risk of the group formed by combining them.

The test statistic quoted by Stone is

$$T_l = \frac{L_1}{L_0}.$$

However, for consistency with other likelihood ratio tests presented in this Chapter the test statistic used is

$$-2\ln\left(\frac{L_0}{L_1}\right) = 2\left\{\sum_{i=1}^{k} E_i - \sum_{i=1}^{k} O_i \ln(E_i) - \sum_{i=1}^{k} \lambda_i E_i + \sum_{i=1}^{k} O_i \ln(\lambda_i E_i)\right\}$$

$$= 2\sum_{i=1}^{k} \{E_i(1-\lambda_i) + O_i \ln(\lambda_i)\}.$$

The sensitivity and specificity are unaffected by this change of test statistic as they are merely in different domains.

Although easier to compute than the test statistic determined from assuming a binomial distribution the above test statistic for Stone's test is still computationally difficult to generate thus Stone's paper provides a less computer intensive alternative test statistic, which is used in the majority of studies. In this case the most exposed group is considered and the relative risk is estimated. Then the two most exposed groups are combined and the relative risk estimated for that. This process continues until the relative risk is estimated for the whole population. The test statistic is then the maximum of these relative risks. Therefore, the test statistic is the maximum likelihood estimator for the first parameter in the order restriction $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_k$ and is given by

$$T_\lambda = \hat{\lambda}_1 = \max_{1 \leq t \leq k} \frac{\sum_{i=1}^{t} O_i}{\sum_{i=1}^{t} E_i}.$$

Stone claims that this is interpreted as "...relating to the radius around the putative source within which the observed relative risk is maximised" when distance is a proxy for exposure and thus determines the partitioning of the study region [236]. Even when distance is used as the determinant of the boundaries of the sub-regions, this statement is not strictly true as the radii are selected by the investigator. Therefore, another radius that is not in this predetermined subset of radii may produce a relative risk that is higher but it will not be selected.

The testing procedure presented above as described by Stone can only be applied to aggregated count data. However, if individual level point data was available, as is the case for congenital anomalies, then grouping this data could mean that some information is lost. Consequently, it may be beneficial to use methodologies that make full use of point data. One such method is presented in Stone's paper and is discussed below.

Stone's test for point data

Stone's method of analysing point data assumes that the exposure level of the cases and controls can be ranked without ties. Suppose

$$X_i = \begin{cases} 1 & \text{if individual with rank i is a case} \\ 0 & \text{otherwise} \end{cases}$$

and let C be the total number of cases and N – C the total number of controls.

The null distribution of the $X_i's$ is that of a random permutation of C 1's and N – C 0's conditional on C and N – C. The alternative hypothesis assumes that they are not random but are influenced by the factor of interest.

The test statistic here is

$$T_p = \max_{1 \le t \le N} \sum_{i=1}^{t} \frac{X_i}{t}.$$

There are a number of problems with implementing this procedure. If the first patient, in other words the patient that has the highest exposure, is a case then the risk is 1 and it is impossible to improve on that so the test statistic is unstable. There is no use of the expected number of cases and, therefore, confounders cannot be adjusted for. If the east and north grid coordinates for the health outcome data are determined from postcodes, as it is with the birth data for Nant y Gwyddon, there is likely to be some individuals with the same coordinates and consequently they will be allocated to the same level of exposure. Individuals with the same exposure level will be tied when ranked, which cannot occur under the assumption of Stone's proposed test for point data, that there must not be any ties.

Therefore, in order to apply this method to situations where some individuals have equal levels of exposure the test statistic has to be adapted to cope with the tied ranks. Suppose there are K possible exposure risks that the individuals are assigned to. These K exposure risks are ranked prior to calculating the test statistic. The adapted version of the test statistic is given by

$$T_p = \max_{1 \le t \le K} \frac{\sum_{i=1}^{t} Y_i}{\sum_{i=1}^{t} n_i}$$

where Y_i is the number of cases with exposure risk of rank i and n_i is the total number of individuals where the exposure risk is ranked i. This approach is equivalent to using T_λ without initially aggregating the data.

Even though the difficulties of applying Stone's proposed test to tied data can be overcome by using this new test statistic the main problem still remains. It is not possible to adjust for confounding factors using either of these test statistics and it is desirable to be able to. Therefore a new parametric method for testing for the spatial association between a health outcome and exposure to a risk factor is introduced below.

2.4.2 New Methodology

Trend test for spatial association

Suppose the risk of a case can be written in the form $c_i + m\gamma_i$ where c_i is the confounder dependent baseline risk and γ_i is the exposure measure. In addition, let

$$X_i = \begin{cases} 1 & \text{if patient i is a case} \\ 0 & \text{otherwise} \end{cases}$$

The hypotheses are

$H_0 : m = 0$ i.e. there is no additional risk attributable to exposure
$H_1 : m \neq 0$ i.e. increased exposure elevates the risk of the health outcome.

The likelihood functions are derived from a Bernoulli distribution as before and given below

$$L_0 = \prod_{i=1}^{N} c_i^{X_i} (1 - c_i)^{1-X_i}$$

$$L_1 = \prod_{i=1}^{N} (c_i + m\gamma_i)^{X_i} (1 - (c_i + m\gamma_i))^{1-X_i} .$$

Hence, the log likelihood under the alternative is

$$l_1 = \ln(L_1) = \sum_{i=1}^{N} \{X_i \ln(c_i + m\gamma_i) + (1 - X_i)\ln(1 - (c_i + m\gamma_i))\}.$$

The maximum likelihood estimator for m is determined by differentiating the above with respect to m and equating it to zero.

$$\frac{\partial l_1}{\partial m} = \sum_{i=1}^{N} \left\{ \frac{X_i \gamma_i}{c_i + \hat{m}\gamma_i} - \frac{(1 - X_i)\gamma_i}{1 - (c_i + \hat{m}\gamma_i)} \right\}$$

$$= \sum_{i=1}^{N} \gamma_i \left\{ \frac{X_i(1-(c_i+\hat{m}\gamma_i)-(1-X_i)(c_i+\hat{m}\gamma_i))}{(c_i+\hat{m}\gamma_i)(1-(c_i+\hat{m}\gamma_i))} \right\}$$

$$= \sum_{i=1}^{N} \left\{ \frac{\gamma_i(X_i-c_i-\hat{m}\gamma_i)}{(c_i+\hat{m}\gamma_i)(1-(c_i+\hat{m}\gamma_i))} \right\} = 0.$$

The solution of the above equation cannot be found algebraically and, therefore, the bisection method was used to determine the maximum likelihood estimator of m.

The test statistic in this situation is

$$-2\ln\left(\frac{L_0}{L_1}\right) = -2\sum_{i=1}^{N}\{X_i \ln(c_i) + (1-X_i)\ln(1-c_i) - X_i \ln(c_i+\hat{m}\gamma_i) - (1-X_i)\ln(1-(c_i+\hat{m}\gamma_i))\}$$

2.4.3 Assessing the Efficiency of the Tests

It is clear from the discussion above that there are a number of options for testing for spatial association between a health outcome and exposure to a specific risk factor. In some situations the accuracy of the geographical information attached to the health outcome data is insufficient for fine subdivision of the study region by exposure in which case there is no option but to use the concentric circle approach [226]. However the concentric circle approach is overly simplistic with a high level of exposure misclassification introduced by the use of only two levels of exposure: exposed and unexposed. Therefore if the geographical information linked to the health outcome is more accurate then either more subdivisions of the study region or even the use of point data is possible in which case either Stone's proposed tests or the trend test for spatial association can be implemented. The relative performance of these tests is assessed below to indicate which test should be adopted for a given situation.

The lack of good quality exposure data prevents the application of these methods to the investigation of the spatial association between the risk of congenital anomalies and exposure to contaminants released from the Nant y Gwyddon landfill site. However, the population pattern around Nant y Gwyddon can be used to test the efficiency of the different testing procedures under an assumed exposure model. The assumed exposure model consists of two point sources of pollution and the risk attributable to exposure decreases with increasing distance from the centre of either of these two point sources. The two point sources are

situated at the location of the two hotspots of elevated relative risk identified by James et al. [10].

The efficiency of a testing procedure is summarised using the sensitivity for different specificities. The specificity is the conditional probability that the testing procedure leads to not rejecting the null hypothesis given that the null hypothesis is true. The sensitivity or power is the conditional probability that the testing procedure leads to rejecting the null hypothesis given that the alternative hypothesis is true. It is clearly desirable for both the specificity and sensitivity to have high values. However, increasing the specificity leads to a decrease in the sensitivity so there has to be a balance between the two. The significance level is equal to 1 – the specificity, i.e. it is the probability of rejecting the null hypothesis when it was in fact true, which is why the significance level is typically chosen to be very small in hypothesis testing. The efficiency of the test is summarised in a receiver operator characteristic (ROC) curve which is a summary of the sensitivity at each possible significance level. A good test will have high values for the sensitivity even when the significance level is small.

1000 simulations were generated under the null hypothesis of no spatial association and also under the alternative hypothesis of spatial association using the study population from the 20km square region centred on Nant y Gwyddon to generate the ROC curves for each of the tests. The simulation process is discussed below.

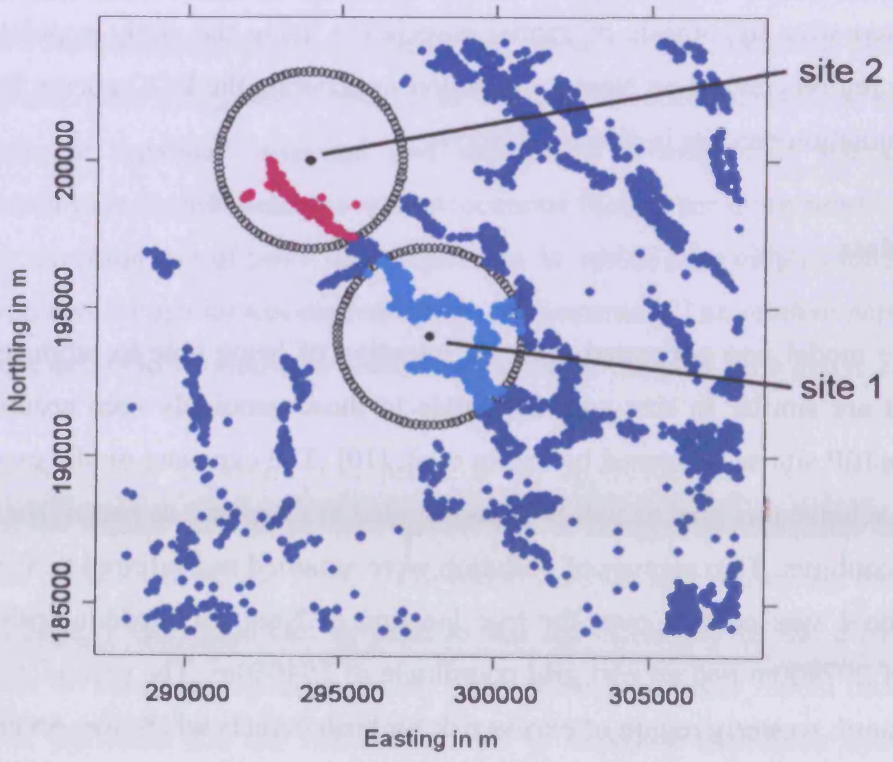
Simulating data

The exposure model was generated with the intention of being able to produce relative risk contours that are similar in size and magnitude to those genuinely seen around the Nant y Gwyddon landfill site as presented by James et al. [10]. The exposure model explained in this section does achieve this goal as will be demonstrated in Chapter 3 as part of the discussion of relative risk contours. Two sources of pollution were assumed and referred to as site 1 and site 2. Clearly site 1 was centred over the true location of Nant y Gwyddon with a north grid coordinate of 297900m and an east grid coordinate of 194050m. The second site was chosen to be near a north westerly region of excess risk for birth defects which was revealed by James et al. with a north grid coordinate of 294000m and an east grid coordinate of 200000m. This choice of sites means that both regions of high (site 1) and low (site 2) population density are represented in the exposed regions. The choice of two sites as opposed to one introduces an

additional level of complexity through which the tests are assessed but not too much complexity. Additionally the number of sites used here coincides with the number of regions over which the relative risk exceeds 2 as seen in James et al.'s paper [10].

The exposure measure for this simulation exercise was chosen to be the distance from the centre of the sites, which is not a good indicator of exposure to contaminants released from the landfill site, since other factors such as topography and wind are important in determining the distribution of substances emitted from the site. However, for the purposes of simulation and test assessment this assumption was thought to be sufficient. The impact of exposure to contaminants from landfill sites does not go beyond 3km from the centre of the site according to Dolk et al. [64]. Therefore, it was assumed that the population living further than that from both sites were unexposed and had no additional risk of birth defects that is attributable to exposure. The remaining population were assumed to have an exponentially increasing additional risk towards the centre of the site.

Figure 2.10: the location of the two exposure sites and all births in the 20km square region centred over Nant y Gwyddon



The location of these two selected sites and their 3km radius regions of exposure are illustrated in Figure 2.10. The blue points represent births that are not exposed to either site, those exposed to site 1 are turquoise and the pink points represent those exposed to site 2.

The total risk for mother i of having a baby with a congenital malformation is

$$p_i = \begin{cases} \alpha_i + \beta\left[\sum_{\forall j}\{\exp(-d_{ij}) - \exp(-3)\}\right] & \text{if } d_{ij} \leq 3 \\ \alpha_i & \text{otherwise} \end{cases}$$

where α_i is the baseline probability of a birth defect and d_{ij} is the distance of individual i from site j. The second term in the first line of the equation represents the risk attributable to exposure and the value of β controls the strength of this. In order to simulate the cases under each hypothesis the values for α_i and β must be known for both the null and alternative hypotheses.

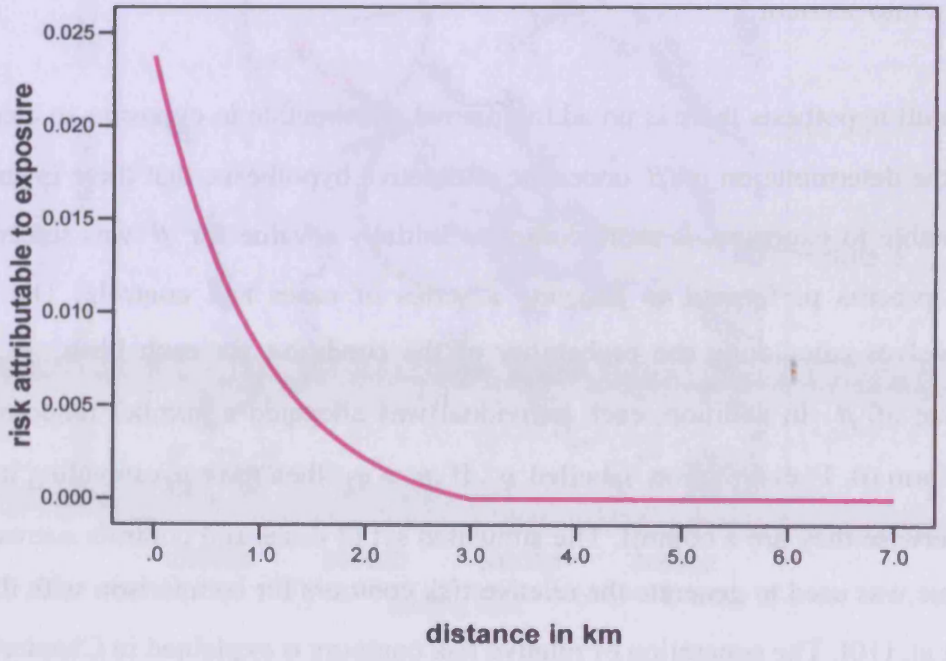
When dealing with real data, a logistic regression model using known confounders would be used to predict the value of α_i for each individual. However, initially α_i was fixed as 0.01 (the approximate risk of birth defects between 1983 and 1997 across the whole of Wales) for all patients to develop the methodology, which could then be adapted later to take the confounders into account.

Under the null hypothesis there is no additional risk attributable to exposure so clearly $\beta = 0$. However, the determination of β under the alternative hypothesis, that there is an additional risk attributable to exposure, is more complex. Initially a value for β was selected and the simulation process performed to generate a series of cases and controls. The simulation process involves calculating the probability of the condition for each birth, p_i, using the chosen value of β. In addition, each individual was allocated a number randomly selected from a uniform (0, 1) distribution, labelled q_i. If $p_i > q_i$, then the corresponding individual is a case, otherwise they are a control. The simulated set of cases and controls corresponding to that β value was used to generate the relative risk contours for comparison with those shown in James et al. [10]. The generation of relative risk contours is explained in Chapter 3.

This process was repeated for a number of different values of β. The chosen value of β should correspond to the relative risk plot which most closely resembled the characteristics of those seen around the Nant y Gwyddon landfill site. If this similarity of relative risk contours is achieved then the alternative hypothesis used in the development and assessment of the testing procedure could plausibly be true. It would be fruitless to assess the performance of the hypothesis test using an inconceivable alternative as the result will be a misleading power that can never be achieved in reality.

It will be shown in Chapter 3 that when $\beta = 0.025$ the relative risk contours for the simulated data are of similar size and magnitude to those seen in James et al. [10]. In fact as discussed in detail later the resultant relative risk surface for a simulation for which it is assumed that $\beta = 0.025$ does not result in elevations that are as high as those seen in the paper by James et al. [10]. Therefore, the risk attributable to exposure could even be more conservative in the simulation than it is in reality thus if these testing procedures were applied to real data for congenital anomalies around the Nant y Gwyddon landfill site the efficiency could be better than is indicated from the simulations.

Figure 2.11: a plot of the relationship between risk attributable to exposure and the exposure measure in the assumed exposure model



Now that β and α_i have been determined it is possible to plot the relationship between the exposure measure, here distance, and the exposure risk, determined by the exposure model. This plot can be seen in Figure 2.11.

Representative values of α_i and β can now be considered known under the null and alternative hypotheses therefore it is possible to perform the simulations as described on page 97. Simulations under each hypothesis enable the calculation of the sensitivity and the specificity, estimated by the proportion of simulations under the alternative that lead to correctly accepting the alternative hypothesis and the proportion of simulations under the null that correctly result in the acceptance of the null hypothesis, respectively. The sensitivity and specificity are both high in a good hypothesis test and, hence, are good indicators of the efficiency of a test. 1000 simulations of cases of congenital malformations under each hypothesis were performed, as this number gives estimates for the sensitivity and specificity to an acceptable level of accuracy [237]. These simulations can now be used to assess the efficiency of the testing procedures presented by Stone and, subsequently, the more complex trend test for a relationship between exposure and the given health outcome using likelihood ratios.

Efficiency of Stone's test

The 1000 simulations of cases under the null and alternative hypotheses, generated as described on page 97, were used to calculate the test statistics for Stone's test. The births were grouped in order of decreasing risk attributable to exposure as shown in Table 2.4. The likelihood ratio based test statistic T_l was calculated for each of the simulations under the null and alternative hypotheses. The distributions of T_l under the null and alternative hypotheses were used to determine the sensitivity for given specificities.

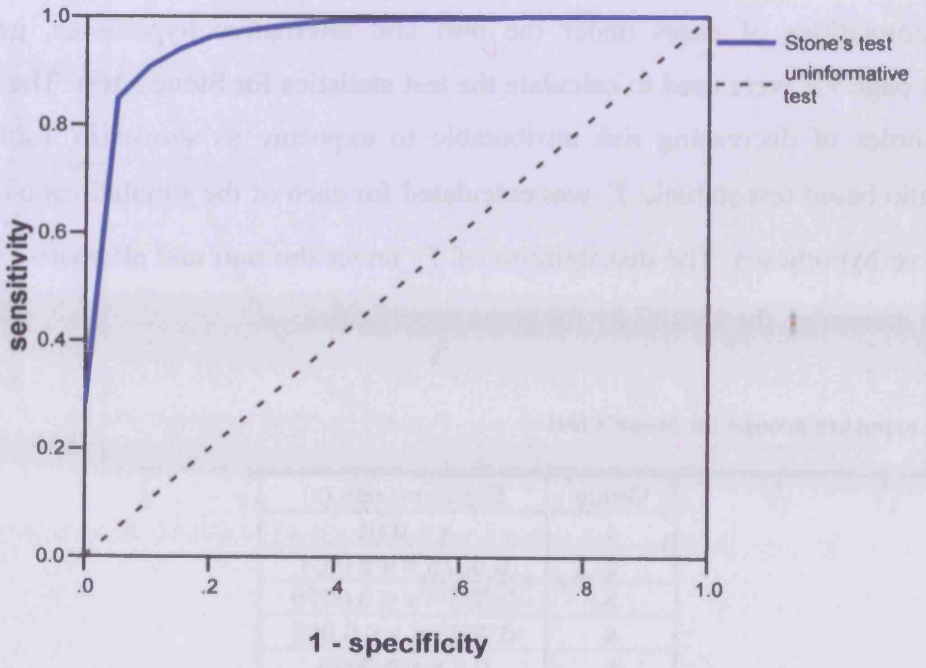
Table 2.4: the exposure groups for Stone's test

Group	Exposure risk (x)
1	$x > 0.01$
2	$0.0075 < x \leq 0.01$
3	$0.005 < x \leq 0.0075$
4	$0.0025 < x \leq 0.005$
5	$0 < x \leq 0.0025$
6	$x = 0$

A ROC curve is a plot of 1 – the specificity against the corresponding sensitivity. A perfect hypothesis test, where there is complete separation between the distributions of the test statistic under the null and alternative hypotheses, has a ROC curve that follows the top and left hand axis on the diagram. An uninformative test follows the 45° line and occurs when the distributions of the test statistics under the two hypotheses are identical.

ROC curves are traditionally used to assess the efficiency of screening procedures utilised in the diagnosis of disease. ROC curves have been adopted here to assess the relative efficiency of the hypothesis testing procedures. Although this is not strictly the correct use of ROC curves they provide a good way of summarising the relative performance of the tests. In the context of hypothesis testing 1 – the specificity relates to the significance level and the sensitivity relates to the power. The power of the tests and the corresponding 95% confidence intervals are quoted at the 5% significance levels to give a sense of the power at a commonly used significance level and the degree of uncertainty surrounding that power. The presentation of the ROC curves and 95% confidence intervals are to give a feel of the relative performance of the tests and are not intended for use in formally testing for differences between the methods.

Figure 2.12: the ROC curve for Stone's test using the simulations from Nant y Gwyddon



The ROC curve for Stone's test using the simulations from the Nant y Gwyddon landfill site is given in Figure 2.12 and is closer to the perfect test scenario than that of the uninformative test indicating that the test is reasonably good. At the 5% significance level the power of Stone's test is 84.9%, which suggests that if the data available is at the area level then the use of Stone's test is reasonable as a way of determining whether or not there is a causal link between the chosen factor and the medical condition.

The confidence interval for the power of Stone's test at the 5% level is (82.6%, 87.0%) which indicates a rough approximation of the error in the estimation of the sensitivity. Clearly, the specificity is also generated from the simulations so there is some error in the selection for the location of the decision line used to calculate the sensitivity for a given specificity. It should be noted that in the generation of the confidence intervals quoted the error in the specificity was assumed to be zero.

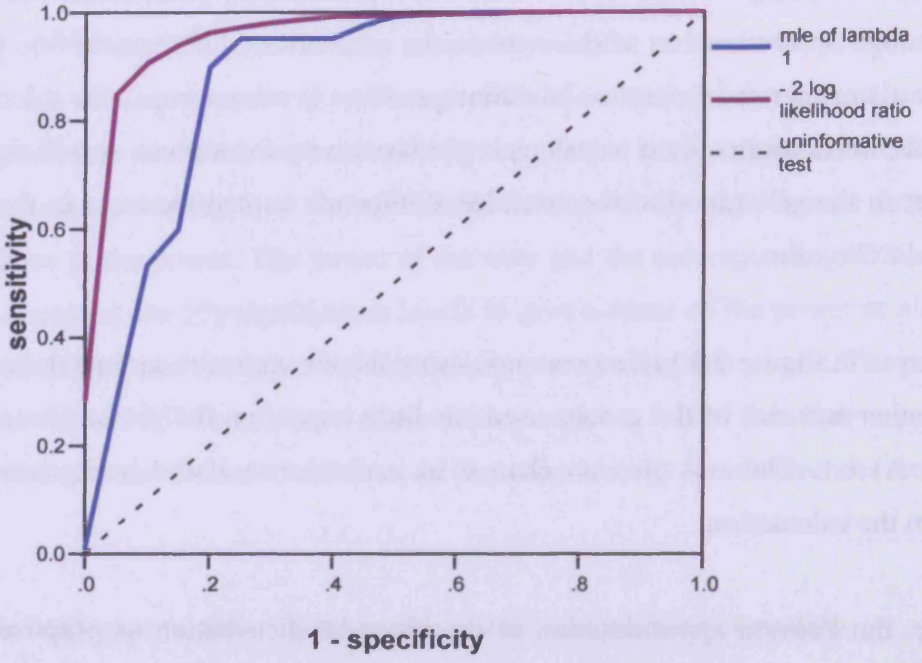
The ROC curve in Figure 2.12 was generated using the six exposure groups defined in Table 2.4. The number and size of the groups used has little impact on the performance of the test, see Section A1.1.1. There is also no change in performance if the background region is omitted from the calculation.

Furthermore, the Poisson approximation of the binomial distribution as proposed by Stone was adopted to generate this ROC curve. If the Poisson approximation is not used and the test statistics are determined using only the binomial distribution referred to above as T_b, the likelihood ratios generated are almost identical to those determined from the Poisson assumption, see Section A1.1.2. Therefore, despite the fact that the binomial distribution is more appropriate for a situation of this type, the Poisson approximation, as expected, is reasonable. This result is advantageous because Stone's approach is less computer intensive than when working without the Poisson approximation and if there are non-constant baseline risks then the Poisson approximation is more efficient in terms of computation time particularly for 2000 simulations.

A further simplification proposed by Stone to reduce the computer intensity of this testing procedure is the use of a simplified test statistic. The ROC curve using this alternative test statistic T_λ is presented in Figure 2.13 alongside that of the more complex likelihood ratio test

statistic, T_l. The power is significantly reduced by the simplification with the sensitivity equal to only 27.9% (95% CI: 25.2%-30.8%) at the 5% significance level. Therefore, it is advised that T_λ is not used when testing for a relationship between exposure to a specified factor and the risk of a medical condition as it is relatively inefficient in detecting an effect.

Figure 2.13: the ROC curve for Stone's test with T_λ compared to testing with the full test statistic using the simulations for Nant y Gwyddon



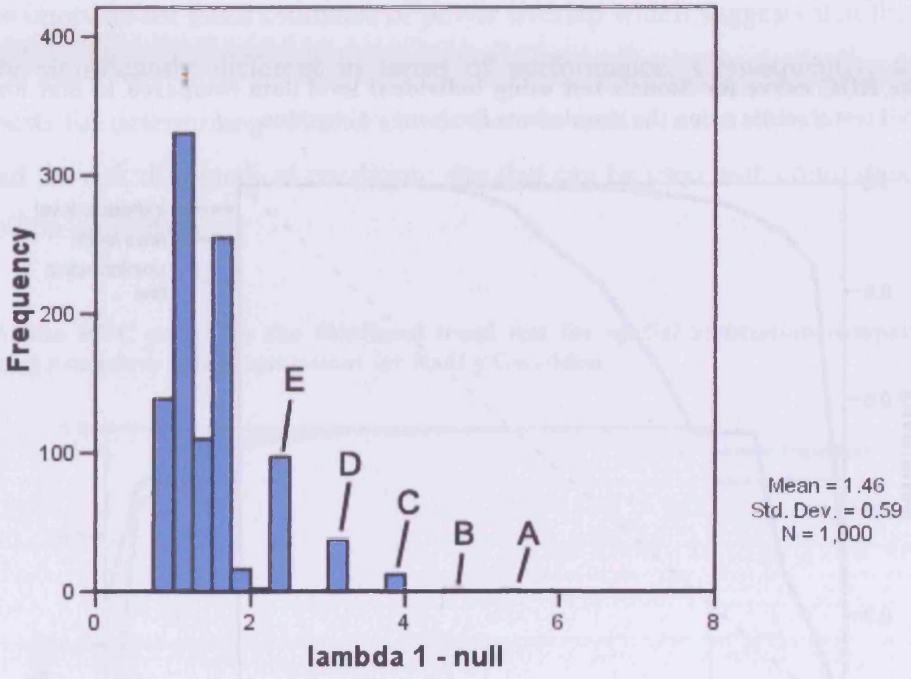
The ROC curve for T_λ is jagged in appearance, this is a result of discontinuities in the distribution of T_λ under the null and alternative hypotheses. Evidence of these punctuated distributions is given in Figure 2.14. The reason for the discontinuities lies in the simulations where there is a high number of simulated congenital anomalies in group 1 which means that T_λ, given by

$$T_\lambda = \max_{1 \le t \le k} \frac{\sum_{i=1}^{t} O_i}{\sum_{i=1}^{t} E_i}$$

is equal to the relative risk in group 1. Bars A, B, C, D and E correspond to the simulations where there are 7 (n = 1), 6 (n = 1), 5 (n = 12), 4 (n = 37) and 3 (n = 97) simulated congenital anomalies in group 1, respectively. This is apparent because an increase of one simulated case

in group 1 increases the relative risk in group 1 by 0.769 which coincides with the size of the gap between these five bars. When T_λ is calculated by combining group 1 with other groups the value of T_λ does not exceed 2 for these 1000 simulations and so the gaps appear in the histogram.

Figure 2.14: the distribution of T_λ for the simulations under the null hypothesis from Nant y Gwyddon



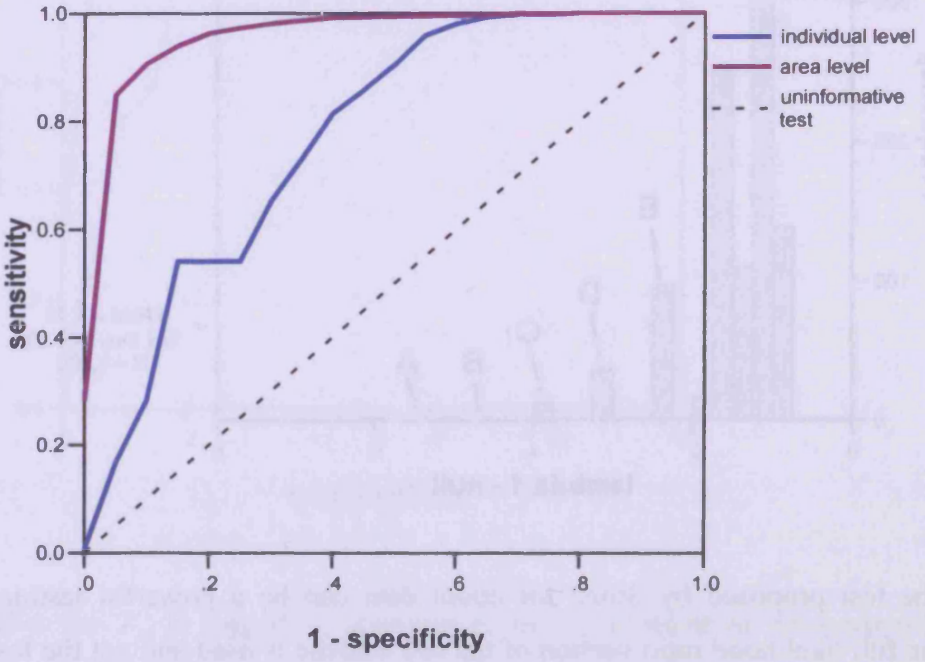
Evidently the test proposed by Stone for count data can be a powerful testing procedure providing the full likelihood ratio version of the test statistic is used and not the less computer intensive, simplified version. Unfortunately, in order to reduce computing time the majority of studies that implement Stone's test use the simplified version of the test statistic and are consequently not taking advantage of the potential power of the test. In all other ROC curves for Stone's test presented in this thesis only the full likelihood ratio test statistic is considered to maximise the power.

Efficiency of Stone's test for point data

In Stone's paper a testing procedure was proposed that dealt with individual level point data as opposed to aggregated count data. The ROC curve for Stone's test for point data was

generated using the simulated data for Nant y Gwyddon under the null and alternative hypotheses. This ROC curve is shown in Figure 2.15 along with the ROC curve for Stone's test for count data. The ROC curve for Stone's test for point data is jagged in appearance like the ROC curve for Stone's test using the simplified test statistic, T_λ. This is because the use of post coding to pinpoint the location of the births means that there are multiple births at each radii and consequently the characteristics of T_p are similar to those when using T_λ only with more exposure groups.

Figure 2.15: the ROC curve for Stone's test using individual level data compared to that for area level data with the full test statistic using the simulations for Nant y Gwyddon

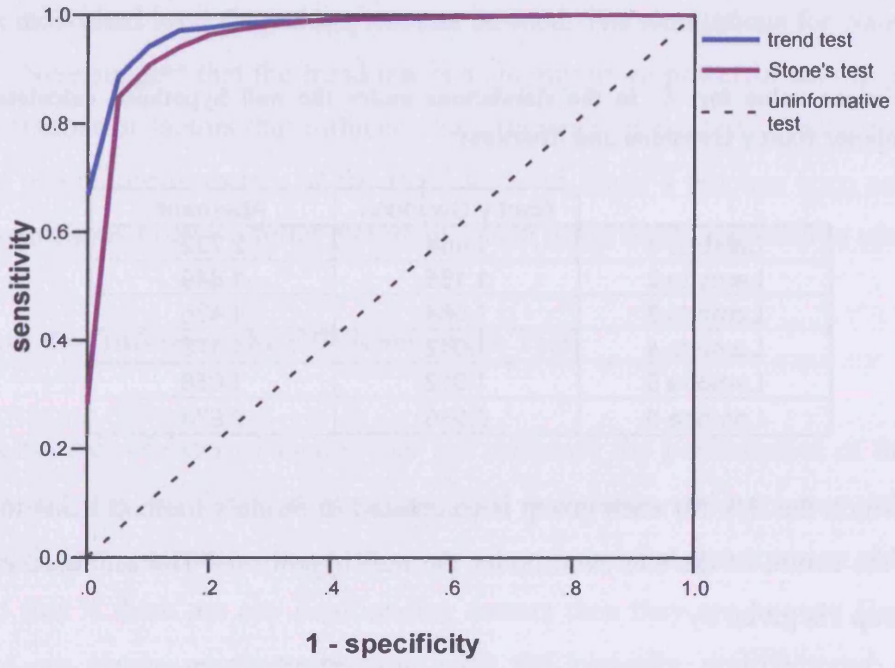


The performance of Stone's test for point data is very poor with a power of only 16.7% (95% CI: 14.5%-19.1%) at the 5% significance level. The reason for the dramatic drop in power for the point data version of Stone's test is the larger overlap observed for distributions of the test statistic under the null and alternative hypotheses compared to that observed with Stone's test for count data. Therefore, it is better to aggregate the data into regions and apply Stone's test for count data than to use this alternative method directly with the point data. The better performance observed for Stone's test using count data means that any future ROC curve comparisons with Stone's test displayed in this thesis only consider the count data version with the full likelihood ratio test statistic.

Efficiency of the trend test for spatial association

The ROC curve for the trend test for spatial association is displayed in Figure 2.16 alongside the equivalent curve for Stone's testing procedure. The trend test is more efficient than Stone's approach with 88.7% (95% CI: 86.6%-90.5%) power at the 5% significance level compared to 84.9% (95% CI: 82.6%-87.0%) with Stone's method. However, the 95% confidence intervals for these estimates of power overlap which suggests that they may not be statistically significantly different in terms of performance. Consequently, there are two powerful tests for determining whether or not there is spatial association between exposure to a factor and the risk of a medical condition: one that can be used with count data and one that can be used on point data.

Figure 2.16: the ROC curve for the likelihood trend test for spatial association compared to that for Stone's testing procedure using simulations for Nant y Gwyddon



The slightly worse performance seen in Stone's test compared to the trend test could be explained by bias in the estimates in λ_i. The risk of the given health outcome should be equal to 1 under the null hypothesis in all regions. However, the risks calculated when applying Stone's test for each of the regions in the simulations under the null hypothesis do not typically have a mean estimated risk of 1. The mean risk for each of the regions in the 1000

simulations under the null hypothesis are given in Table 2.5 for both Nant y Gwyddon and also for Abernant, which is another landfill site in Wales with a much smaller population and is discussed in detail later.

It is clear from Table 2.5 that the risks deviate from 1 with the risks for Abernant deviating more severely from 1 because the population is smaller and therefore the variation is larger. The mean risk for the inner circle is substantially higher than 1 at 1.464 and 2.732 for Nant y Gwyddon and Abernant, respectively, and the mean risks for each group decline monotonically with increasing distance from the site until the mean risks actually fall below 1 at 0.986 and 0.974. These mean risks are calculated using data simulated under the null hypothesis and they should all be approximately equal to 1 with a gradient of 0 but this is not the case. The observed increasing risk towards the centre of the site under the null hypothesis means that the estimates for λ_i under the null hypothesis have estimates similar to those expected under the alternative hypothesis. Hence, Stone's testing procedure is not as efficient as it should be.

Table 2.5: the mean value for λ_i in the simulations under the null hypothesis calculated using the minimax formula for Nant y Gwyddon and Abernant

	Nant y Gwyddon	Abernant
Lambda 1	1.464	2.732
Lambda 2	1.165	1.849
Lambda 3	1.084	1.436
Lambda 4	1.042	1.175
Lambda 5	1.012	1.058
Lambda 6	0.986	0.974

The way in which the risk for each group is calculated in Stone's method leads to increasing risk towards the centre of the site even under the null hypothesis. The unbiased estimator of the risk in group i is given by

$$\hat{\lambda}_i = \frac{O_i}{E_i}$$

whereas in Stone's test the risk in group i is calculated using the minimax formula given by

$$\hat{\lambda}_i = \min_{s \le i} \max_{t \ge i} \frac{\sum_{r=s}^{t} O_r}{\sum_{r=s}^{t} E_r}$$

so that the constraint that $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_k$ is met. It is this minimax formula that leads to the biased values for the risk in each group. This was not shown by Stone and a detailed explanation of why it is biased is given in Section A1.12.

Effectively by meeting the constraint that $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_k$ the minimax formula ensures that the risks are either equal throughout all the groups or they increase towards the centre of the site. Therefore, in a situation where the risk is increasing away from the centre of the site the values of λ_i are set to be equal in all groups. Furthermore, in a simulation where the risk undulates with decreasing distance from the site, the value of λ_i will increase with decreasing distance. Consequently, the risk either increases towards the centre of the site or is equal in all groups even under the null hypothesis, which has been observed.

If the data is only available at the area level then Stone's approach is the only option for testing for a relationship between exposure and the health outcome. However, if the data is available at individual level then either test can be used. The simulations for Nant y Gwyddon considered above suggest that the trend test is a slightly more powerful testing procedure but there are a number of factors that influence the efficiency of the tests. Therefore, in the next section, the relative performance of the trend test and Stone's test has been assessed under each of the situations to give a fuller picture of which is the better approach to take.

2.4.4 Factors that Influence the Efficiency of the Tests

There are a number of external factors that can influence the performance of the hypothesis testing. The test was developed using a specified population and exposure risk which will change in each new application. In addition, it was assumed that the exposure risk is correctly defined and that if there are any confounding factors then they are known. However, these assumptions are rarely going to be true with the typically multifactorial aetiology of conditions and the limitations in modelling exposure. Therefore, this subsection is devoted to discussing the extent to which these factors affect the performance of the hypothesis tests.

Adjusting for known confounders

In the simulation process it was assumed that the baseline risk was 0.01 for all mothers in the population, in other words, all mothers were at equal risk if there was no additional risk attributable to the exposure of interest. However, in reality this is not true, the baseline risk varies from mother to mother because other factors including socioeconomic status, gender of child, maternal age, hospital of birth and year of birth affect a mother's risk of a congenitally malformed birth [272]. These non-constant baseline risks mean that if an additional risk does exist, as a result of exposure, then it will be obscured by the variation in the baseline.

Hence, in order to consider the risk attributable to one specified factor alone, the remaining factors must be adjusted for in the calculation of the likelihood ratios used in the testing procedures. In Stone's test, this is achieved by changing the expected number of congenital malformations used in the calculation of the likelihood ratios. Therefore, rather than the expected number of birth defects being 0.01 multiplied by the number of births in the region, it is calculated by summing the predicted probabilities for all births in the region that were determined from the fitted logistic regression model. In the trend test, the value of c_i for birth i is the corresponding predicted probability from the logistic regression model and not equal to 0.01 for all births.

In order to assess the efficiency of the hypothesis tests when adjustment for known confounding is required, the simulations were repeated as before under the null and alternative hypotheses, however, the values of α_i were equal to the predicted probability for the i^{th} mother calculated using the logistic regression model. These two new sets of 1000 simulations under each of the hypotheses were used to produce the ROC curve for each of these tests when there is adjustment for known confounders.

The ROC curves for the testing procedures with adjustment for known confounders are given in Section A1.5. The performances of the hypothesis tests are similar to when the baseline was assumed to be constant and, consequently, no confounder adjustment was required. At the 5% significance level Stone's test has a power of 74.7% (95% CI: 71.9%-77.3%) when adjusting for known confounding, which is almost a 10% drop from the situation where there was a constant baseline in the simulation process. The trend test, on the other hand, experienced

almost no change in power at the 5% significance level where the sensitivity was 89.1% (95% CI: 87.0%-90.9%). The ROC curves for Stone's test and the trend test when known confounders are adjusted for are compared in Section A1.11 revealing that the trend test performs better under this scenario. In fact, at the 5% significance level, the sensitivity is higher by nearly 15% in the trend test compared to Stone's approach.

Effect size

The effect size varies between different risk factors and between health outcomes. For example, in terms of the impact of landfill sites on the risk of congenital malformations the effect size is relatively large, with reports of an increased risk of 33% within 3km of 21 sites in the UK [5]. On the other hand, as later discussed, when dealing with sex ratios the effect size can be as small as 0.5%. The effect size will influence the performance of the test because as the effect size reduces the test becomes less sensitive.

In the exposure model given on page 97 the value of β controls the strength of the risk attributable to exposure. Consequently, to investigate the impact of the effect size the simulations of cases of birth defects under the alternative hypothesis were performed again with a weaker and a stronger effect size. The testing procedure was then performed with these two new alternative hypotheses.

The weaker exposure effect has a β value of 0.0125, half that used before, and for the stronger exposure effect the β value was doubled to 0.05. The ROC curves for each of the varying effect sizes are presented in Section A1.6 and in both tests the efficiency is reduced with a decreasing effect size. When the effect size is doubled ($\beta = 0.05$) the power at the 5% level is 100% (95% CI: 99.6%-100%) for the trend test and 98.6% (95% CI: 97.7%-99.2%) for Stone's test from this finite set of 1000 simulations. In most cases the effect size would not realistically be this strong. However, with the halved effect size ($\beta = 0.0125$) the tests both have a reduced power relative to the situation where the effect size is close to what is expected for congenital anomalies in terms of exposure to the Nant y Gwyddon landfill site. The sensitivity of Stone's test drops down to 33.9% (95% CI: 31.0%-36.9%) at the 5% level whereas the corresponding power for the trend test is 44.0% (95% CI: 41.0%-47.1%) which despite being a little higher is still not very good.

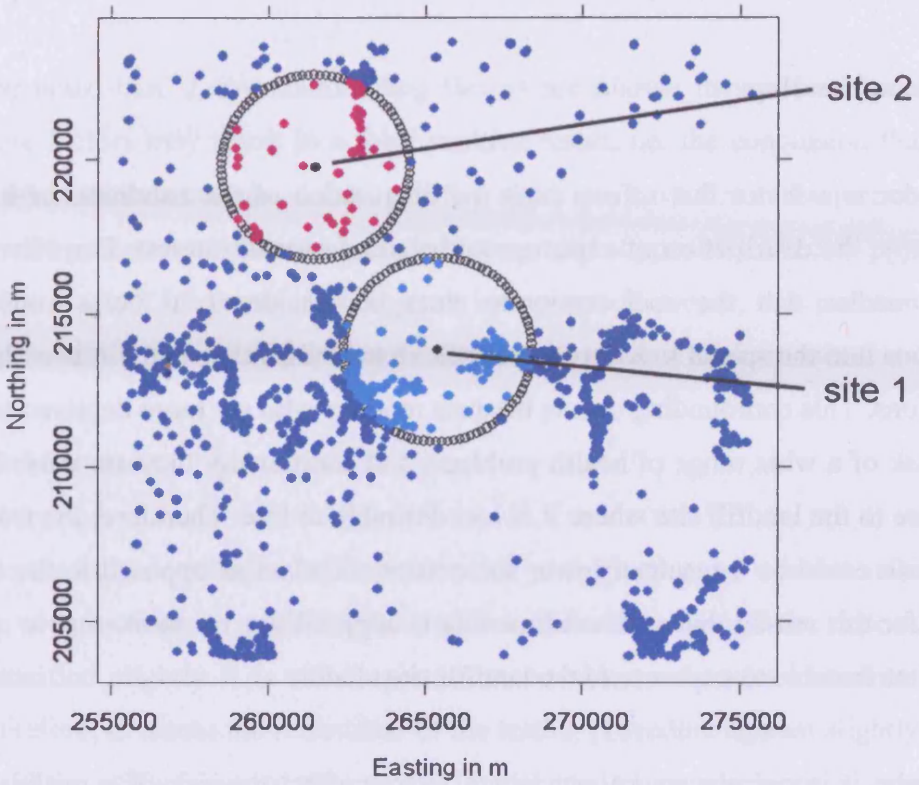
Therefore, if the effect size is too small then both testing procedures are unlikely to be powerful enough to detect a causal link. However, as the effect size diminishes the importance of identifying a spatial association reduces too. Regardless of effect size the trend test still out-performs Stone's method at the 5% level, although, there is little difference between the ROC curves for any effect size considered here.

Population size

The region surrounding the Nant y Gwyddon landfill site is densely populated and as a result by taking all births between 1983 and 1997 whose mothers resided in the 20km square region centred over the site a population of 41,337 is selected. However, this dataset is reasonably large and if a sparsely populated region was taken then the population would be much smaller. Abernant landfill site is located in such an area, where the population is only 8,159 under the same specifications which is less than a fifth of that around Nant y Gwyddon. In addition, the ribbon-like structure of the population around Nant y Gwyddon is not present around Abernant where the population is more randomly scattered. Therefore, by comparing the performance of the test using the population of Abernant to that of Nant y Gwyddon it will be possible to get some idea of the impact of the population size and distribution on the efficiency of the tests.

Abernant is a landfill site near a town 16 miles North of Swansea, called Ammanford, located in Wales. The population pattern in Abernant is shown in Figure 2.17 along with the location of the two exposure sources. The relative location of the exposure sites is identical to that for Nant y Gwyddon with site 1 now centred over Abernant. The points representing the population clearly show the lower population density of Abernant, particularly within 3km of the exposure sources.

Figure 2.17: the location of the two exposure sites and all births around Abernant



Simulations were performed on the Abernant population using the exact same method used on Nant y Gwyddon, as described on page 97, and the ROC curves for the two testing procedures were generated, see Section A1.7. The efficiency of the test is greatly reduced by the smaller population this is because the statistics are based on fewer numbers and therefore the sampling variation in the test statistics is larger. The greater variance results in a smaller degree of separation between the distributions under the null and alternative hypotheses, therefore, the tests do not perform as well as they do with Nant y Gwyddon. The drop in sensitivity between using simulations from Abernant and using those from Nant y Gwyddon is large however the trend test is still more efficient at the 5% significance level with a power of 25.2% (95% CI: 22.6%-28.0%) compared to 19.6% (95% CI: 17.3%-22.2%) for Stone's.

The size and distribution of the population clearly plays an important part in the absolute efficiency of the testing procedure. If the population size is small then the tests lose efficiency. However, if there is a sufficient population size included in the study then the tests provide a powerful way of identifying a causal link. Therefore, the loss of efficiency could be remedied by using multiple equivalent sites and this should particularly be considered if the effect size

is small. Additionally using the trend test instead of Stone's test increases the power of the testing procedure with small population sizes.

Unknown confounding

A confounder is a factor that affects both the distribution of the condition of interest and, independently, the distribution of exposure to the risk factor of interest. For example, it was mentioned earlier that the socioeconomic class is considered to be a confounder for investigations into the spatial link between exposure to landfill sites and births with congenital malformations. This confounding occurs because mothers who are more deprived tend to have a greater risk of a wide range of health problems and additionally they are more likely to be located close to the landfill site where it is less desirable to live. Therefore, the reason for the increased risk could be a result of lower socioeconomic class as opposed to the landfill site itself. It is for this reason that such confounding is adjusted for, i.e. to attempt to quantify the risk that is attributable to exposure to the landfill site alone.

Unfortunately, it is not always known which factors affect the risk of a medical condition particularly with conditions like congenital anomalies where little is known about them. Additionally, if all risk factors were known then there would be no need to investigate the health outcome. Hence, it is unlikely to be possible to adjust for all confounding. In order to assess the impact of unknown confounding on the testing procedure the data was simulated with the assumption that there were confounding factors that influence the risk of the health outcome. However, the testing procedure was performed without adjusting for confounding as it would be in a situation where the confounding factors were unknown. This was achieved by simulating with the baseline risks equal to the predicted probabilities from the logistic regression model. The testing procedures were then implemented without adjustment for confounding. Subsequently, the hypothesis tests were assessed as before using ROC curves. The ROC curve comparisons can be seen graphically in Section A1.8.

When there is unknown confounding Stone's test performs poorly with only a 30.1% (95% CI: 27.3%-33.0%) power at the 5% significance level. However, the trend test is not influenced as strongly with a corresponding power of 87.8% (95% CI: 85.6%-90.0%). The ROC curves for the two tests are compared directly in Section A1.11 and reveal that the trend

test is more able to identify a causal link between exposure to a given risk factor and the health outcome when the health outcome has a relatively unknown multifactorial aetiology.

It should be noted that if any confounding factors are known they should be adjusted for because these factors may result in a false positive result, i.e. the conclusion that there is a causal link when there is not one. A false positive result can have dire consequences for public health as it will generate unnecessary anxiety amongst the supposedly affected population and may lead to redundant interventions being put into place. However, this problem affects all the testing procedures equally and does not influence the relative performance.

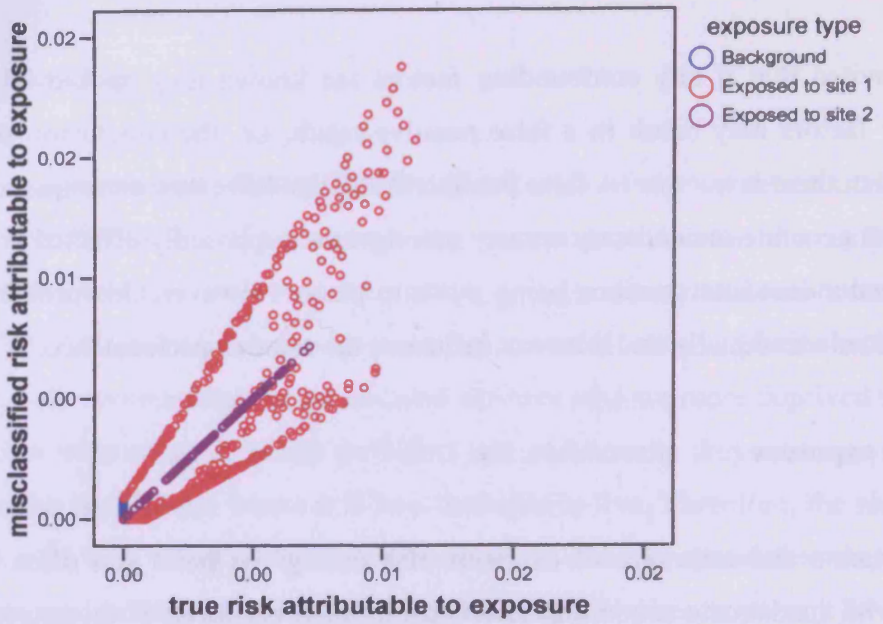
Misclassified exposure

The measurements and estimates of exposure risk at a given point can often be incorrect, particularly with inadequate modelling procedures. However, even if the exposure risk has been misclassified slightly it is still important to be able to detect a causal effect if it is present. Therefore, to assess the robustness of the testing procedure against slightly inaccurate exposure risks the performance of the test was investigated using misclassified exposure risks. Clearly, if the accuracy of the exposure risks is very poor then they should not be used in investigations.

A source of error that can occur with the exposure model given on page 97 is that the coordinates given for one of the sources of pollution could be incorrect. For example, the coordinates of the gate for Nant y Gwyddon could have been taken as the location of site 1 instead of the centre of the landfill site. Hence, when generating the misclassified exposure risks, site 1 was moved half a kilometre northwest of its original location to the coordinates of the gate.

The true exposure risks are plotted against these misclassified risks attributable to exposure in Figure 2.18. The mothers exposed to site 1 now have different risks because of the movement of the centre of this site. Some of the population has even crossed over from being in the exposed group to being unexposed and vice versa. However, those mothers exposed to site 2 have the same exposure risk under the true and incorrect classifications.

Figure 2.18: the scatterplot between true risk attributable to exposure and misclassified risk attributable to exposure



The ROC curves for the testing procedures when exposure has been misclassified in the way indicated are compared to those for when exposure is correct in Section A1.9. A misclassification of this magnitude does not greatly affect the performance of the hypothesis tests. The trend test loses some efficiency and the power is 84.6% (95% CI: 82.2%-86.7%) at the 5% significance level. However, there is little change in the ROC curve for Stone's test when exposure has been misclassified to this degree with a power of 83.5% (95% CI: 81.1%-85.7%) at the 5% significance level. The inability of Stone's test to react to this level of misclassification is a worrying point since it could give rise to false positives.

Spatial distribution of the risk factor

The spatial distribution of a risk factor of interest plays an important role in the determination of the existence of spatial association between exposure to that risk factor and the specified health outcome. If the risk factor is spatially homogeneous over the area under scrutiny then it is impossible to investigate such a relationship using information from that region. The simulations, discussed on page 97, were repeated where exposure site 2 was assumed not to exist. Therefore, the resultant data is based on a more uniform distribution of exposure risks where more mothers are assigned the baseline risk of 0.01 with more people in the unexposed region and fewer individuals in the exposed region.

The removal of the second exposure source means that the relative risk values will decrease around that source and as a result increase in the remaining regions of the area of study. Therefore, the β value had to be lowered to 0.02 in these simulations to compensate for this and ensure that the relative risk contours produced around site 1 are still of a similar size and magnitude to the observed relative risk contours for birth defects around Nant y Gwyddon. This concept is explained in detail in Chapter 3.

The ROC curve comparisons in Section A1.10 reveal that the reduction in the number of exposure sources has a negative impact on the performance of the testing procedure. The trend test has a sensitivity of 74.9% (95% CI: 72.1%-77.5%) at the 5% significance level whereas the power of Stone's test falls to even lower than that at 66.4% (95% CI: 63.4%-69.3%) at the 5% level. This drop in efficiency is a result of the additional uniformity introduced into the spatial distribution of the risk factor hindering the ease with which any spatial association can be identified. Hence, if the exposure risk is similar for all individuals in the study population then there is no point in investigating further as no relationship can be determined from it.

The ROC curves for the two tests under this more uniform spatial distribution of exposure risks are compared in Section A1.11. The difference in efficiency is even more pronounced between the two tests when there is less spatial variation in the risk factor. Therefore, if the risk factor does not vary hugely over the study region then it is advised that the trend test is used as opposed to Stone's testing procedure.

2.5 Parametric Approach to Spatial Analysis

In this section a range of statistical methods for spatially analysing health outcomes have been discussed in detail. In conclusion to this discussion, when investigating the relationship between exposure to a risk factor and a given health outcome the recommended approach is as follows. Initially, map the risk of the health outcome which if the data is provided at an area level is straight forward. However, if the data is at the individual level aggregation of data points is required either using concentric circles, when a putative point source is involved, otherwise administrative boundaries can be used. The main problem with this approach is that if the data is provided at the individual level then the choice of aggregation is arbitrary and the

boundaries used to define regions influence the appearance of the map and consequently manipulate conclusions based on it.

Once the risk has been mapped, if the data seems to have a spatially heterogeneous distribution, then the likelihood based test for homogeneity of risk can be applied, as described in Section 2.3, to determine whether this spatial pattern can be explained by random variation. However, as with the mapping procedure this hypothesis test requires the data to be at area level. If the data is provided at an individual level then the data will have to be aggregated and the choice of aggregation is arbitrary. The boundaries chosen for the regions could influence the conclusions made. For example, there may be a non-random, heterogeneous spatial distribution of risk within one of the regions but across that region the risk can only be depicted as uniform therefore the test will not detect it.

After results have been obtained from the homogeneity test, if the conclusion is that the spatial patterns are unlikely to be explained by random error then possible risk factors should be considered and spatial association tested. The concentric circle method is the most commonly used approach to assessing the impact of exposure to landfill on the risk of birth defects. However this method is overly simplistic. The use of only two regions in the analysis means that little information can be gathered on the dose-response relationship and any variation that occurs within either of the two regions will be overlooked. Additionally, the choice of boundary for the two regions influences the result. Therefore, even though it may be the only option available to investigators, when the geographical information linked to the health outcome is not of a high accuracy, the concentric circle approach is not recommended here.

Stone's test, as discussed in Section 2.4.1, must be used if the data provided is at area level. However, if the data is given as a series of individuals each with a pinpointed spatial location, as is the case for the birth defect data used above, then the trend test for spatial association (see Section 2.4.2) should be considered as it is more powerful under a wide range of situations as discussed in Sections 2.4.3 and 2.4.4. The test for homogeneity, Stone's test and the trend test all are parametric and, consequently, assume a standard statistical distribution for the health outcome. This assumption may well be correct but the estimated parameters may be wrong and, hence, the conclusions drawn from the testing procedures will also be inaccurate.

In James et al.'s paper an alternative approach of investigating the risk of congenital malformations around landfill sites is considered [10]. The risk is mapped using kernel density methodology which requires individual level data but does not involve any form of aggregation nor is a statistical distribution assumed. Therefore, the difficulty of arbitrary boundary choice is overcome whilst preserving the non-identifiable nature of the mapping of risk. The paper only maps risk but the possibility of evolving the methodology to test for the homogeneity of risk and the spatial association between a health outcome and exposure to a risk factor is mentioned in its discussion. Thus, in the next Chapter, the kernel density methodology is considered and developed to provide a complete alternative to the process of spatially investigating a health outcome, as discussed in this Chapter.

Chapter Three: Non-Parametric Methodology for Spatial Analysis

In the previous Chapter a parametric approach to the spatial analysis of a health outcome with consideration of possible risk factors was presented. In this Chapter a non-parametric alternative is considered for each step of the spatial analysis of a health outcome. These steps are as follows:

1. mapping the risk of the health outcome over the study region;
2. testing for spatial homogeneity; and
3. testing for the spatial association between the risk of the health outcome and putative risk factors.

The first step of mapping the risk of the health outcome can be achieved non-parametrically using kernel density estimation, a non-parametric method of estimating the probability density function of a specified variable. Kernel density methods can be used to generate surfaces for the density of observed cases of a given health outcome and also the surface for the density of expected cases. The ratio of these two surfaces gives the relative risk surface for the health outcome of interest. Contours can be mapped based on the relative risk surfaces indicating regions over which the observed risk of cases is greater than the expected risk. The existence of regions over which the relative risk surfaces are elevated suggests that there is spatial heterogeneity in the risk of the health outcome. Kernel density methodology has been used to map congenital malformations [10], cancer [185] and sex ratios [186]. Kernel density estimation and the generation of relative risk contours are discussed and the risk of congenital malformations around the Nant y Gwyddon landfill site is mapped in this Chapter.

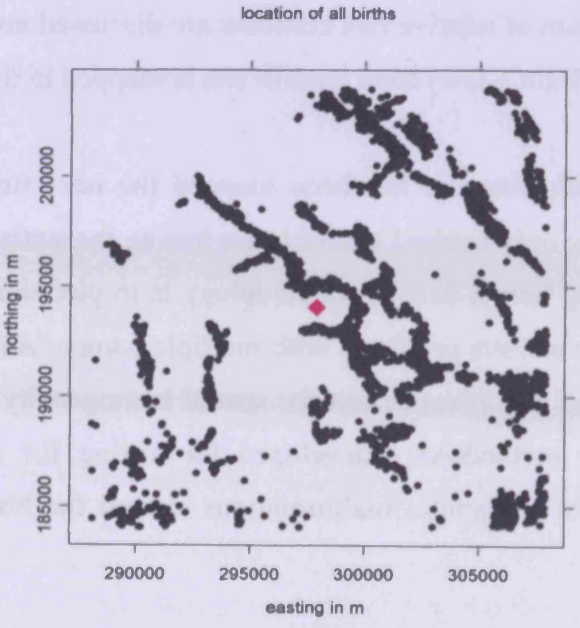
Once the risk of the health outcome has been mapped the next step is to test for spatial homogeneity. Currently the only method available for testing the statistical significance of the spatial pattern of risk using kernel density methodology is to plot significance contours. The significance contours have serious problems with multiple comparisons and do not provide a p-value. Therefore, a global significance test for spatial homogeneity has been developed as part of this project. The methodology developed for testing for spatial homogeneity is illustrated using the risk of congenital malformations around the Nant y Gwyddon landfill site.

If the global significance test provides evidence of spatial heterogeneity in the risk of the health outcome then the next step in the spatial analysis is to test for a spatial association between the health outcome and possible risk factors. In this thesis a new methodology is developed for testing the spatial association between the health outcome and a risk factor that is based on kernel density methodology. The performance of the test is assessed and compared with that of the parametric methodology discussed in the previous Chapter using ROC curves. The ROC curves are generated using the simulations for congenital malformations around the Nant y Gwyddon landfill site, as performed in Chapter 2.

3.1 Mapping Risk

The first step in the spatial analysis of a health outcome is to map the risk. The risk is a function of the study population and the cases of the disease. Figure 3.1 indicates the location of the study population by means of a scatterplot and includes all births in the 20km square area centred over the Nant y Gwyddon landfill site represented by the pink diamond. The pixilation on the graphics produced in SPlus and the use of postcodes to determine the location of each of the births mean that many of the plotted points overlap each other so it is unclear what the density of births is in some areas. This lack of clarity can be overcome by density estimation.

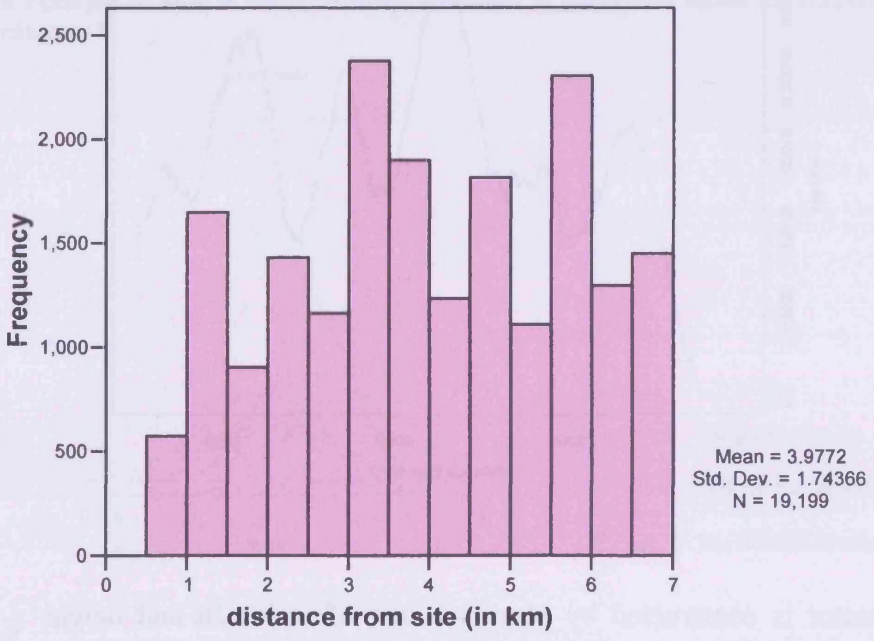
Figure 3.1: the location of all births around Nant y Gwyddon



A density estimate of the probability density function $f(x)$ is the result of applying any method that estimates $f(x)$ for any given x. A parametric density estimate is generated by assuming a parametric form and subsequently estimating the parameters required to define the density. For example, the Gaussian density is defined by μ and σ, which are estimated by the sample mean and sample standard deviation, denoted by $\bar{x}$ and s, respectively. In contrast, the form of a non-parametric estimate is determined by the data itself and no parametric form is assumed.

There are three commonly used non-parametric density estimators: histograms, naïve estimators and kernel density estimators. The mathematics behind these density estimators and their limitations are discussed further in Appendix 2.

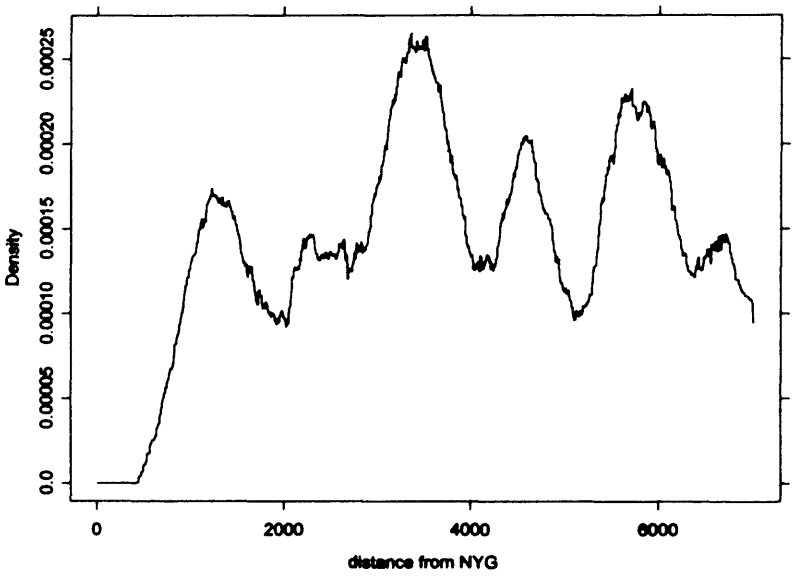
Figure 3.2: the histogram for the distance of births from Nant y Gwyddon in the circular region with a 7km radius centred over the Nant y Gwyddon site (origin = 0km, number of bins = 14)



Histograms are the oldest form of density estimator and are still the widest used. Typically they provide some indication of the shape of the underlying density function by partitioning the variable of interest into intervals of width h, called the bandwidth, and displaying the counts of the observations that lie in each interval. An example of a histogram is given in Figure 3.2 and was produced in SPSS where the variable presented on the y-axis is the count

in each interval. Despite the use of this construction method in some statistical computing packages, the typical approach is to display the relative density on the y-axis which is calculated by dividing the count for each interval by the bandwidth. Thus in the resultant histogram the area in each bar represents the count in the corresponding interval and the area underneath the histogram is equal to the sample size. However, this approach to constructing histograms does not result in a true density estimator and the variable represented on the y-axis should be the relative density divided by the total sample size so that the total area beneath the histogram is constrained to 1. Unfortunately, histograms pose a number of problems as they are affected by the origin and bandwidth and the lack of mathematical form provides no means to estimate an optimal bandwidth. Additionally, discontinuities in the profile cause difficulties if derivatives are required.

Figure 3.3: the naïve estimator for the distance of births from the centre of Nant y Gwyddon in the circular area with radius 7km centred over the site (h = 300)

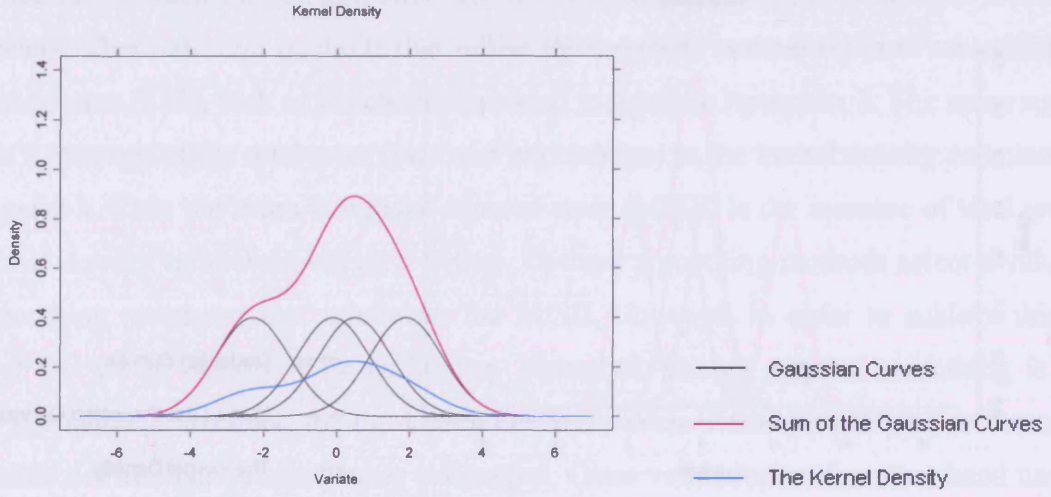


A naïve estimator is constructed by placing a box of width 2h and height $\frac{1}{2nh}$ on each observation and for any given x value summing the contributions from each box to obtain the estimate. The boxes' width and height are chosen so that when the heights of the boxes are summed at each point along the x-axis the total area beneath the resultant curve is equal to 1, as required for an estimator of the probability density function. Naïve estimators overcome the problem of the arbitrary nature of the origin as it is now determined by the data. However, the naïve estimator is very jagged in appearance making it difficult to compare with other naïve

estimators and the choice of bandwidth is still arbitrary. An example of a naïve estimator is given in Figure 3.3.

Kernel density estimators are similar to naïve estimators, however, rather than boxes, unimodal, symmetric curves (e.g. Gaussian curves) are the foundation of their construction. Figure 3.4 illustrates the construction of a kernel density estimator based on Gaussian curves centred over each observation, shown in black, with a standard deviation equal to the value of the smoothing parameter, h, here equal to 1. The sum of the heights of the component kernel densities, taken at each point on the x-axis, is shown in pink. The total height is then divided by the number of observations to give the average height of the curves across the x-axis, this is the kernel density estimator and is displayed in blue. It is clear that although the component kernel densities are Gaussian in form, the resultant density estimator is determined by the observed points and therefore can take any shape. Thus, kernel density estimators are non-parametric in nature.

Figure 3.4: an example of how a kernel density estimator is generated based on Gaussian curves with smoothing parameter 1



The ideal smoothing parameter should be the smallest value that smoothes sufficiently. If the smoothing parameter is too high it over smoothes hiding detail by increasing bias, which is caused by displacing density from the peaks into the troughs. An example of this can be seen in Figure 3.5 which displays the same information as Figure 3.4 but with a smoothing parameter of 1.5. The Gaussian curves have a larger variance so they are wider and have

lower peaks. Thus, the result is that the kernel density has a lower peak and higher tails than in Figure 3.4 where the smoothing parameter is only 1.

Figure 3.5: an example of how a kernel density estimator is generated based on Gaussian curves with smoothing parameter 1.5

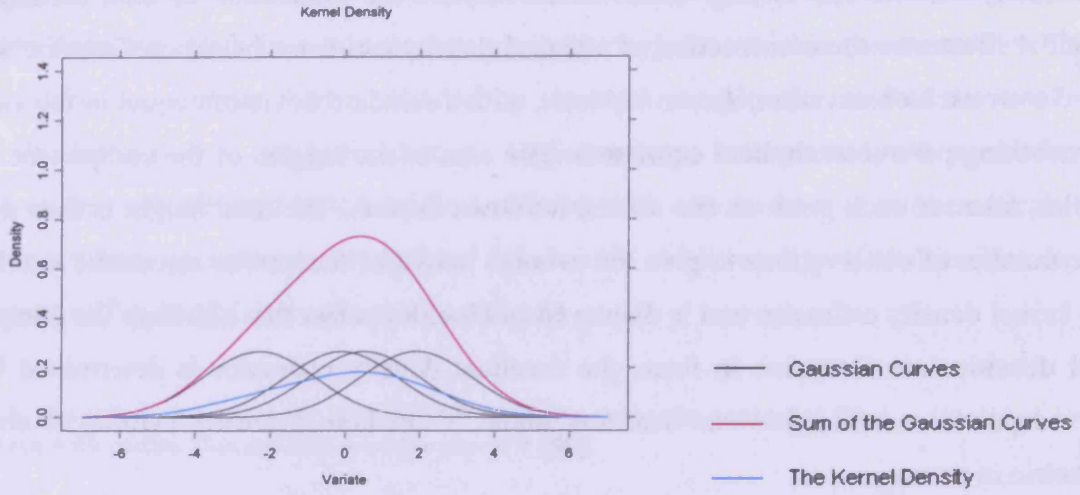
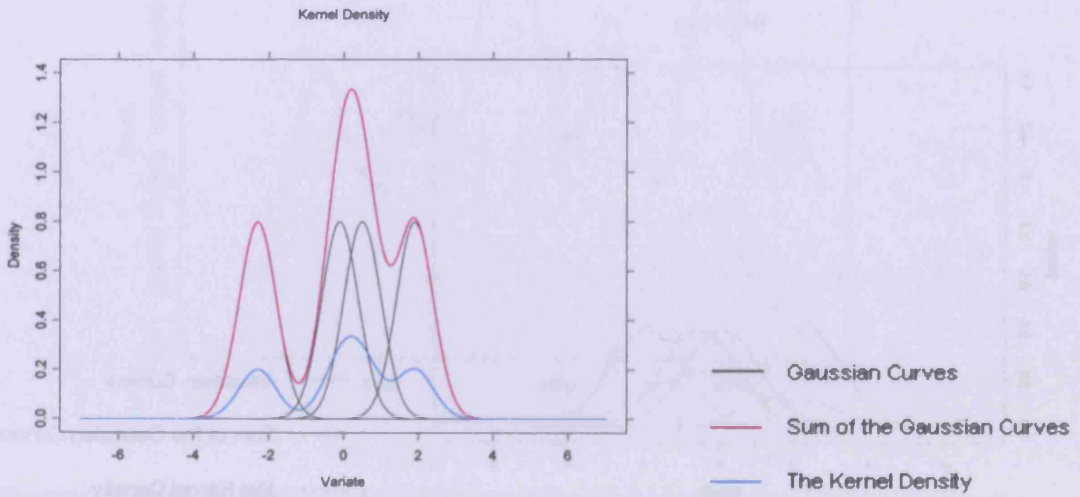


Figure 3.6: an example of how a kernel density estimator is generated based on Gaussian curves with smoothing parameter 0.5



The kernel density estimator in Figure 3.4 shows more detail, whereas the higher smoothing parameter reduces the level of detail in Figure 3.5. However, if the smoothing parameter is too low the kernel density estimator shows too much detail by increasing variability. Figure 3.6 displays the same data as Figure 3.4 but the smoothing parameter is only 0.5. The Gaussian curves have a smaller variance making them narrower with higher peaks. Therefore, the curves do not overlap as much between observations so the kernel density reveals more of the

variation in the data, i.e. too much structure is seen. Figure 3.6 has three peaks whereas Figure 3.4 has one peak with a shoulder, hence, it is possible that some of the features in Figure 3.6 are not genuine and only a result of random noise.

If the smoothing parameter chosen is too large then the kernel density is over smoothed and density is redistributed from the peaks into the troughs. Conversely, if it is too small then the kernel density presents too much detail making it difficult to see the fundamental structure of the true density.

There are two types of error that can occur in density estimation: bias and imprecision. The level of bias present in the density estimator is determined by the degree to which the density is redistributed from the peaks into the troughs. If the bias is too high this is referred to as over smoothing and is associated with a high smoothing parameter. The imprecision is reflected in the noise of the density estimator. If the imprecision is too high there is under smoothing as a result of a low smoothing parameter.

The trade off between imprecision and bias can be used to determine the optimal smoothing parameters. There are two methods that utilise this concept: normal optimal smoothing and cross validation [183], both of which are discussed in detail in Appendix 2. The mean squared error is a measure of the total error (bias and imprecision) in the kernel density estimator at a given point x. Thus the mean integrated squared error (MISE) is the measure of total error in the kernel density estimator over all x values. Optimal smoothing methods select a value for the smoothing parameter that minimises the MISE. However, in order to achieve this, the form of the kernel density must be known. Therefore, normal optimal smoothing is used where normality is assumed for the kernel density, which if true would lead to the use of parametric not non-parametric density estimation. Cross validation on the other hand uses the integrated squared error instead of the MISE. Both cross validation and normal optimal smoothing provide a global optimum value which may result in undesirable local features. Thus the values given for the optimal smoothing parameter from both these methods are treated as guidelines.

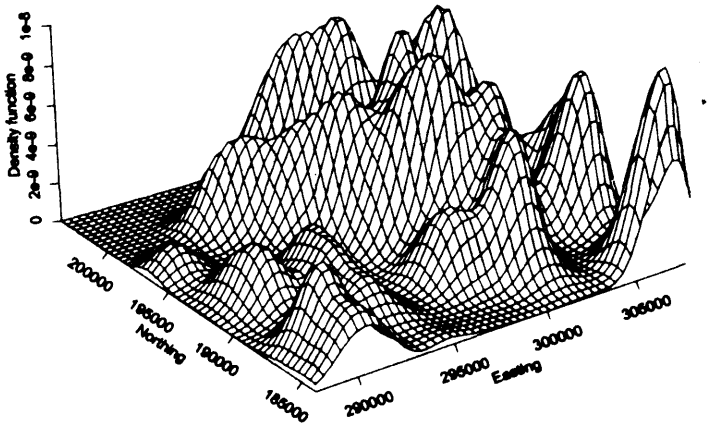
Therefore, in order to select a smoothing parameter used to generate the kernel density the values for the optimal smoothing parameter are calculated using both normal optimal smoothing and cross validation. Then different values for the smoothing parameter that are

close to those given in the normal optimal smoothing and cross validation approaches are used to generate kernel densities. The kernel densities for the different smoothing parameters are compared and the most desirable plot that shows sufficient detail but not too much variability is then selected giving the optimal smoothing parameter. This is a subjective method but it is a good way of selecting the smallest possible smoothing parameter that shows sufficient structural detail. An example of this procedure is provided in Appendix 2.

An alternative approach is to have a variable bandwidth which is low in regions where there is a lot of data and high in regions where there is less data. The umbrella term for such approaches is adaptive techniques and not only does this approach reduce the bias but it avoids the problems of using a globally appropriate smoothing parameter to address a localised problem.

The location of all births around Nant y Gwyddon based on the east and north grid coordinates can be treated as a bivariate distribution of points. This bivariate distribution can be estimated using a two-dimensional kernel density surface where it is assumed that the two variables are independent [184]. The method of construction used here for a two-dimensional kernel density comprises plotting the observations on a graph of the east gird coordinate versus the north grid coordinate, placing a bivariate Gaussian surface over each observation and averaging the heights of these surfaces at each point.

Figure 3.7: the perspective plot of the kernel density surface for the location of all births around Nant y Gwyddon



A two-dimensional kernel density, using the east and the north grid coordinates of each of the births as the two explanatory variables, as seen in Figure 3.1, produces peaks in areas with a relatively high density of births. Figure 3.7 is a perspective plot of the kernel density surface for the location of all births generated with a smoothing parameter of 0.5km for both the east and the north grid coordinates. 0.5km is the smoothing parameter as used in the paper by James et al. which was selected using the method explained above [10]. In the case of geographical location this equal smoothing of the east and north grid coordinates is desirable, however, in other situations this is not the case as seen later with the kernel densities used in the non-parametric discriminant analysis.

The peaks in Figure 3.7 are higher where there is a dense region of points in Figure 3.1, as was expected. However, this three-dimensional method of plotting a kernel density provides a poor visual representation of the information it summarises. A contour plot based on a bivariate kernel density can be produced by drawing contours where the density height is equal to 25%, 50% and 75% of the maximum height of the kernel density, labelled 25, 50, and 75, respectively. Contours for kernel densities are referred to as probability contours. This form of presentation presents a much clearer picture of population density patterns.

Figure 3.8: the probability contours for all births around Nant y Gwyddon

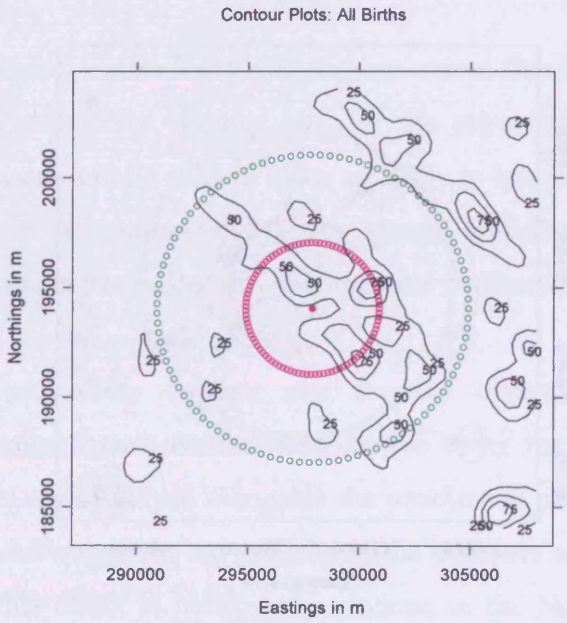


Figure 3.8 is the probability contour plot based on the kernel density for the location of all births, shown in Figure 3.7. The contours clearly indicate that the areas of high population density form a ribbon-like pattern near Nant y Gwyddon with some smaller areas of denser population in the surrounding area. The probability contours reproduce, as expected, the pattern shown by the areas with a high density of points in Figure 3.1 and it displays the information more clearly. The pink dot indicates the centre of the Nant y Gwyddon site and the pink and green circles indicate distances of 3km and 7km from the site centre, respectively.

Clearly, the distribution of the population around Nant y Gwyddon is not uniform. Therefore, if landfill has no influence on the risk of congenital anomalies then there will still be a pattern in the distribution of congenital malformations induced by the population pattern alone. In other words, there will be areas with higher numbers of birth defects where there is a higher population density. In addition, if there is a causal effect there still may not be a peak in the density surface for congenital anomalies over the Nant y Gwyddon site itself because Figure 3.1 clearly shows that there is no population there.

Figure 3.9: the probability contour plot for observed congenital anomalies as defined by SAHSU around the Nant y Gwyddon site

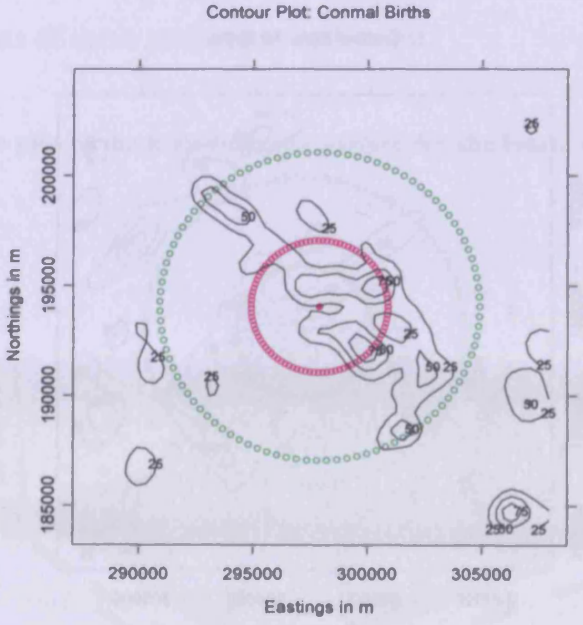


Figure 3.9 is the probability contour plot for the congenital anomalies around Nant y Gwyddon. The underlying kernel density is generated as before, but the bivariate Gaussian

surfaces are formed only for those mothers who gave birth to a baby with a congenital anomaly using the same criteria as that used by SAHSU [68]. Clearly, the probability contours for birth defects closely follow those for all births therefore population is a major factor in the pattern of the contours.

The number of congenital malformations is smaller than the total number of births and this smaller sample means that a larger smoothing parameter than 0.5km might be appropriate (this can be deduced from equation A2.1 in Appendix 2). However, when comparing two kernel density surfaces it is recommended that the smoothing parameters should be equal so that the two surfaces are equally smoothed [183]. This enables ease of comparison as the size of the smoothing parameter is not a factor in the difference between the two plots. Therefore, the use of the smoothing parameter of 0.5km is maintained.

In order to identify whether there is an increased risk of birth defects surrounding the Nant y Gwyddon site, the probability contour plot for the expected density of congenital anomalies needs to be mapped. The probability contours for the expected density of congenital anomalies are determined by the underlying population pattern and the spatial pattern of known confounders. Therefore, the kernel density of expected congenital malformations is estimated by that of all births with adjustment for known confounders.

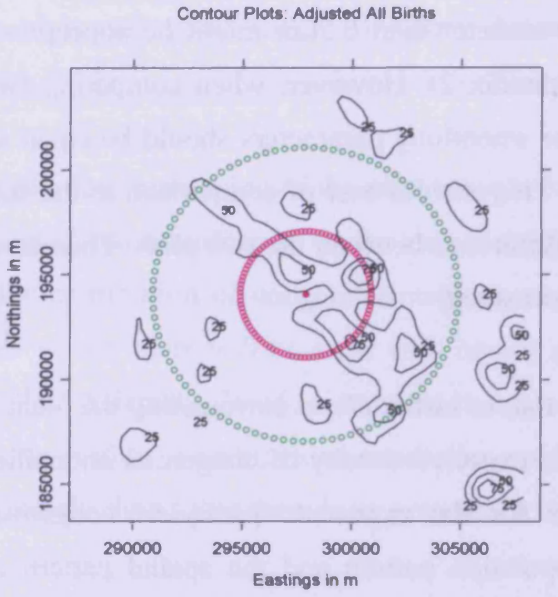
The predicted probabilities from the logistic regression model fitted in Chapter 2 were used as weights to generate a probability contour plot for the expected density of birth defects. Weights were used because a child who is twice as likely to have a congenital malformation based on their exposure to confounding factors should contribute twice as much to the probability contours of the expected density of congenital malformations.

Figure 3.10a is the probability contour plot for the expected density of congenital malformations adjusted for known confounding factors using the method described above. These adjusted contours are displayed alongside the unadjusted contours presented in Figure 3.10b. There are clear differences in the patterns of the contours revealing the impact of the confounding factors. This effect is particularly obvious in the North East which relates to births in the Prince Charles hospital that has a lower reported rate of congenital anomalies than East Glamorgan and the Princess of Wales hospitals where the majority of the remaining

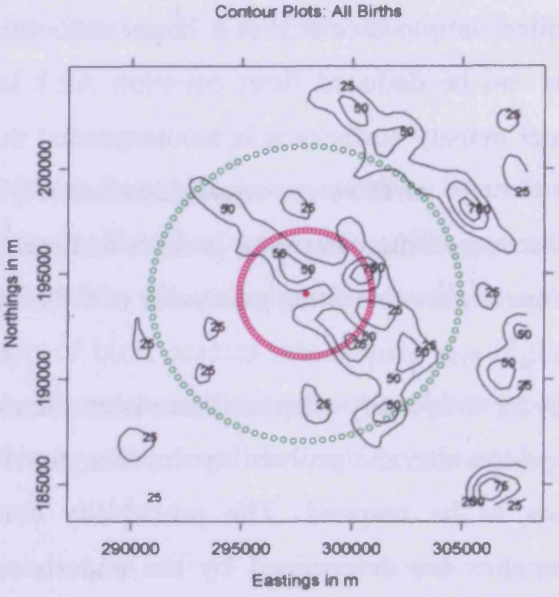
births were reported. The impact of hospital of birth on the risk of congenital anomalies is caused by different attitudes towards reporting.

Figure 3.10: the probability contours for a) the expected density of congenital malformations and b) the density of all births around Nant y Gwyddon

a)



b)



The aim in this subsection is to generate a method of mapping the risk of a health outcome over a geographical region and, in particular, mapping the risk of birth defects around Nant y Gwyddon. Comparing the kernel densities for the location of the congenital anomalies and all births adjusted for confounding factors will indicate where there are more congenital anomalies than expected, i.e. whether there is any excess that might be attributable to proximity to Nant y Gwyddon. A contour plot which summarises the spatial properties of areas where there is an elevated risk of congenital anomalies would be useful. This information can be illustrated by generating the relative risk surface. Suppose the observed density of congenital malformations is denoted by the function f and the expected density of congenital malformations is denoted by the function g. Then the relative risk at point y can be defined as

$$RR(y) = \frac{f(y)}{g(y)}.$$

If the height of the relative risk surface is equal to 1 then the observed density of birth defects in that area is the same as the expected density. However, if the height of the relative risk surface is above 1 then the density of birth defects is greater than expected and if it is below 1 then the density of birth defects is less than expected. The relative risk surface is independent of the size of the local population since the densities are standardised for sample size, because by definition the two kernel densities defining the ratio must integrate to 1.

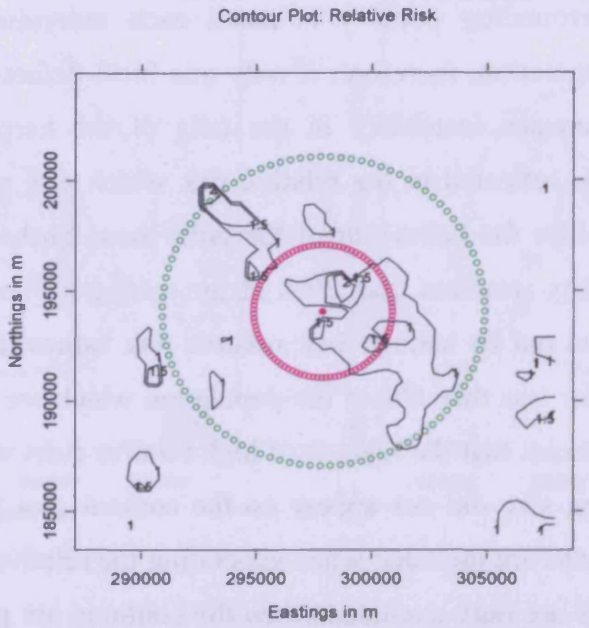
If a local population surrounding point y is small each individual represents a larger proportion of the local population, therefore, if only one birth defect occurs it will create a high relative risk. This creates instability in the tails of the kernel density, where the population is lower, and is reflected in the relative risk which will possess greater random variation. In order to stabilise the behaviour of the tails, those births that occurred in areas where the population density was less than 20% of the maximum height for the population kernel density were filtered out by setting their relative risk values to zero. 20% may seem high but it only accounts for less than 8% of the population which are distributed in the main in more rural areas. This meant that the regions of high relative risks which occurred because of a small local population size did not appear on the contour plot for the relative risk. It should be noted that all births are included when generating the relative risk surface and those in sparsely populated areas are only excluded when the contours are plotted. The filtering of data by population density is discussed in more detail later on.

Figure 3.11 is the relative risk contour plot obtained from the ratio of the expected density of congenital malformations adjusted for confounding factors compared to the observed density of congenital anomalies. The relative risk contours are plotted along points where the relative risk is equal to 1, 1.5 and 2: this is because 1 is the point where the density of birth defects is as expected, 2 is where the density is twice that expected, which is a big increase, and 1.5 is the midpoint between the two. All three contours give a clear indication of the rate of change of the risk surface.

The relative risk contours in Figure 3.11 indicate that in some regions there is a greater risk of a birth defect in relation to other areas. There is an area within 3km of the centre of Nant y Gwyddon where the relative risk reaches over 2, indicating that there is a large increase in risk of congenital malformations near to the site. Areas that have this property of elevated relative risk are referred to as hotspots. There are other areas further from the site where the relative

risk is also high. However, excluding the hotspot in the North West these regions of high relative risk are much smaller in extent and magnitude than those nearer the site. The hotspot in the North West may be the result of a secondary unknown source of additional risk. The larger the defined area of relative risk and the greater the magnitude of the relative risk values in that area, the less likely it is to be a random event.

Figure 3.11: the relative risk contours for congenital malformations around the Nant y Gwyddon site



The relative risk contours provide a powerful tool for mapping the risk of a given health outcome, which is more sophisticated than mapping the SER's as discussed in Chapter 2. The contours are not constrained by any form of aggregation, which in Chapter 2 was found to potentially influence the conclusions drawn from the mapping process. Both the relative risk and the map of SER's reveal that the risk of congenital malformations increases towards the centre of Nant y Gwyddon. However, the map of SER's only indicates that the risk of a birth defect is elevated within 1.5km of the centre of the site and between 6.5 and 7km from the landfill whereas the relative risk contours provide accurate and spatially defined locations for areas over which the risk is elevated. The other advantage to using the kernel density method is that it can be used regardless of whether or not there is a putative point source of pollution involved, hence, it is ideal for use with the sex ratio example which is the focus of Chapter 5. Therefore, if the data is provided as a series of individuals each with their own geographical location the kernel density approach to mapping risk is recommended.

It has been ascertained from the relative risk contours that the risk of congenital malformations is not spatially homogeneous. However, as with the SER maps it is not clear from the mapping procedure alone whether it is likely that these patterns are a result of random variation or not. Therefore, the kernel density method needs to be developed to test for the homogeneity of risk using specified levels of significance.

3.2 Testing for Homogeneity of Risk

The second stage in the spatial analysis of the risk of a health outcome, after the risk has been mapped, is to determine whether or not these patterns are merely the result of random variation. If they are unlikely to be random events then there may be some risk factor that influences the health outcome which explains the spatial patterns observed. In Chapter 2 a test for the homogeneity of risk was presented based on likelihood ratio methodology, however, this can be achieved using kernel density techniques as well.

3.2.1 Significance Contours

One method of determining where the risk of congenital malformations is significantly different to that expected is to generate significance contours [183]. The significance contours are generated by taking the difference of the square root of the density estimates of the observed and the expected birth defects and standardising them by dividing by an approximate standard deviation term i.e. the spatial significance between the estimated densities $\hat{f}$ and $\hat{g}$ at a point x is assessed by the function below

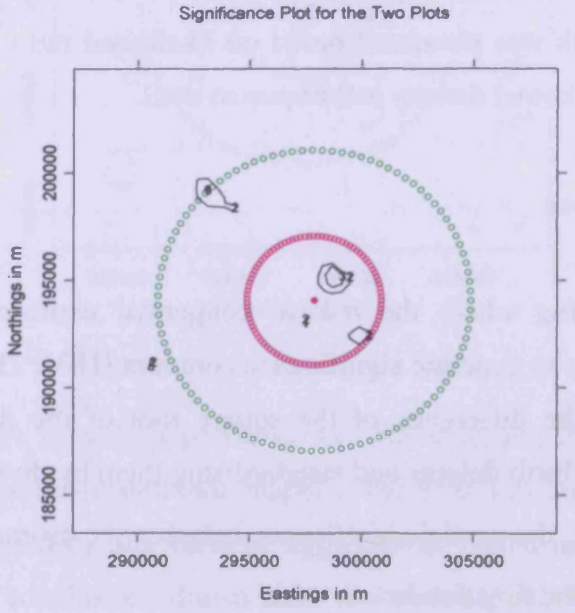
$$significance = \frac{\sqrt{\hat{f}(x)} - \sqrt{\hat{g}(x)}}{\sigma}.$$

This standardising process gives the differences in standard deviation units. The square root of the density estimate has useful statistical properties as revealed by the algebra presented in Section A3.3. The standardised differences can be plotted in what are referred to as the significance contours, which indicate the regions where the expected and observed densities of congenital malformations are significantly different. The significance contours are plotted along points where the difference between the two surfaces is equal to 2 and 3 standard

deviations, in a direction indicating more birth defects than expected. These contours are labelled 2 and 3, respectively.

Figure 3.12 is the significance contour plot for congenital malformations around the Nant y Gwyddon landfill site. There are two large hotspots shown on the plot within 3km of the centre of the site indicating that around those points the risk of birth defects is significantly larger than expected. However, there is another large hotspot that lies outside this 3km region. The existence of regions where the risk of congenital malformations is significantly elevated suggests that the risk of birth defects is not homogeneously distributed and that this spatial pattern of risk is unlikely to be a random event.

Figure 3.12: the significance contours for the difference between the expected and observed density of birth defects around Nant y Gwyddon



Unfortunately, there is a substantial problem in determining whether or not spatial patterns are random events using significance contours in this way. Significance contours are based on standardised differences that are calculated at each point and do not represent a global measure of significance. These calculations of standardised differences at multiple points lead to problems with multiple comparisons. Thus, care needs to be exercised when using significance contours because the risk of getting a false positive result could be high. However, an alternative method of determining whether or not spatially heterogeneous risk is

a random event has been developed that overcomes the problem of multiple comparisons to some extent and this method is described below.

3.2.2 Global Significance Test for Regions of Excess Relative Risk

The global significance test is used to determine whether a region of elevated risk is likely to be a random event or not. If the probability of such a hotspot occurring at random (p-value) is small (conventionally < 0.05) then the hotspot is unlikely to be a result of random variation and may be the result of a risk factor influencing the health outcome. If there are hotspots that could be a product of some risk factor then the spatially heterogeneous risk is unlikely to be a random event. The focus on the hotspot as a whole as opposed to individual points as considered when using significance contours overcomes the problem of multiple comparisons. In addition, this testing procedure is global and not based on point measures, hence, the term global significance test.

The initial step in the global significance test is to produce relative risk contours for the health outcome under investigation to detect regions of excess risk, which has already been achieved as part of the first stage proposed in the spatial analysis. A threshold relative risk (r) is then chosen and any region over which the relative risk is greater than r is considered to be a hotspot. 1000 simulations of cases are performed for the study population under the null hypothesis, that is when the risk of the health outcome is equal to the predicted probability from the logistic regression model fitted with all known risk factors. The dataset containing the simulations is tagged to identify individuals within each of the defined hotspots.

There are three potential global significance tests each of which has its own strengths and weaknesses. Test 1 is based on whether the relative risks for all individuals living within the boundary of the hotspot are greater than the chosen threshold relative risk (r). The p-value for test 1 is the proportion of simulations under the null hypothesis for which this is true. Test 1 is influenced primarily by the size of the hotspot because as the hotspot increases in size the chance of getting a risk surface which is uniformly greater than the chosen threshold becomes progressively less likely.

Test 2 is based on the mean relative risk amongst individuals in the hotspot under investigation. Therefore, the proportion of simulations for which the mean relative risk value

of individuals living within the identified region exceeds that determined from the observed data provides the p-value for test 2. Test 2 is not as heavily influenced by the size of the region and allows for more variation of relative risks within the hotspot boundary. However the higher the observed relative risk is for each individual within the boundary of the hotspot the less likely it is that the simulated mean relative risk is higher and hence a small p-value is expected.

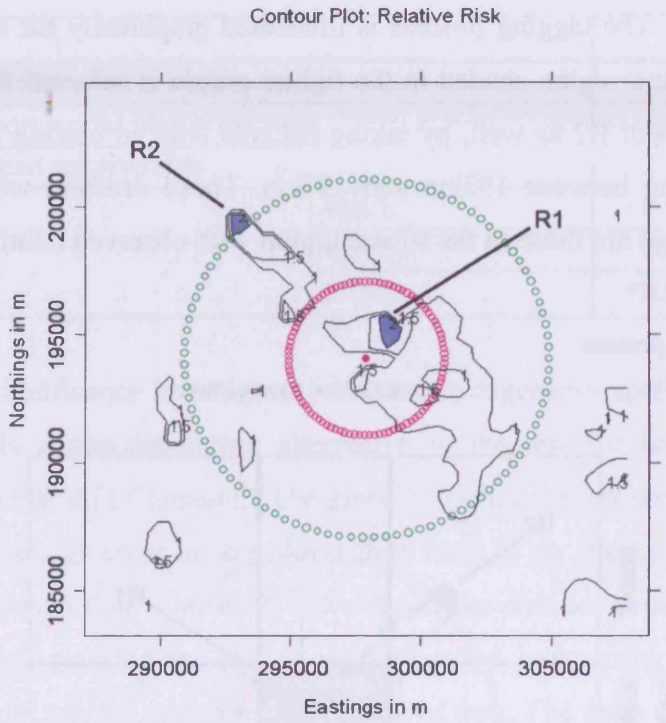
Finally, test 3 involves the number of cases of the health outcome within the boundary of the hotspot. Thus, the p-value is equal to the proportion of simulations for which the number of congenital malformations exceeds that determined from the observed data. Clearly, test 3 is not influenced by the risk surface other than the fact that the risk surface is used to determine the boundary of the hotspot. However, test 3 is important as it is controlled by the number of cases within the hotspot. This emphasis on the number of cases ensures that a hotspot that occurs as a result of a low population density as opposed to a high density of cases does not give a significant result.

All three tests are required to determine whether a hotspot is globally significant or not since a combination of all three considers the size of the hotspot, magnitude of the relative risks within the hotspot and whether these relative risks are consistently high. If all three tests have positive results (p-value < 0.05) then conclusions from the global significance test is that the hotspot under investigation is unlikely to be a random event, in which case, there may be a risk factor that influences the health outcome that has yet to be considered. If at least one of the tests has a p-value greater than 0.05, then the corresponding hotspot is not considered to be statistically significant. The true significance level lies between 0.05^3 and 0.05 because of the criterion that all three p-values from the three tests must be less than 0.05. The combination of the three different tests to produce the global significance test means that large hotspots over which the relative risk values are of a high magnitude and within which there are a large number of cases are more likely to give a significant result, as desired.

3.2.3 Applying the Global Significance Test for Spatial Homogeneity to Nant y Gwyddon

The global significance test is illustrated below with the two most prominent hotspots in the vicinity of the Nant y Gwyddon site with relative risks that exceed 2, labelled R1 and R2. The locations of these hotspots are marked in Figure 3.13.

Figure 3.13: the location of the hotspot R1 and R2



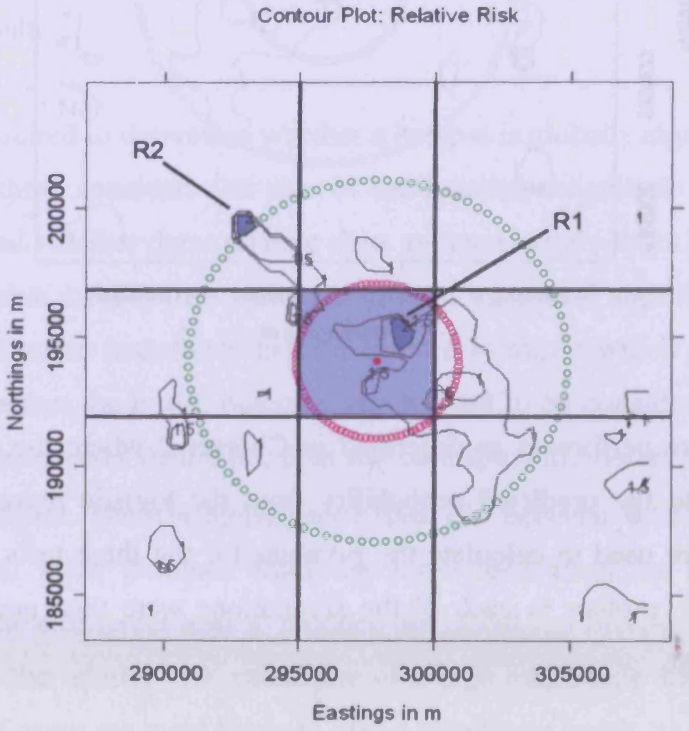
1000 simulations were performed as described in Chapter 2 where the risk of a congenital anomaly was equal to the predicted probability from the logistic regression model. These 1000 simulations were used to calculate the p-values for the three tests. The corresponding relative risks for each mother in each of the simulations were then needed to perform the global significance test.

The program for creating relative risk contours involves selecting a series of grid points and subsequently calculating the height of the kernel density surfaces for the simulated congenital anomalies and that of the expected congenital malformations, here assumed to be that of all births, at the location of each grid point. The height of the relative risk surface at each grid point is then determined by dividing the corresponding density for the simulated birth defects by that of the expected birth defects. This program can be modified so that rather than

calculating the height of the relative risk surface at each grid point the height is determined at the location of the residence of each mother in the population. Therefore, the output for one simulation can be converted to a series of relative risks for each of the births in the population calculated from the simulated cases. This calculation of relative risks was performed for the observed data and for the 1000 simulations.

The dataset was tagged to mark which of the births are located within the boundary of hotspot R1 and those for R2. The tagging process is illustrated graphically for hotspot R1 below in Figure 3.14. The square region shaded in the lighter purple is selected first to avoid tagging births that are in hotspot R2 as well, by taking patients with an easting between 295km and 300km and a northing between 192km and 197km. Those mothers who reside within the boundary of the hotspot are those in the square region with observed relative risk above 2.

Figure 3.14: the tagging process



Once the dataset has been tagged it is possible to calculate the p-value for each of the three tests. The results of the three tests are given in Table 3.1 for hotspots R1 and R2. The p-values are estimated using 1000 simulations and consequently there is sampling error in these estimates. The confidence interval for a p-value of 0.006 based on 1000 simulations is (0.0036, 0.0084). All three p-values are well below 0.05 for both hotspots. Therefore, the

hotspots are unlikely to have occurred as a result of random variation. Hence, there is evidence that the non-homogeneous spatial pattern is not random, which motivates the need to investigate possible risk factors that could explain the distribution of risk that has been observed.

Table 3.1: the results of the global significance test for hotspots R1 and R2 of excess risk of congenital malformations around Nant y Gwyddon

Hotspot		R1	R2
Number of births		359	310
Number of congenital malformations		16	11
Observed mean relative risk		2.222	2.100
p-value	test 1	<0.001	0.001
	test 2	0.001	0.001
	test 3	<0.001	0.006

The test for global significance investigates whether heterogeneous spatial patterns in risk are random events. It is a non-parametric alternative to the test for homogeneity based on likelihood ratios described in Chapter 2. The global significance test requires individual level data and if the data are given at an area level then there is no alternative but to follow the testing procedure discussed in Chapter 2. However, if the data are provided at an individual level then the global significance test is recommended over using the likelihood ratio approach as it does not rely on arbitrary aggregation of data. The main advantage of using the global significance test is that the relative risk contours required in its application are powerful tools for mapping risk and surpass the use of SER's calculated over pre-specified regions to determine the spatial distribution of the risk of a given health outcome.

The result of the global significance test suggests that there may be risk factors other than those already accounted for in the logistic regression model that influence the risk of birth defects. In Chapter 2 a method for testing for spatial association between exposure to a given risk factor and a health outcome is discussed. In the next section an alternative approach is considered using kernel density methodology.

3.3 Testing for a Spatial Association between Exposure and the Health Outcome

In the previous section it was concluded that there is likely to be a risk factor, other than those five that have already been adjusted for, that influences the spatial pattern of risk for birth defects. One possible risk factor that is discussed extensively in the literature is exposure to landfill. Therefore, the next step is to test the hypothesis that landfill sites influence the risk of congenital malformations. Kernel density techniques have been utilised to generate a non-parametric procedure for testing for spatial association between the health outcome and the risk factor and this hypothesis test is discussed in detail in this section. The efficiency of the kernel density test is assessed using ROC curves. The performance of the kernel density test is compared with that of the standard testing procedure as outlined by Stone and the trend test for spatial association developed in Chapter 2. Real exposure data unfortunately was not forthcoming and so the methodology was developed by means of a simulation study.

3.3.1 Simulating Data

The model for exposure outlined in Chapter 2 was considered where the total risk for birth i of having a congenital malformation is

$$p_i = \begin{cases} \alpha_i + \beta\left[\sum_{\forall j}\{\exp(-d_{ij}) - \exp(-3)\}\right] & \text{if } d_{ij} \le 3 \\ \alpha_i & \text{otherwise} \end{cases}$$

where α_i is the baseline probability of a birth defect, β controls the strength of the risk attributable to exposure and d_{ij} is the distance of individual i from site j. This model was used to simulate cases under the null and the alternative hypotheses so that the methodology could be developed. The 1000 simulations that were generated under the null and alternative hypotheses in Section 2.4.1 were also used here to enable direct comparison of the relative performance of the tests.

The exposure model was generated with the intention of being able to produce relative risk contours that are similar in size and magnitude to those genuinely seen around Nant y Gwyddon. The contours for the relative risk of congenital malformations around Nant y Gwyddon are shown in Figure 3.11. When $\beta = 0.025$ the relative risk contours, seen in Figure 3.15a, are similar in size and magnitude to those seen in Figure 3.11 albeit with a

slightly different spatial pattern. In fact in the simulation under the alternative hypothesis the relative risk surface does not exceed 2 in regions close to the exposure sources whereas the observed relative risk surface does exceed 2 within 3km of Nant y Gwyddon. Therefore, these simulations that are used to assess the efficiency of the testing procedure could suggest that the testing procedure is less efficient than it would be when dealing with observed data from Nant y Gwyddon. The relative risk contours for the simulation are not as extensive around site 2 as they are around site 1 as the population is smaller in that region, where there is no population there will be no relative risk contours plotted.

Figure 3.15: the relative risk contours for a simulation of congenital malformations under a) the alternative hypothesis with beta = 0.025 and b) the null hypothesis with beta = 0

a) **b)**

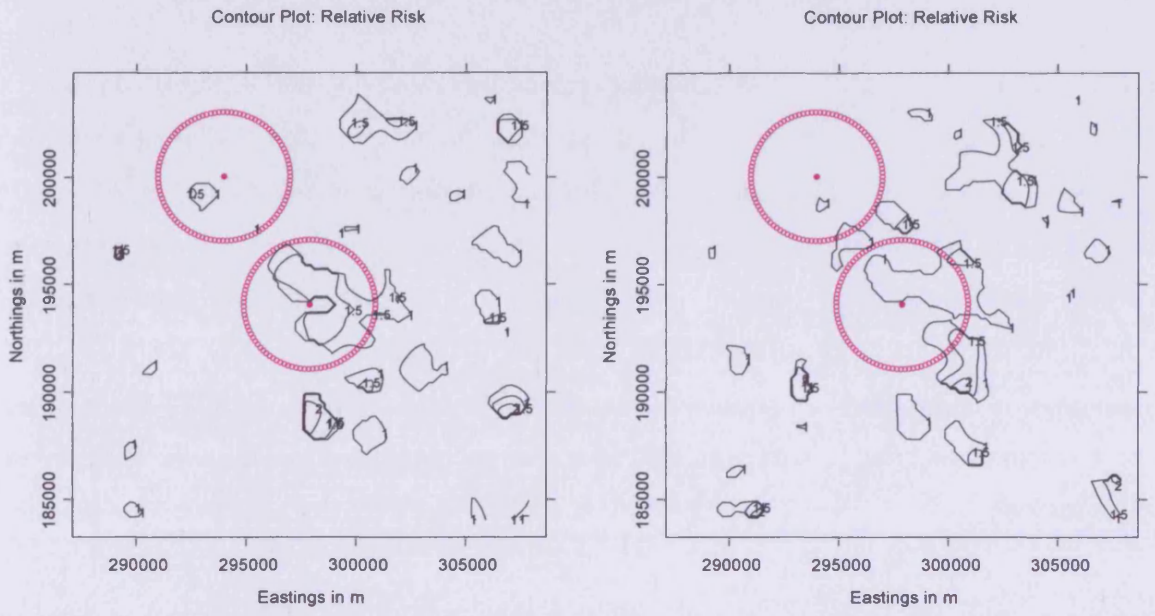


Figure 3.15b displays the relative risk contours under the null hypothesis with $\beta = 0$. The increase in relative risk seen in Figure 3.15a around the two sites is not present in Figure 3.15b where the relative risk contours are not as elevated. However, the existence of some contours indicates that relative risks can be over 1 suggesting the presence of elevated risks in some regions. It can be deduced that these increases in risk occur as a result of random fluctuations in the risk surface as it was assumed in the simulations that there was no excess risk.

3.3.2 Properties of the Relative Risk Contours

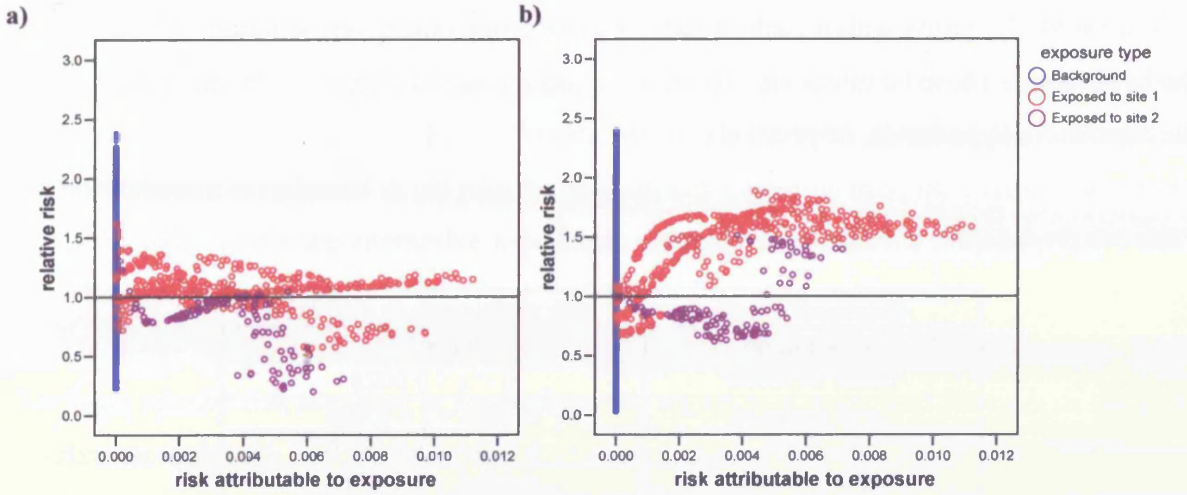
In the development of a testing procedure, for investigating the impact of exposure to a risk factor on the risk of a given health outcome, the properties of the relative risk contours under the null and alternative hypotheses needed to be considered to determine which test statistics should be used.

Modification of the program for creating relative risk contours can be performed as it was for the global significance test to create a series of relative risks for each of the births in the population from the cases simulated under the relevant hypothesis. The smoothing parameter used for the relative risk calculations was 0.5km as before. The impact of changing the smoothing parameter is discussed later.

In the creation of the relative risk contours regions of low population density, which led to unstable relative risk values, were filtered out. For the same reason, the relative risks of mothers in regions where the height of the kernel density for the location of all births is lower than 20% of its maximum height are not included in the analysis. However it should be noted that the whole population is used in the simulation of cases and in the relative risk calculations. It is after the relative risks have been calculated that they are filtered out. Although, this is possible in a population the size of Nant y Gwyddon, 41,337 births, it may have a detrimental effect on the analysis of a smaller population, which will be discussed further later on.

It is clear from the exposure model that the risk attributable to exposure is given by subtracting the baseline risk from the overall risk of a baby being born with a congenital malformation, i.e. the risk attributable to exposure for birth i is $p_i - \alpha_i$. The risk attributable to exposure under the alternative hypothesis for each birth can be plotted against the corresponding relative risk value in a scatterplot. Figure 3.16a and Figure 3.16b are examples of these scatterplots under the null and alternative hypotheses, respectively. It should be noted that the transformation and standardisation of the relative risk values in the scatterplots were considered but ultimately they did not influence the performance of the non-parametric testing procedure and so the relevant detail on these processes is omitted from this thesis.

Figure 3.16: the scatterplot between the risk attributable to exposure under the alternative hypothesis and the relative risks for one simulation under a) the null hypothesis and b) the alternative hypothesis



The nature of the kernel density means that two individuals who are located very close to each other will have very similar relative risk values as a result of the smoothing process. Therefore, the relative risks of two individuals located close together are not independent. This is reflected in the striations running through the scatterplots, which represent traversing along the radii of the exposed regions, following the rise and fall in the relative risk surface.

In addition, in the scatterplots the variation of the plotted points falls with increasing exposure risk. Imagine circles centred over the site all with different radii, clearly, a circle with a small radius has a smaller circumference than a circle with a larger radius. These smaller circumferences, therefore, relate to mothers living closer to the centre of the site and, thus, those with a higher risk attributable to exposure. The decrease in circumference limits the population size in these higher risk areas, thus, reducing the likelihood of seeing estimates for mothers whose residence is close to the centre of the site. Therefore, the number of points in the scatterplot drops with increasing risk resulting in a lower variability.

In Figure 3.16a and Figure 3.16b the high density of points that represent births in the unexposed region, where the risk attributable to exposure is zero, makes it difficult to assess the distribution of the relative risk values in this unexposed region. Boxplots of the relative risks were generated for a simulation under the null hypothesis and for a simulation under the alternative to offer a clearer picture of the distribution of relative risk in the background region. Exposure groups were defined and used to partition the population for the purpose of

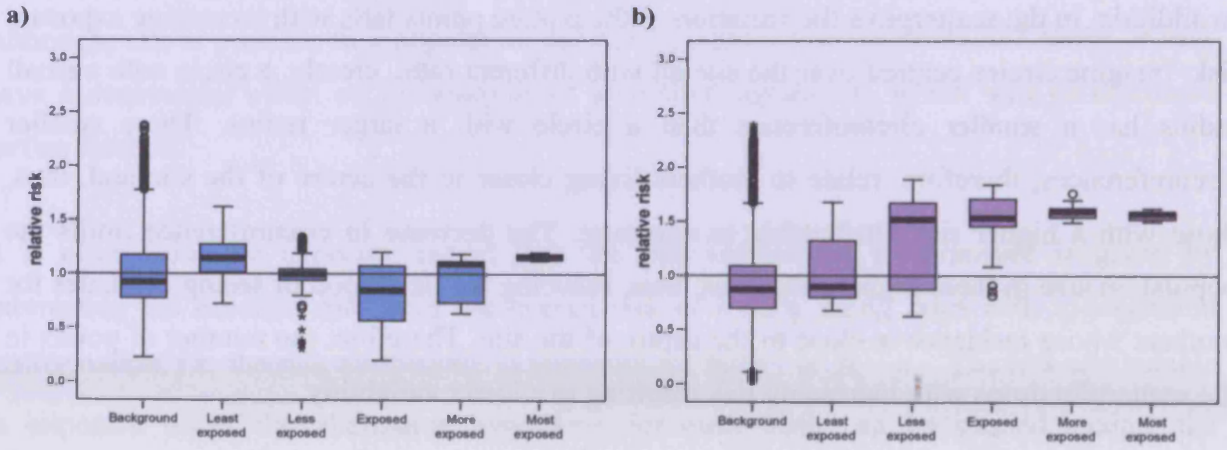
creating the boxplots. The risk attributable to exposure was used to determine the exposure groups which are defined in Table 3.2. Boxplots were then generated for the relative risk values for all the births within each of the exposure groups using one simulation from each of the hypotheses. These boxplots are shown in Figure 3.17a and Figure 3.17b under the null and the alternative hypotheses, respectively.

Table 3.2: the definition of the exposure groups

Exposure group	Risk attributable to exposure
Background	$x = 0$
Least exposed	$0 < x \leq 0.0025$
Less exposed	$0.0025 < x \leq 0.005$
Exposed	$0.005 < x \leq 0.0075$
More exposed	$0.0075 < x \leq 0.01$
Most exposed	$x > 0.01$

The median relative risk under the null hypothesis should be approximately equal in all exposure groups. The deviation of the medians seen in Figure 3.17a from equality reflects the strong random component in the simulation process.

Figure 3.17: the boxplots for the relative risk for one simulation under a) the null hypothesis and b) the alternative hypothesis by exposure group

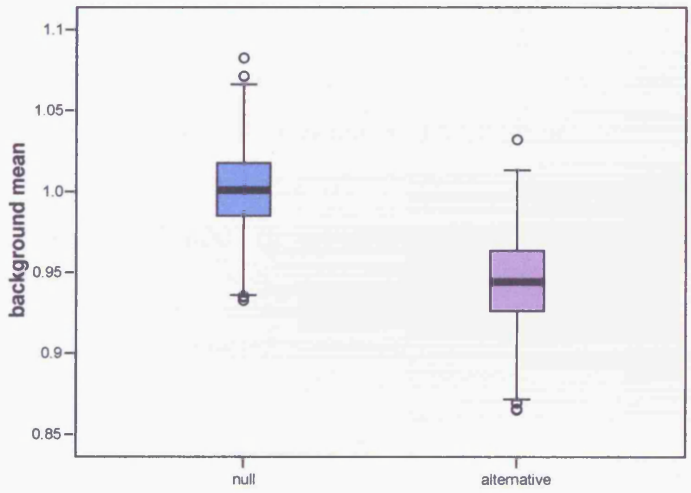


The median relative risk value in the unexposed background region is lower under the alternative hypothesis in comparison to the null. This is to be expected because of the nature of the kernel density estimators used to determine the relative risk values represented in the plots. The kernel density is constructed by averaging across the heights of two dimensional Gaussian surfaces centred over each observation in the population. By definition, each

Gaussian surface has a total volume of 1 beneath it, therefore, it follows that the volume beneath the kernel density surface is also 1. This criterion means that when the height of the kernel density surface increases at the locations of the simulated cases, the height must fall in the remaining regions to maintain a volume of 1. The kernel density for the expected cases, used in the denominator for calculating the relative risk, is fixed for all simulations. Therefore, this phenomenon is reflected in the relative risk surface. Consequently, the excess cases in the exposed regions under the alternative hypothesis are responsible for the drop of relative risks in the background regions.

The mean relative risk for those in the background region was calculated for each of the 1000 simulations under the null hypothesis and this was repeated for each of the 1000 simulations under the alternative. Figure 3.18 presents the boxplot for the mean relative risk in the unexposed background region for the 1000 simulations under the null hypothesis against that of the 1000 simulations under the alternative hypothesis. It is clear from this plot that the mean relative risk in the unexposed region is generally lower under the alternative hypothesis than under the null hypothesis.

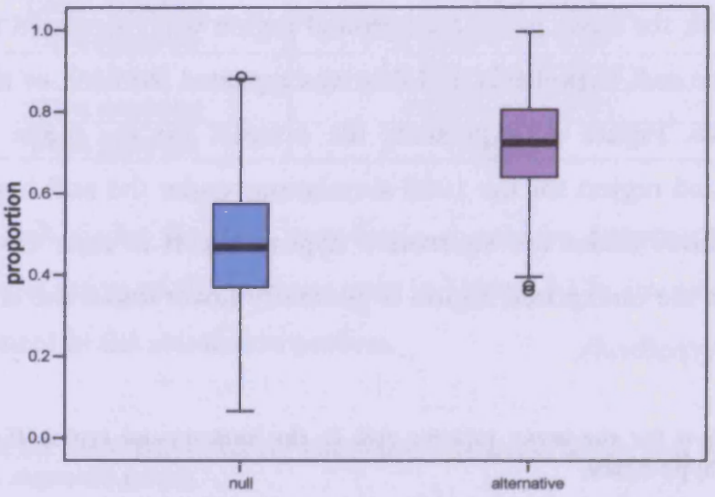
Figure 3.18: the boxplots for the mean relative risk in the background region for the 1000 simulations under each of the two hypotheses



In Figure 3.16a the relative risk values vary around the value of 1, which corresponds to no increased risk, with approximately equal numbers of relative risk values above and below that reference point. However, in contrast the relative risk values in Figure 3.16b are elevated in the exposed region so the majority of the points in the exposed region represent relative risks

above 1. The proportions of relative risks over 1 in the exposed regions were calculated for each of the 1000 simulations under the null hypothesis and for each of the 1000 simulations under the alternative hypothesis. These proportions are displayed in boxplots in Figure 3.19 for the null and alternative hypotheses separately. The proportion of relative risks over 1 is typically larger under the alternative hypothesis where mothers are exposed to an increased risk resulting in these higher relative risk values.

Figure 3.19: the boxplots for the proportion of relative risks over 1 in the exposed region for the 1000 simulations under each of the two hypotheses



The points representing mothers in the exposed regions form no obvious trend in Figure 3.16a. However, in contrast, the relative risk values in the exposed regions of Figure 3.16b are elevated, manifesting in a positive trend in the scatterplot. The slope of the points in the exposed region of the scatterplot was calculated for all 2000 simulations.

The ordinary least squares estimator was used to calculate the slopes. Suppose Y_i represents the relative risk of mother i who resides within 3km of the centre of either site and x_i represents their risk attributable to exposure then the fitted line is $Y_i = mx_i + c$ where m and c are the slope (gradient) and intercept, respectively. The ordinary least squares estimate for the value of m is defined as

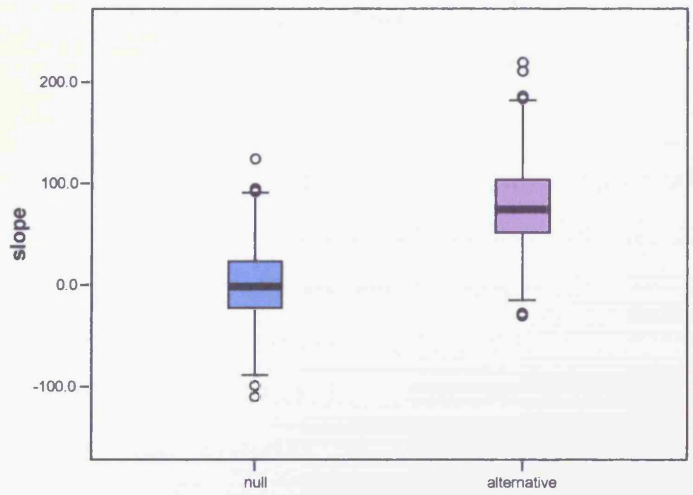
$$\hat{m} = \frac{\sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y}}{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2}$$

where $\bar{x}$ is the mean risk attributable to exposure for all mothers in the exposed regions and $\bar{y}$ is the mean relative risk value for all exposed mothers [273].

A program was written for the purpose of calculating the least squares estimator for the 1000 simulations of cases under each of the hypotheses. This calculation assumes equal variance for all points which may not be true and in that case a weighted least squares estimator should be used. This will be discussed in detail later on.

The resultant gradients were plotted in the boxplots displayed in Figure 3.20 contrasting the values under the two hypotheses. A comparison of the two boxplots indicates that as with the proportion of relative risks above 1 in the exposed region the gradient of the points for the same area in the scatterplot is generally higher under the alternative hypothesis.

Figure 3.20: the boxplots for the slope in the exposed region of the scatterplot for the 1000 simulations under each of the two hypotheses



It should be noted that some of the slopes calculated under the alternative hypothesis are negative despite the fact that an exposure effect with a positive trend was assumed in the simulations. This is reflective of the random assignment of cases in the simulation process and the relative rarity of a case, both of which can lead to lower relative risks close to the centre of the site, a problem that is exacerbated by the smoothing process.

The analysis described above indicates that several characteristics of the scatterplots between risk attributable to exposure and relative risk differ under the two hypotheses. These characteristics are as follows:

1. background mean – the mean relative risk value in the unexposed region;
2. proportion – the proportion of relative risk values above 1 in the exposed region; and
3. slope – the gradient of the straight line fitted through the points corresponding to the exposed population in the scatterplot between risk attributable to exposure and relative risk.

The background mean is lower when the alternative hypothesis is true because the elevations in the relative risk surface in regions of high exposure are compensated by a trough in the relative risk surface in regions of low exposure. These elevations in the relative risk surface in the exposed regions are not present under the null hypothesis and therefore the background mean is higher under the null than under the alternative. The elevations in the relative risk surface in regions of higher risk attributable to exposure under the alternative hypothesis also result in both a higher slope and an increased proportion of relative risks above 1 under the alternative compared to under the null.

The differences in these three characteristics in the scatterplots under the null and alternative hypothesis mean that each of them can be used as test statistics in a hypothesis test for the spatial association between the health outcome and the risk factor. Clearly the proportion and the slope are dependent and therefore it is questionable as to whether the proportion of relative risks above 1 in the exposed region is required at all. However, in Figure 3.16b there is an obvious non-linear relationship between the exposure risk and the relative risk and consequently proportion provides a useful extra discriminant that compensates to some extent for fitting a straight line to a clear non-linear trace to calculate the slope.

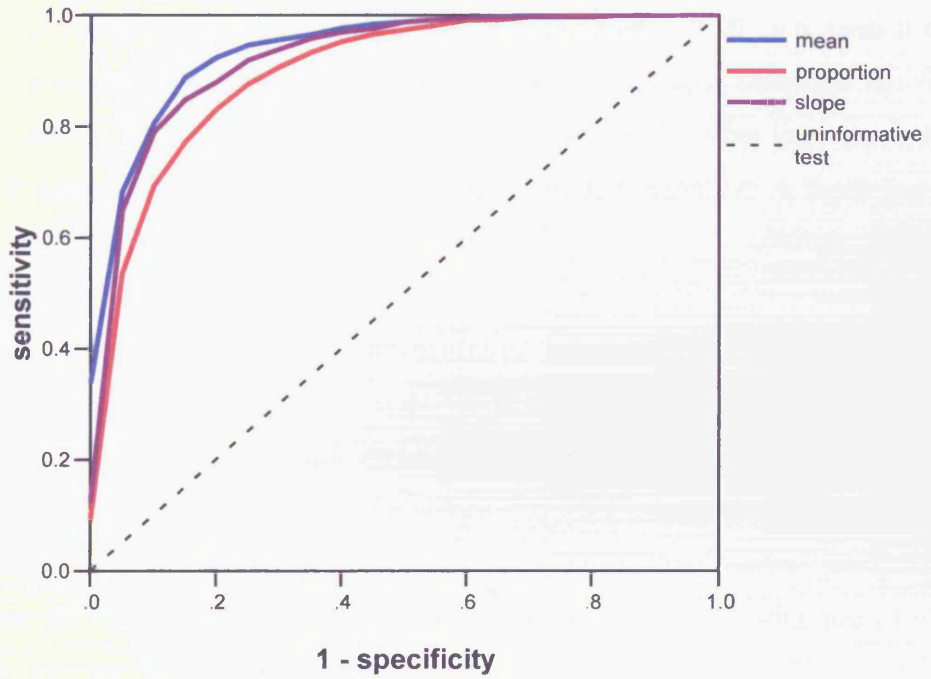
3.3.3 Kernel Density Test for Spatial Association

The three test statistics determined above from the relative risk contours can be used in a hypothesis test to determine whether or not there is spatial association between exposure to a risk factor and the health outcome. The test statistics can be used on their own to form the basis of the hypothesis but to increase efficiency they can be combined in one test using discriminant analysis.

Testing with one test statistic

Performing the testing procedure with only one test statistic is straight forward. 1000 simulations are performed under the null hypothesis that there is no additional risk attributable to exposure, which when using observed health data will require predicted probabilities from a logistic regression model fitted with known confounders. The chosen test statistic is calculated for the observed health data and for each of the 1000 simulations. The p-value is then the proportion of simulations that had a test statistic value at least as extreme as that for the observed health data. Therefore, when considering the slope or the proportion it is the proportion of simulations with higher values than the observed value of the test statistic. However, as explained earlier when dealing with the background mean, it is the proportion of simulations with lower values than the observed situation.

Figure 3.21: the ROC curve for the kernel density test for spatial association with one test statistic using the simulations from Nant y Gwyddon



The ROC curve summarising the efficiency of the three test statistics using the sensitivity and specificity of the test is shown in Figure 3.21. These ROC curves indicate that the tests are reasonably good for this effect size. Background mean is the most efficient test statistic

because it is based on the relative risks of 32,484 mothers whereas the other two test statistics are calculated from the exposed population of only 5,716. The hypothesis test using proportion is fairly crude and this is reflected in its poorer performance in comparison to that using slope. It should be noted that the sensitivities are estimated using 1000 simulations under each hypothesis and therefore are subject to sampling error which means that the differences in efficiency between test statistics is not necessarily statistically significant at all points in the ROC curve, this is discussed in more detail later on. The discussion which follows indicates how these tests can be bettered by using discriminant analysis to combine more than one statistic for use in one test.

Discriminant analysis

Discriminant analysis centres on the assignment of an observation $\mathbf{x}$ of unknown origin to one or more distinct groups using the value of the observation [274]. In this case the observation is the vector of values of the required statistics for the actual cases of the condition. There are two possible groups this observation can belong to: those where an exposure effect exists and those where it does not. If the observation $\mathbf{x}$ is assigned to the former group then the null hypothesis is not accepted providing evidence that spatial association between exposure and the health outcome may exist. However, if it is assigned to the latter then the null hypothesis is accepted and there is evidence that exposure to the specified factor does not increase the risk of the health outcome.

Two possible methods of carrying out discriminant analysis are investigated here: the non-parametric method and the parametric method. Here the non-parametric method involves estimating the distribution of the observation $\mathbf{x}$ under the null and alternative hypotheses using kernel density estimators. In contrast the parametric method assumes that the joint probability function is bivariate normal with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ which are estimated and then Fisher's Linear Discriminant Function is used for the decision process as described by Lachenbruch [274]. However, this relies on the covariance matrices of the two populations being equal and if this is not true then the boundary is not linear and other approaches are needed.

In keeping with the kernel density methodology, non-parametric discriminant boundaries based on kernel density estimators are presented here. Parametric discriminant analysis was

also considered and was discussed in Appendix 3 alongside the testing procedures required to assess the validity of the assumptions: the test for bivariate normality [275] and Box's M test [276]. The efficiencies of the two distinct approaches are similar as seen in Appendix 3. Both in this chapter and in Appendix 3, hypothesis testing using the background mean and proportion is considered for illustrative purposes only. Consequently, **x** refers to the column vector of length two comprising the values of these two test statistics.

The scatterplot of the background mean against the proportion for each of the 1000 simulations under the two hypotheses is shown in Figure 3.22a. The blue points represent the simulated values of **x** (background mean, proportion) under the null hypothesis and the pink those under the alternative.

Figure 3.22: a) the scatterplot between the background mean and the proportion under the null and the alternative hypotheses and b) the probability contour plots for the joint distributions of the background mean and the proportion along with the discriminant boundary

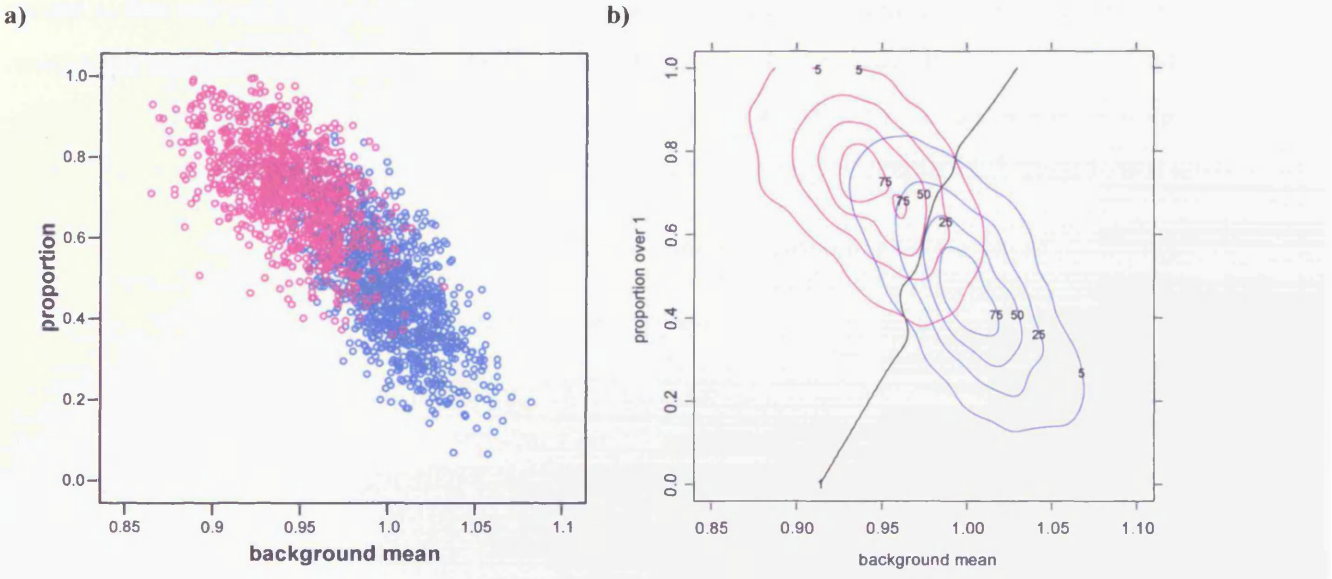


Figure 3.22b displays the probability contour plots for the corresponding kernel densities based on the points plotted in the scatterplots in Figure 3.22a, these kernel densities under the two hypotheses are used to determine the discriminant boundary. The blue lines represent the contours under the null hypothesis and the pink those under the alternative. The discriminant boundary in this approach is determined by the value of the ratio of the heights of the two kernel densities. In the diagram it is represented by the black line and is drawn where the ratio is equal to 1, in other words, where the two kernel densities have equal height, for illustrative

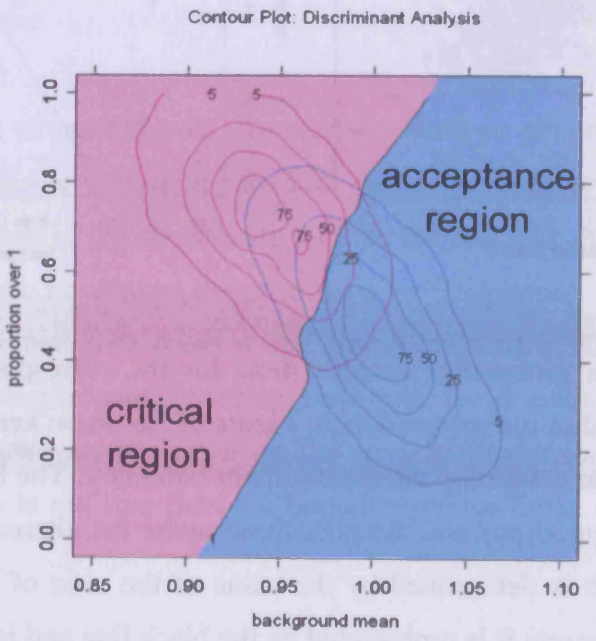
purposes. In Figure 3.22b it is apparent that the discriminant boundary fluctuates slightly. These small deviations are more obvious with a smaller effect size than they are in this situation. This has no untoward affects on the discriminant analysis as discussed in Appendix 3. The smoothing parameters used in creating these boundaries were calculated using the normal optimal smoothing method. Although, it has been previously mentioned that the use of normal optimal smoothing is not always valid, it is a quick way of selecting a rough value for the smoothing parameter. The choice of smoothing parameter has been investigated further and the results are discussed later.

The ratio of the heights of the two kernel densities is referred to as the 'height ratio' and is defined as

$$\frac{\text{the height of the kernel density for } \mathbf{x} \text{ under the alternative hypothesis}}{\text{the height of the kernel density for } \mathbf{x} \text{ under the null hypothesis}}.$$

The acceptance region is the area where the height ratio is below the value represented by the discriminant boundary as in this region the height of the kernel density under the null is higher than it is under the alternative hence the null is more likely to be true. Conversely, the critical region is where the height ratio is above 1 where the alternative hypothesis is more likely to be true. Figure 3.23 illustrates the definition of these two regions graphically.

Figure 3.23: the acceptance and critical regions



The efficiency of this non-parametric approach to discriminant analysis can be assessed using a ROC curve. The specificity is estimated by the percentage of simulations under the null hypothesis whose **x** values are located in the acceptance region. The corresponding sensitivity is estimated by the percentage of simulations under the alternative hypothesis whose **x** values are located in the critical region.

The discriminant boundary can be moved to change the specificity so that the corresponding sensitivity can be determined. Supposing the height ratios for each point determined by the values of **x** for the 1000 simulations under the null hypothesis are listed in ascending numerical order. If the specificity is $(100-\alpha)\%$ then the discriminant boundary is moved to where the height ratio is equal to the $\frac{(100-\alpha)}{100}N^{th}$ height ratio in this ordered list. This value is referred to as the critical height ratio. Therefore, the sensitivity is the proportion of **x** values under the alternative hypothesis located at points where the height ratio exceeds the critical height ratio. The ROC curve can be generated by altering the specificity and calculating the corresponding sensitivity for each selected value.

Testing with two test statistics

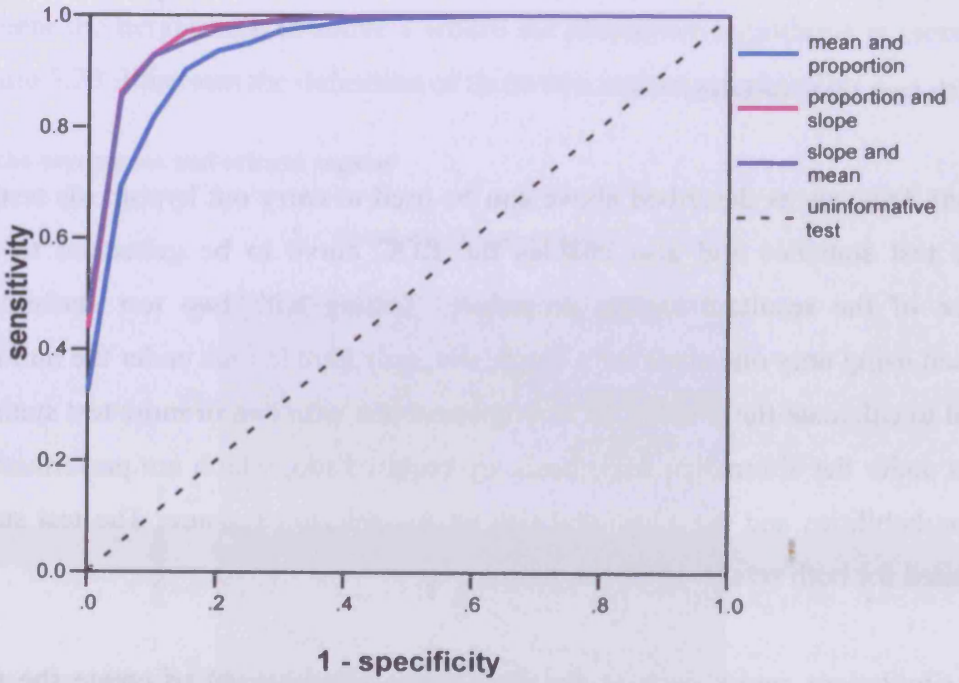
Discriminant Analysis as described above can be used to carry out hypothesis testing with a pair of the test statistics and also enables the ROC curve to be generated to assess the performance of the resultant testing procedure. Testing with two test statistics is more involved than using only one since for a single test, only simulations under the null hypothesis are required to calculate the p-value. In a component test with two or more test statistics, 1000 simulations under the alternative hypothesis are required too, which are performed using the predicted probabilities and the additional risk attributable to exposure. The test statistics are then calculated for both sets of 1000 simulations.

The 1000 simulations under each of the hypotheses can be used to create the probability contours for the joint distributions of the test statistics and the discriminant boundary with a 5% significance level can be generated. Plotting the test statistics for the observed health data on the resultant diagram indicates whether it lies in the acceptance or critical regions. Clearly, if the point lies in the critical region there is evidence that there is a causal link and if it lies in

the acceptance region there is no evidence of a relationship. The p-value is given by the proportion of height ratios for the 1000 simulations under the null hypothesis that are as extreme as the height ratio for the observed health data.

The ROC curve summarising the efficiency of hypothesis testing with a pair of test statistics is shown in Figure 3.24. These are again good hypothesis tests for comparing environmental exposure and the risk of the health outcome. The tests using background mean and proportion; and proportion and slope have similar efficiency. However, the performance of the hypothesis test using slope combined with background mean is not as good. This reduction in efficiency is a result of a lower degree of separation between the joint distributions of background mean and slope under the two hypotheses relative to the other combinations of test statistics. The use of two test statistics as opposed to one increases the efficiency of the testing procedure because more information is used when two test statistics are involved.

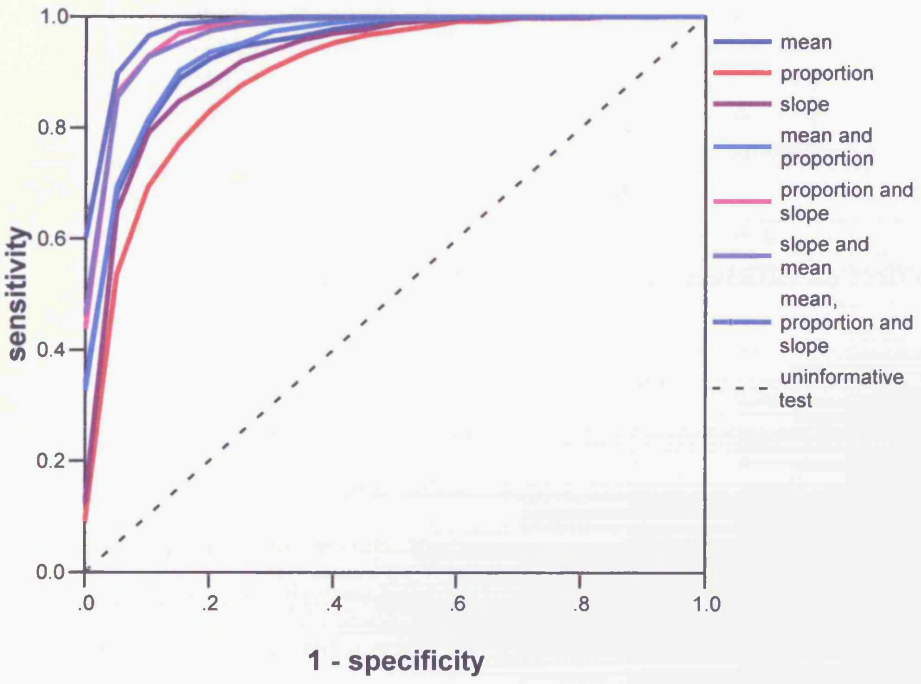
Figure 3.24: the ROC curve for the kernel density tests for spatial association using pairs of test statistics based on the simulations from Nant y Gwyddon



Testing with three test statistics

The improvement in test performance caused by using two test statistics instead of one can be surpassed by using all three test statistics. Discriminant analysis is used when testing with three test statistics in much the same way as it was with two test statistics. However, the joint distributions of the test statistics under the two hypotheses cannot be illustrated easily as they are three-dimensional images and the discriminant boundary is no longer a line but a surface. Therefore, it is recommended that the test is performed by calculating the p-value and if the p-value is less than 0.05 then there is some evidence of spatial association between exposure to the risk factor and the health outcome.

Figure 3.25: the ROC curve for the kernel density test for spatial association with all possible combinations of test statistics based on the simulations from Nant y Gwyddon



In Figure 3.25 the efficiency of the hypothesis tests with one, two and three statistics are compared. This clearly indicates that by increasing the number of test statistics the efficiency of hypothesis testing is improved. The sensitivity values presented in the ROC curve were estimated using 1000 simulations under each of the hypotheses. The use of a finite number of simulations means that there is sampling error in these estimates which is quantified for the 5% significance level using 95% confidence intervals for the sensitivity given in Table 3.3.

The location of the critical boundary used to set the significance level to 5% is also determined by the 1000 simulations under each of the hypotheses and consequently there is additional sampling error in the significance level quoted. However, the confidence intervals quoted in Table 3.3 should still give a good idea about the relative performance of the tests. The consideration of the confidence intervals indicate that although there is not much difference between the sensitivity values for each test statistic combination it is clear that when implementing this non-parametric approach, all three test statistics should be combined to maximise the power of the test.

Table 3.3: the sensitivity with 95% confidence intervals for the kernel density test for spatial association at the 5% significance level based on 1000 simulations from Nant y Gwyddon

Test statistics	Sensitivity (95% CI)
Mean	0.683 (0.654, 0.711)
Proportion	0.536 (0.505, 0.567)
Slope	0.649 (0.619, 0.678)
Mean and proportion	0.694 (0.665, 0.722)
Proportion and slope	0.863 (0.840, 0.883)
Slope and mean	0.855 (0.832, 0.876)
Mean, proportion and slope	0.898 (0.878, 0.915)

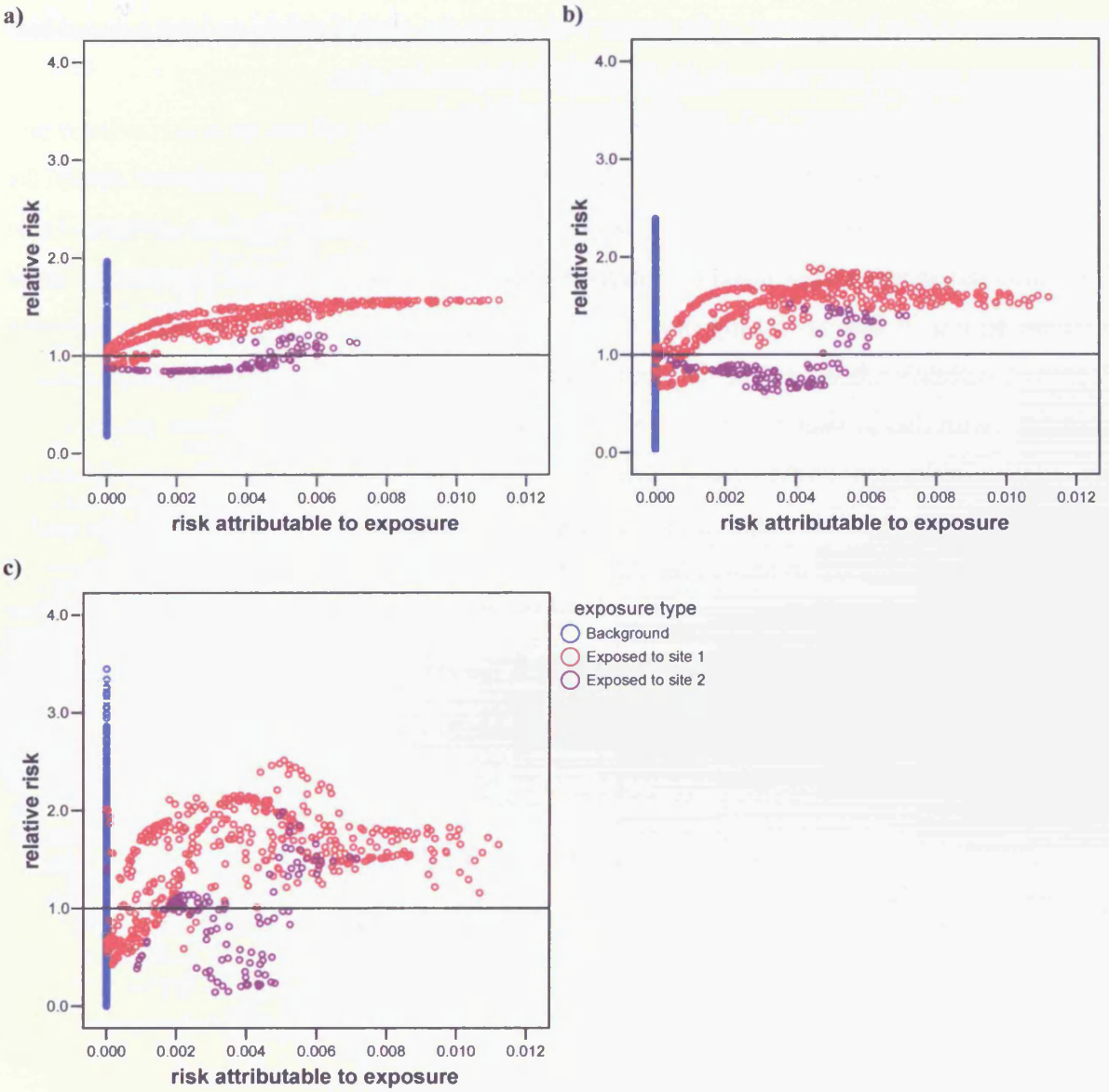
3.3.4 The Effect on Efficiency of Changes to the Testing Procedure

The testing process described above is clearly a powerful way of determining whether or not there is a statistically significant spatial association between the exposure risk and the health outcome. However, there are a number of steps that could have been implemented differently in the testing procedure. For example, different smoothing parameters could have been selected in the generation of the kernel densities used to derive the relative risk surface. In the testing procedure described above the relative risks were filtered to exclude relative risks calculated in regions of low birth density and the test could be applied without this filter. Weighted least squares can be used instead of ordinary least squares in the calculation of the slope. Furthermore, different values for the smoothing parameters used in the non-parametric discriminant analysis can be considered. Clearly, it is important to know whether modifying the approach slightly affects the performance of the test. Each of the changes that can be implemented to the testing procedure is discussed and the resultant alteration in the performance of the test is commented on below.

Smoothing parameters in the relative risk surface

The smoothing parameter controls the degree of smoothing in the kernel density estimation, as discussed previously, if it is increased then the estimator becomes more biased and if it is reduced then the precision of the estimator falls. Previously, the smoothing parameter selected was 0.5km, here, smoothing parameters of 0.25km and 1km are also considered to assess the impact of the smoothing parameter on the testing procedure.

Figure 3.26: the scatterplot between the risk attributable to exposure and the relative risks simulated under the alternative with a smoothing parameter of a) 1km, b) 0.5km and c) 0.25km



The scatterplot between the risk attributable to exposure and the relative risks simulated under the null hypothesis is presented in Figure 3.26 using a decreasing smoothing parameter from a) to c). The impact of the smoothing parameter on a kernel density estimator is discussed in Section 3.1. The fall in imprecision and increase in bias associated with using a larger smoothing parameter is reflected in the reduced variability of the points and the less prominent peak in the exposed region observed in Figure 3.26a where the smoothing parameter used was 1km.

Similarly, the greater imprecision and reduced bias associated with a smaller smoothing parameter is reflected in Figure 3.26c where the smoothing parameter used was 0.25km. The greater imprecision is apparent in the larger variance in the plotted points and the reduced bias in the higher relative risk values in the peak for the exposed region.

The ROC curves for testing using all three smoothing parameters mentioned above are compared in Section A1.2 and the sensitivities at the 5% significance level are given in Table 3.4. Although there is little change in performance, the testing procedure is generally more efficient with a higher smoothing parameter as the additional smoothing results in a higher degree of separation between the distributions for the test statistics under the two hypotheses. However, with this greater separation the bias is increased so that the resultant estimates for the relative risks are further from their true values. After the hypothesis test has been performed the next step is to quantify the effect size, therefore, lower bias is desirable and is achieved with a lower smoothing parameter.

Table 3.4: the sensitivity and 95% confidence intervals at the 5% significance level for the kernel density test with smoothing parameters of 1km, 0.5km and 0.25km

	Sensitivity (95% CI)		
Test statistics	1km	0.5km	0.25km
Mean	0.745 (0.717, 0.771)	0.683 (0.654, 0.711)	0.594 (0.563, 0.624)
Proportion	0.531 (0.500, 0.562)	0.536 (0.505, 0.567)	0.588 (0.557, 0.618)
Slope	0.729 (0.700, 0.756)	0.649 (0.619, 0.678)	0.628 (0.598, 0.657)
Mean and proportion	0.739 (0.711, 0.765)	0.694 (0.665, 0.722)	0.673 (0.643, 0.701)
Proportion and slope	0.874 (0.852, 0.893)	0.863 (0.840, 0.883)	0.858 (0.835, 0.878)
Slope and mean	0.871 (0.849, 0.890)	0.855 (0.832, 0.915)	0.829 (0.804, 0.851)
Mean, prop and slope	0.907 (0.887, 0.924)	0.898 (0.877, 0.915)	0.914 (0.895, 0.930)

Therefore, the initial criterion of selecting a smoothing parameter based on balancing the bias and imprecision of the underlying kernel density estimator is preferential to choosing a

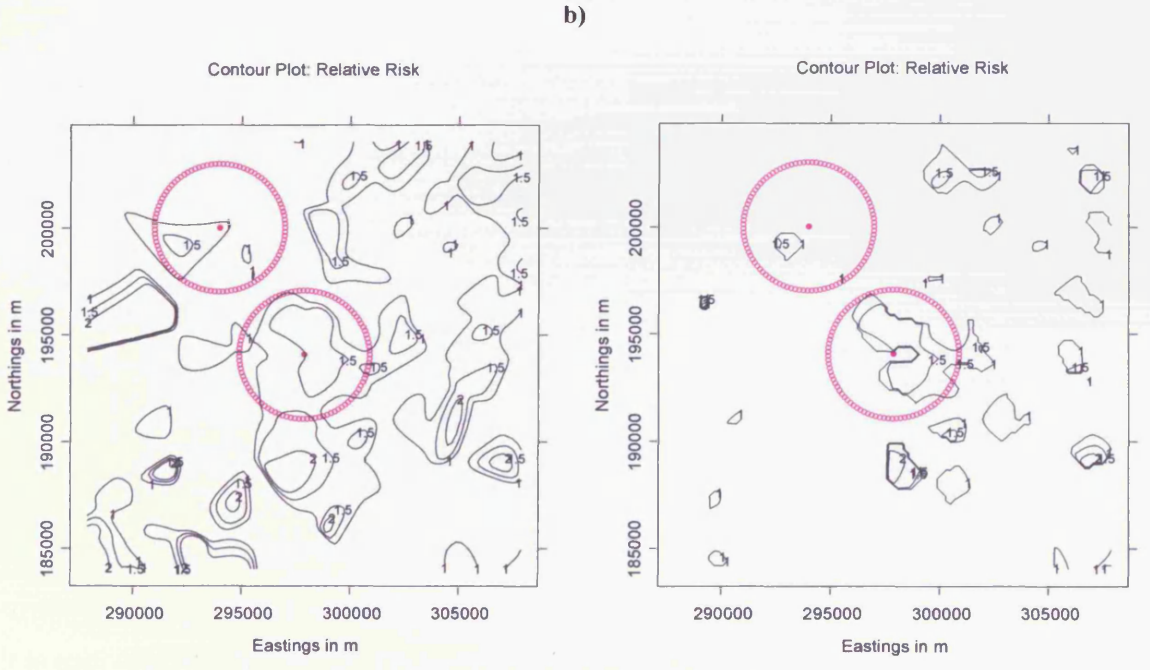
smoothing parameter according to the resultant efficiency of the hypothesis test. Hence, the smoothing parameter of 0.5km is an appropriate compromise in both respects.

Filtering the dataset

In the generation of relative risk contours the relative risk surface was fitted using all births, however, when the contours were plotted relative risk values for regions of low population density were filtered out. The filtering process was implemented so that the inflated relative risks caused by isolated cases in regions of low population do not obscure the true effect in the contour plot. It is for this reason that James et al. concluded that it was necessary to filter out these areas of sparse population density [10].

The relative risk contours for a simulation under the alternative hypothesis are given in Figure 3.27 with no filtering and with filtering. Although, the elevated relative risk values within 3km of the site are still present they are no longer as prominent with many other relative risk hotspots, usually of small size, appearing in regions that coincide with low population density. Therefore, it is sensible to filter out these regions to reduce the noise in the plot before plotting the contours so that the more major regions of high risk are clearly visible.

Figure 3.27: the relative risk contours for a simulation of congenital anomalies under the alternative hypothesis with a) no filtering and b) with filtering



The relative risk values calculated for regions of sparse population density are filtered out in the testing procedure because it removes relative risk values with the greatest sampling variation. However, in terms of hypothesis testing, by omitting these mothers from the testing procedure the population size is reduced, which could in turn have a negative affect on the performance of the resultant test. In the case of Nant y Gwyddon the population size falls from 41,337 to 38,200 by applying the filter, a landfill site with a different population pattern could see a greater or lower percentage drop. Hence, in order to assess the impact of filtering out regions of low population density on the testing procedure the ROC curves were generated without any omissions from the dataset. These ROC curves are compared with those for the original filtered data initially shown in Figure 3.25 in Section A1.3.1. The sensitivities for each of the test statistic combinations at the 5% level are compared in Table 3.5 for filtering and no filtering.

The ROC curves and sensitivity values do not differ greatly when filtering is not implemented. Testing with the background mean is more efficient when the data is not filtered. This improvement in efficiency is a result of the increased population reducing random error in the calculation of the background mean. The performance of testing with the other two statistics is not affected to the same degree as the population in the exposed region, the population on which their calculation is based, is not hugely changed by the filtering process as the population density is generally higher than it is in the background area.

Table 3.5: the sensitivity and 95% confidence intervals at the 5% significance level for the kernel density test with and without filtering

	Sensitivity (95% CI)	
Test statistics	Filtered	Unfiltered
Mean	0.683 (0.654, 0.711)	0.796 (0.770, 0.820)
Proportion	0.536 (0.505, 0.567)	0.544 (0.513, 0.575)
Slope	0.649 (0.619, 0.678)	0.653 (0.623, 0.682)
Mean and proportion	0.694 (0.665, 0.722)	0.796 (0.770, 0.820)
Proportion and slope	0.863 (0.840, 0.883)	0.858 (0.835, 0.878)
Slope and mean	0.855 (0.832, 0.876)	0.880 (0.858, 0.899)
Mean, prop and slope	0.898 (0.878, 0.915)	0.900 (0.880, 0.917)

In a situation where the population density is lower in the exposed regions, testing with slope and proportion will be affected as was seen for Abernant, discussed in more detail later on, where the ROC curves for this landfill site, seen in Section A1.3.2, and the sensitivities at the

5% significance level given in Table 3.6 indicate that testing with one or two test statistics is generally slightly more efficient when the data has not been filtered. However, for Abernant filtering improves the performance of the hypothesis test which combines all three test statistics. This contradicts the results for Nant y Gwyddon where filtering data reduces the efficiency of the test. However, it should be noted that the changes are minimal so it makes little difference in test performance whether the data is filtered or not.

Table 3.6: the sensitivity and 95% confidence intervals at the 5% significance level for the kernel density test with and without filtering for Abernant

	Sensitivity (95% CI)	
Test statistics	Abernant - Filtered	Abernant - Unfiltered
Mean	0.114 (0.096, 0.135)	0.156 (0.135, 0.180)
Proportion	0.102 (0.085, 0.122)	0.118 (0.100, 0.140)
Slope	0.157 (0.136, 0.181)	0.141 (0.121, 0.164)
Mean and proportion	0.174 (0.152, 0.199)	0.181 (0.158, 0.206)
Proportion and slope	0.197 (0.174, 0.223)	0.241 (0.216, 0.269)
Slope and mean	0.232 (0.207, 0.259)	0.233 (0.208, 0.260)
Mean, prop and slope	0.390 (0.360, 0.421)	0.347 (0.318, 0.377)

Weighted least squares

The calculation of the gradients of the straight line fitted through the points in the exposed region of the scatterplot was performed using ordinary least squares regression which assumes that the random errors have constant variance. Figure 3.28 is the scatterplot between population density and the standard deviation of the 1000 relative risks simulated under the null hypothesis for each mother in the Nant y Gwyddon populations. This scatterplot indicates that the standard deviation of the errors is dependent on the population density and is not therefore constant. Weighted least squares regression does not assume that the variances are constant, hence, the slopes were calculated using weighted least squares to investigate the impact it has on the testing process. The maximum likelihood estimator of the gradient of a line using weighted least squares is given by

$$\hat{m} = \frac{\sum_{i=1}^{n} \frac{1}{\sigma_i^2} \sum_{i=1}^{n} \frac{x_i y_i}{\sigma_i^2} - \sum_{i=1}^{n} \frac{x_i}{\sigma_i^2} \sum_{i=1}^{n} \frac{y_i}{\sigma_i^2}}{\sum_{i=1}^{n} \frac{x_i^2}{\sigma_i^2} \sum_{i=1}^{n} \frac{1}{\sigma_i^2} - \sum_{i=1}^{n} \frac{x_i}{\sigma_i^2} \sum_{i=1}^{n} \frac{x_i}{\sigma_i^2}}.$$

A program was written in SPlus to calculate the gradients using this approach.

Figure 3.28: the scatterplot between population density and the standard deviation of the 1000 relative risks simulated under the null for each mother in the Nant y Gwyddon population

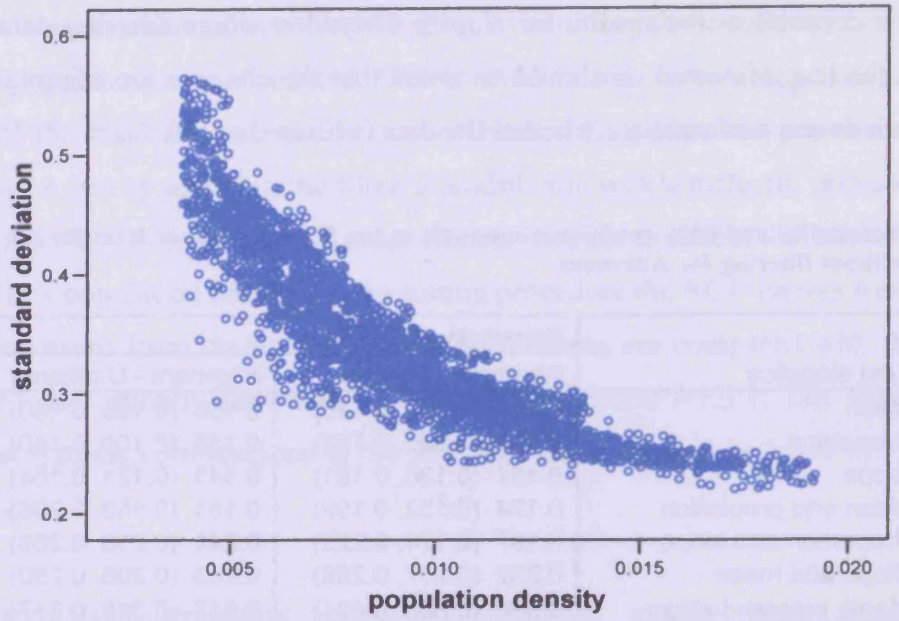
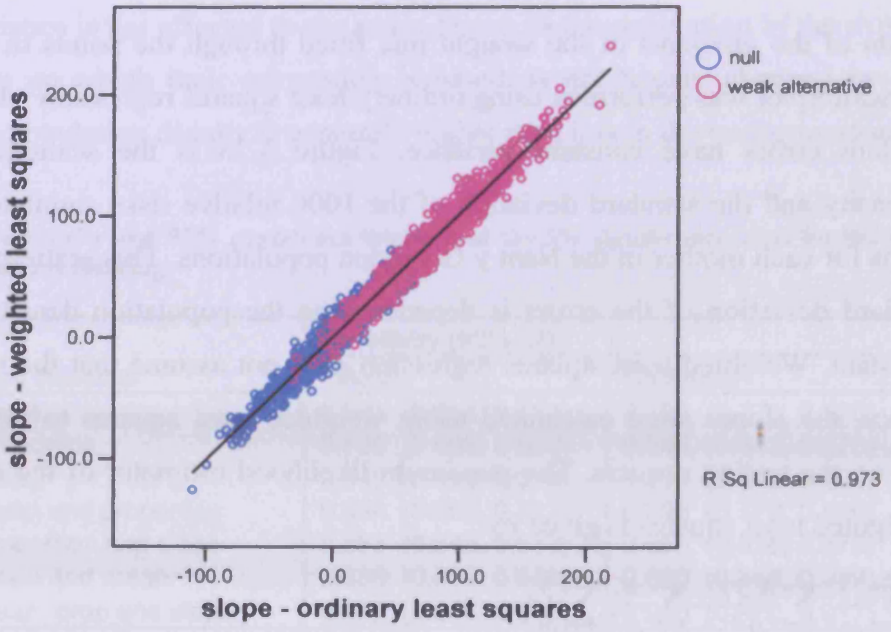


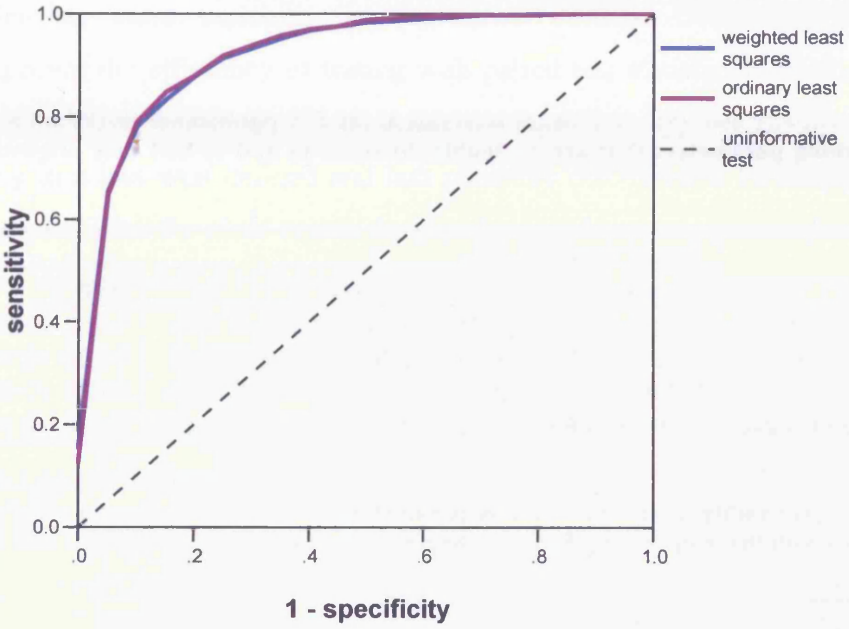
Figure 3.29: the scatterplot of the slopes calculated using ordinary least squares against those calculated using weighted least squares



The slopes calculated under the two methods for each simulation is plotted in Figure 3.29 which indicates that using weighted least squares regression has only a marginal affect on the

values generated for the slopes. Figure 3.30 indicates that the marginal change in values for the slope does not have an impact on the performance of the testing procedure and therefore for the purposes of hypothesis testing it is sufficient to use ordinary least squares.

Figure 3.30: the ROC curve for testing using slope comparing ordinary least squares with weighted least squares



Smoothing parameters in discriminant analysis

The smoothing parameters used for the discriminant analysis discussed in Section 3.3.3 are given in Table 3.7. These were selected by applying normal optimal smoothing and rounding to a sensible figure. The hypothesis tests with two test statistics were reassessed when the smoothing parameters were twice and half the size of those used previously to investigate the impact of the smoothing parameters on the efficiency of the testing process.

Table 3.7: the smoothing parameters used initially for discriminant analysis

Test statistic	Smoothing parameter
background mean	0.007
proportion	0.04
slope	10

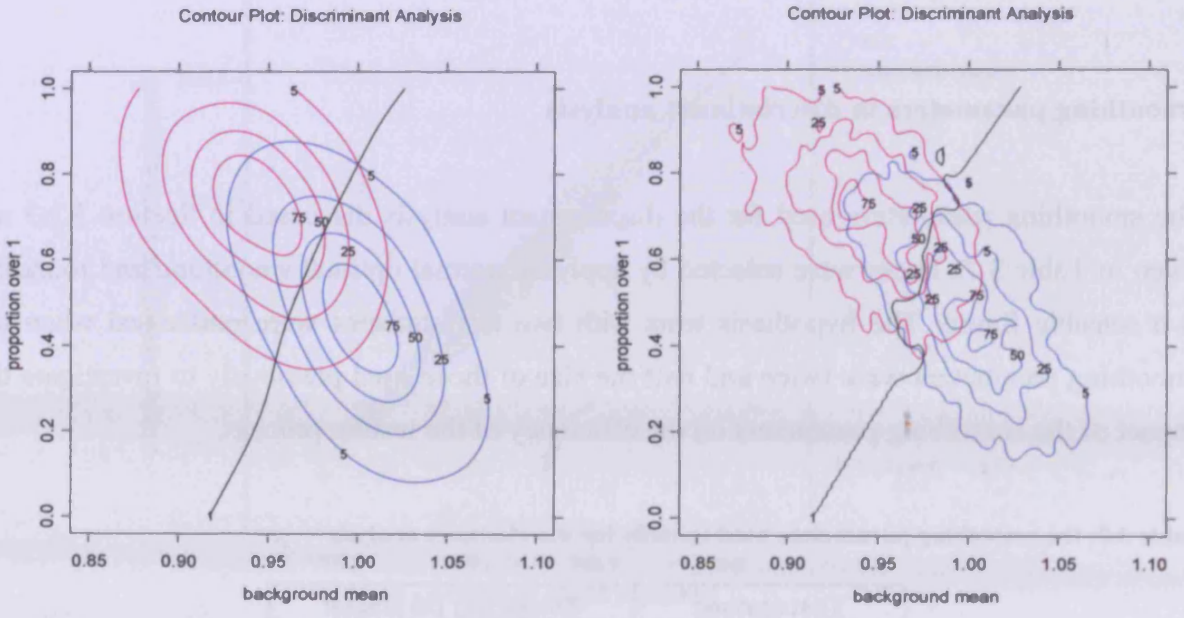
The ROC curves for the different smoothing parameters are compared in Section A1.4 and the sensitivities at the 5% significance level are given in Table 3.8. The ROC curves do not vary much with doubling or halving the smoothing parameters but when the smoothing parameter is reduced the efficiency of the test increases. However, despite the improved performance from a smaller smoother, care must be taken not to take smoothing parameters that are too high or low and the smoothing parameter selected using normal optimal smoothing is recommended.

Table 3.8: the sensitivity and 95% confidence intervals at the 5% significance level for the kernel density test using smoothing parameters that are a) double, b) equal to and c) half that originally used in the discriminant analysis

	Sensitivity (95% CI)		
Test statistics	Double	Original	Half
Mean and proportion	0.686 (0.657, 0.714)	0.694 (0.665, 0.722)	0.732 (0.704, 0.759)
Proportion and slope	0.850 (0.827, 0.871)	0.863 (0.840, 0.883)	0.883 (0.862, 0.902)
Slope and mean	0.838 (0.814, 0.860)	0.855 (0.832, 0.876)	0.881 (0.860, 0.900)
Mean, prop and slope	0.884 (0.863, 0.902)	0.898 (0.878, 0.915)	0.988 (0.979, 0.993)

Figure 3.31: the probability contours and non-parametric discriminant boundary for combining the background mean with the proportion when the smoothing parameters are a) doubled and b) halved

a) **b)**



When the smoothing parameter is increased the kernel density for the joint probability distribution of the test statistics undergoes a greater degree of smoothing with density shifted from the peaks into the tails. The kernel densities are flatter as a result which is reflected in

the contours that spread further across the plane, see Figure 3.31a. If the smoothing parameter is too large then the kernel densities are completely flat with all height ratios equal to one under both hypotheses making it impossible to differentiate between the two distributions and consequently the hypothesis test is uninformative.

When the smoothing parameter is decreased then the densities undergo a smaller degree of smoothing, leaving more variability in the kernel density. Decreasing the smoothing parameter improves the efficiency of testing with paired test statistics according to the ROC curves in Section A1.4. However, the contours are jagged in appearance as seen in Figure 3.31b resulting in a less well defined and less plausible discriminant boundary. The smaller the smoothing parameter the more erratic the discriminant boundary becomes often leading to more than one unique critical region or more than one unique acceptance region which is undesirable.

In conclusion, the most appropriate smoothing parameters for use in the discriminant analysis are the smallest possible smoothing parameters that produce a sensible discriminant boundary that partitions the domain for the test statistics into precisely two regions under this criterion those given by normal optimal smoothing seem sufficient.

3.3.5 Factors that Influence the Efficiency of the Tests

In Section 2.2.4 a number of factors that influence the efficiency of Stone's test and the trend test were highlighted and discussed. The factors considered included adjustment for known confounders, effect size, population size, unknown confounding, misclassified exposure and the spatial distribution of the risk factors. These factors will also affect the performance of the kernel density test for spatial association in a similar way and this phenomenon is discussed in this section. The simulations involved in each situation are the same as those correspondingly used in Chapter 2.

Adjusting for known confounders

In the simulation process used to assess the kernel density test so far it was assumed that all births had an equal baseline risk of having a congenital anomaly. However, in reality there are confounding risk factors that lead to some births having a higher baseline risk of a congenital

anomaly than others. In this situation the relative risk surface used in the testing procedure must be adjusted for known risk factors. In order to assess the impact of adjusting for known confounding on the efficiency of the testing procedure cases are simulated with a variable baseline and the test is performed with adjusted relative risk surfaces.

The ROC curves for testing when an adjustment for known confounding has been made are given in Section A1.5 and the sensitivities at the 5% significance level are given in Table 3.9. Some of the power of the testing procedure is lost when adjustment for known confounding factors is made but the test is still a powerful one regardless. This is probably due to extra variability being induced because of the spatial variation in the confounding variables.

Table 3.9: the sensitivity and 95% confidence intervals at the 5% significance level for the kernel density test with no confounding and with adjustment for known confounding

	Sensitivity (95% CI)	
Test statistics	No Confounding	Confounding with Adjustment
Mean	0.683 (0.654, 0.711)	0.448 (0.417, 0.479)
Proportion	0.536 (0.505, 0.567)	0.432 (0.402, 0.463)
Slope	0.649 (0.619, 0.678)	0.604 (0.573, 0.634)
Mean and proportion	0.694 (0.665, 0.722)	0.552 (0.521, 0.583)
Proportion and slope	0.863 (0.840, 0.883)	0.794 (0.768, 0.818)
Slope and mean	0.855 (0.832, 0.876)	0.722 (0.693, 0.749)
Mean, prop and slope	0.898 (0.878, 0.915)	0.851 (0.828, 0.872)

Effect size

The effect size differs for different health outcomes and risk factors. The relative risks simulated under the null and three different alternative hypotheses were plotted against the corresponding risk attributable to exposure in Figure 3.32. Figure 3.32b is a boxplot of the relative risks by exposure group in a simulation performed with a weaker effect size (β = 0.0125) and does not differ hugely from the boxplot in Figure 3.32a which displays relative risks in a simulation under the null hypothesis. Figure 3.32c represents a simulation under the original effect size (β = 0.025) and displays a greater disparity with the simulation under the null than there was between the null and the weaker effect size. The boxplot for a simulation under a stronger effect size (β = 0.05) is given in Figure 3.32d where the deviation from the pattern seen under the null hypothesis is maximised.

Figure 3.32: the boxplot for the relative risks by exposure group in one simulation under a) the null, b) the weaker alternative, c) the original alternative and d) the stronger alternative hypotheses

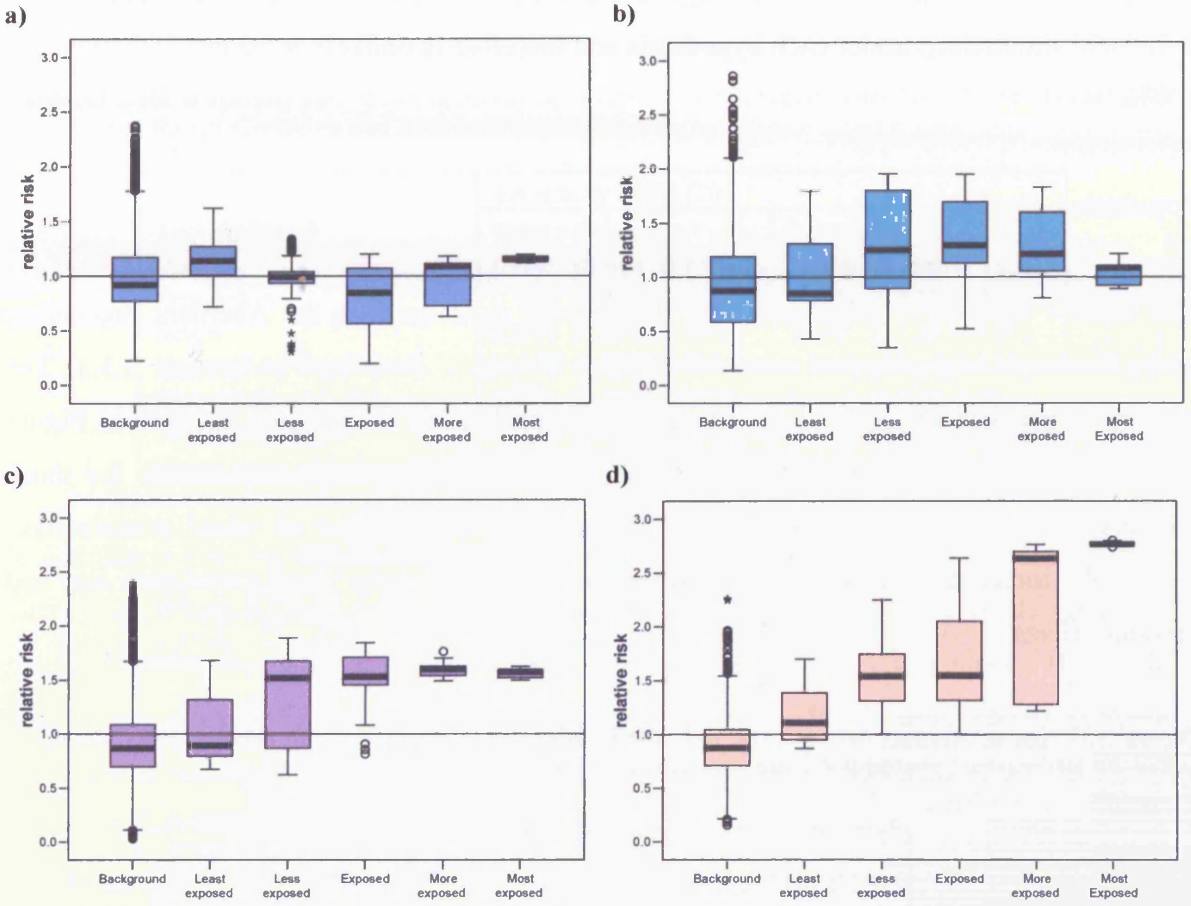


Table 3.10: the sensitivity and 95% confidence intervals at the 5% significance level for the kernel density test with a weak effect size, the original effect size and the strong effect size

	Sensitivity (95% CI)		
Test statistics	Weaker Effect Size	Original Effect Size	Stronger Effect Size
Mean	0.337 (0.308, 0.367)	0.683 (0.654, 0.711)	0.993 (0.986, 0.997)
Proportion	0.256 (0.230, 0.284)	0.536 (0.505, 0.567)	0.830 (0.806, 0.852)
Slope	0.263 (0.237, 0.291)	0.649 (0.619, 0.678)	0.978 (0.967, 0.985)
Mean and proportion	0.354 (0.325, 0.384)	0.694 (0.665, 0.722)	0.993 (0.986, 0.997)
Proportion and slope	0.508 (0.477, 0.539)	0.863 (0.840, 0.883)	1.000 (0.996, 1.000)
Slope and mean	0.470 (0.439, 0.501)	0.855 (0.832, 0.876)	0.999 (0.994, 1.000)
Mean, prop and slope	0.630 (0.600, 0.660)	0.898 (0.878, 0.915)	1.000 (0.996, 1.000)

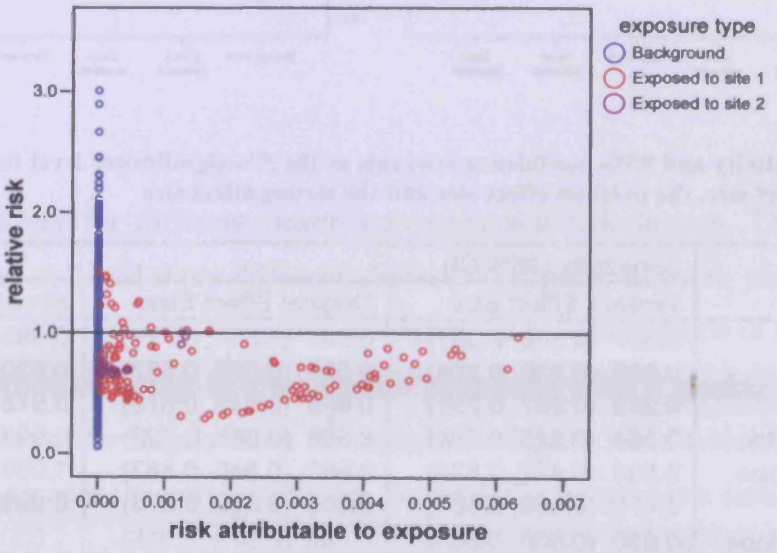
The ROC curves are given in Section A1.6 and the sensitivities at the 5% significance level are given in Table 3.10. The smaller the effect size the less powerful the test becomes. The reason for this is that if the risk attributable to exposure is increased then the relative risk

values under the alternative hypothesis deviate further from those simulated under the null hypothesis. This greater deviation results in a higher degree of separation between the distributions of the test statistics calculated under the two hypotheses. It should be noted that testing with the stronger effect size suggests that the power is 100% however this is based on only 1000 simulations under each hypothesis and therefore is unlikely to be precisely equal to 100%.

Population size

The relative risks were calculated from simulations performed on the Abernant population using the exact same method used on Nant y Gwyddon described in Section 3.3.3. The scatterplot between the relative risks and the risk attributable to exposure is given in Figure 3.33. The relative risks are not elevated above 1 suggesting no increased risk for those mothers with additional risk attributable to exposure. The absence of raised relative risks despite the increased risk in the exposed region is a result of the small population size and random effects.

Figure 3.33: the scatterplot between the risk attributable to exposure and the relative risks simulated under the alternative hypothesis for the population of Abernant



The ROC curves are compared in Section A1.7 for the Nant y Gwyddon and Abernant populations and the sensitivities at the 5% significance level are given in Table 3.11. The efficiency of the test is greatly reduced by the smaller population because the statistics are

based on fewer numbers and, therefore, the variation in the test statistics is larger. The greater variance results in a smaller degree of separation between the distributions under the null and alternative hypotheses, therefore, the tests do not perform as well as they do with Nant y Gwyddon.

Table 3.11: the sensitivity and 95% confidence intervals at the 5% significance level for the kernel density test with the Nant y Gwyddon and Abernant populations

	Sensitivity (95% CI)	
Test statistics	Nant y Gwyddon	Abernant
Mean	0.683 (0.654, 0.711)	0.114 (0.096, 0.135)
Proportion	0.536 (0.505, 0.567)	0.102 (0.085, 0.122)
Slope	0.649 (0.619, 0.678)	0.157 (0.136, 0.181)
Mean and proportion	0.694 (0.665, 0.722)	0.174 (0.152, 0.199)
Proportion and slope	0.863 (0.840, 0.883)	0.197 (0.174, 0.223)
Slope and mean	0.855 (0.832, 0.876)	0.232 (0.207, 0.259)
Mean, prop and slope	0.898 (0.878, 0.915)	0.390 (0.360, 0.421)

In addition to the loss of efficiency the performance of each of the tests relative to the others alters also. When the hypothesis tests were performed using the Nant y Gwyddon population the best single test statistic was background mean, however, here, it is slope. This shift in relative performance is a result of the changing population distribution. The proportion of relative risk values involved in the calculation of each test statistic changes, therefore, the relative efficiency of each test statistic changes. However, this is not an issue if all three statistics are combined which is clearly beneficial for a population of this small size.

Unknown confounding

In order to assess the impact of unknown confounding on the testing procedure, the simulations were performed with varying baseline risks. The relative risks were then calculated without adjustment for confounding factors. Subsequently the hypothesis tests were assessed as before using ROC curves, the ROC curve comparisons can be seen in Section A1.8 and the sensitivities at the 5% significance levels are given in Table 3.12.

The testing procedure becomes less efficient if confounding is introduced into the data because the non-uniform baseline risks introduce further random variation into the relative risk surfaces resulting in a smaller degree of separation between the distributions of the test statistics under the two hypotheses as expected.

Table 3.12: the sensitivity and 95% confidence intervals at the 5% significance level for the kernel density test with no confounding, unknown confounding and known confounding with adjustment

	Sensitivity (95% CI)		
Test statistics	No Confounding	Unknown Confounding	Known Confounding
Mean	0.683 (0.654, 0.711)	0.608 (0.577, 0.638)	0.448 (0.417, 0.479)
Proportion	0.536 (0.505, 0.569)	0.348 (0.319, 0.378)	0.432 (0.402, 0.463)
Slope	0.649 (0.619, 0.678)	0.605 (0.574, 0.635)	0.604 (0.573, 0.634)
Mean and proportion	0.694 (0.665, 0.722)	0.624 (0.594, 0.654)	0.552 (0.521, 0.583)
Proportion and slope	0.863 (0.840, 0.883)	0.769 (0.742, 0.794)	0.794 (0.768, 0.818)
Slope and mean	0.855 (0.832, 0.876)	0.773 (0.746, 0.798)	0.722 (0.693, 0.749)
Mean, prop and slope	0.898 (0.878, 0.915)	0.831 (0.807, 0.853)	0.851 (0.828, 0.872)

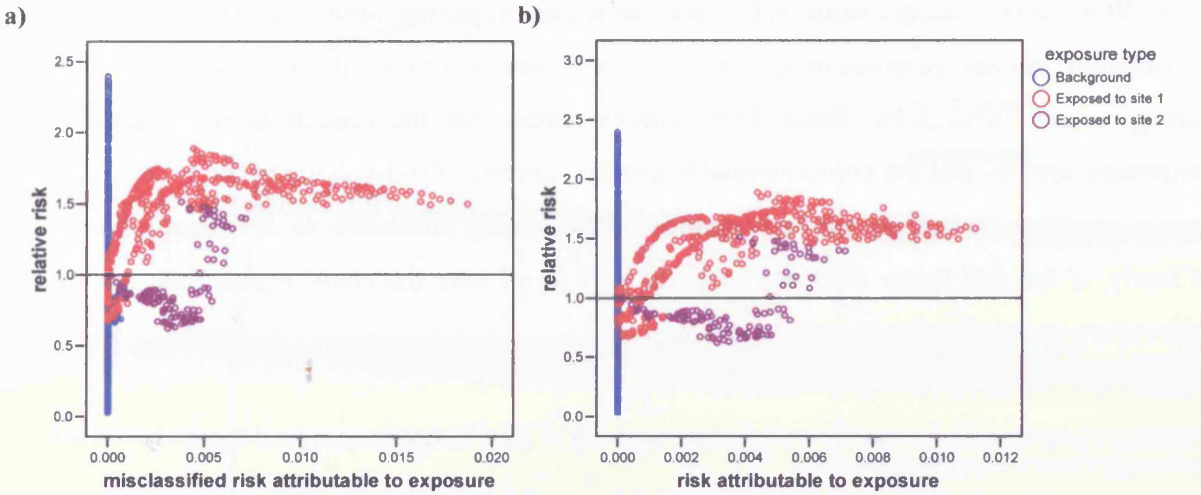
However, what is surprising is that the test appears to be more efficient when there is unknown confounding compared to where the confounding is known and can therefore be adjusted for, except for when the test involves using proportion, the crudest of the test statistics. Despite this increase in efficiency it is not advisable to neglect to adjust for known confounders as this could lead to accepting the alternative hypothesis when it may not be true. Therefore, this could mean that additional risk in the vicinity of the site could be attributed to exposure to the landfill site when in fact it was merely a result of the spatial distribution of these known confounding factors.

In conclusion, it is important to gather all information of possible confounders and make an adjustment for them as part of the testing procedure otherwise they could lead to an incorrect outcome being accepted with undesirable consequences for public health.

Misclassified exposure

The misclassified exposure model, as used in Chapter 2 to investigate the impact of misclassified exposure, was used in the kernel density test. Assuming these misclassified exposure risks are accurate leads to the horizontal shift of points in the scatterplot between risk attributable to exposure and the relative risks. This shift is apparent in Figure 3.34a where the true exposure risks in Figure 3.34b are replaced by the misclassified ones.

Figure 3.34: the scatterplot between the relative risks for one simulation under the alternative hypothesis and a) the misclassified exposure risk and b) the true exposure risk



The ROC curve for testing when the exposure risk has been misclassified is compared to that for when the risk attributable to exposure is correct in Section A1.9 and the sensitivities at the 5% significance level are given in Table 3.13. A misclassification of this magnitude does not greatly affect the performance of the hypothesis tests. The background mean is the most robust test statistic against a small error in quantifying the exposure risk and is, therefore, the best statistic to use if there is some doubt surrounding the accuracy of the exposure risk model. Hypothesis testing using slope is severely affected by this slight deviation as the horizontal shift affects the gradient of the scatterplot. Consequently, if it is suspected that the exposure risks are not accurate it should be avoided unless combined with another test statistic. However, in the case of proportion the testing procedure performs slightly better with the misclassification, which casts aspersions on the usefulness of proportion as a test statistic.

Table 3.13: the sensitivity and 95% confidence intervals at the 5% significance level for the kernel density test with true exposure risks and misclassified exposure risks

	Sensitivity (95% CI)	
Test statistics	True Exposure	Misclassified Exposure
Mean	0.683 (0.654, 0.711)	0.669 (0.639, 0.698)
Proportion	0.536 (0.505, 0.567)	0.569 (0.538, 0.599)
Slope	0.649 (0.619, 0.678)	0.507 (0.476, 0.538)
Mean and proportion	0.694 (0.665, 0.722)	0.699 (0.670, 0.727)
Proportion and slope	0.863 (0.840, 0.883)	0.807 (0.781, 0.830)
Slope and mean	0.855 (0.832, 0.876)	0.810 (0.785, 0.833)
Mean, prop and slope	0.898 (0.878, 0.915)	0.851 (0.828, 0.872)

Spatial distribution of the risk factor

The ROC curve comparisons for when only one exposure source is considered in the simulation process are given in Section A1.10 and the sensitivities at the 5% significance level are given in Table 3.14. These ROC curves reveal that the reduction in the number of exposure sources and the consequentially smaller number of individuals in the exposed region has a negative impact on the performance of the testing procedure as discussed previously. Clearly, if the risk factor does not spatially vary at all over the study region then the testing procedure cannot be considered for that region.

Table 3.14: the sensitivity and 95% confidence intervals at the 5% significance level for the kernel density test with 1 and 2 exposure sources

	Sensitivity (95% CI)	
Test statistics	2 Exposure Sources	1 Exposure Source
Mean	0.683 (0.654, 0.711)	0.489 (0.458, 0.520)
Proportion	0.536 (0.505, 0.567)	0.305 (0.277, 0.334)
Slope	0.649 (0.619, 0.678)	0.561 (0.530, 0.592)
Mean and proportion	0.694 (0.665, 0.722)	0.458 (0.427, 0.489)
Proportion and slope	0.863 (0.840, 0.883)	0.667 (0.637, 0.696)
Slope and mean	0.855 (0.832, 0.876)	0.653 (0.623, 0.682)
Mean, prop and slope	0.898 (0.878, 0.915)	0.800 (0.774, 0.824)

3.3.6 Comparisons with Parametric Tests

The kernel density test for spatial association has maximised efficiency when all three test statistics are considered. In Figure 3.35 the ROC curve for the kernel density test with all three test statistics is compared to that of the trend test for spatial association and Stone's test with the simulations described in Section 2.4.3. Although the kernel density test for spatial association is clearly an improvement on Stone's test, there is not a lot of difference between the kernel density test and the trend test for spatial association using the Nant y Gwyddon data.

The sensitivities at the 5% significance level for the three tests are given in Table 3.15. The kernel density test for spatial association is slightly better with a high power of 89.8% at the 5% significance level. Furthermore, the kernel density approach has an inbuilt powerful contouring tool for mapping the health risk and this visual advantage is a very important one.

Figure 3.35: the ROC curve for the kernel density test for spatial association compared to that of Stone's test and the trend test for spatial association using simulations for Nant y Gwyddon

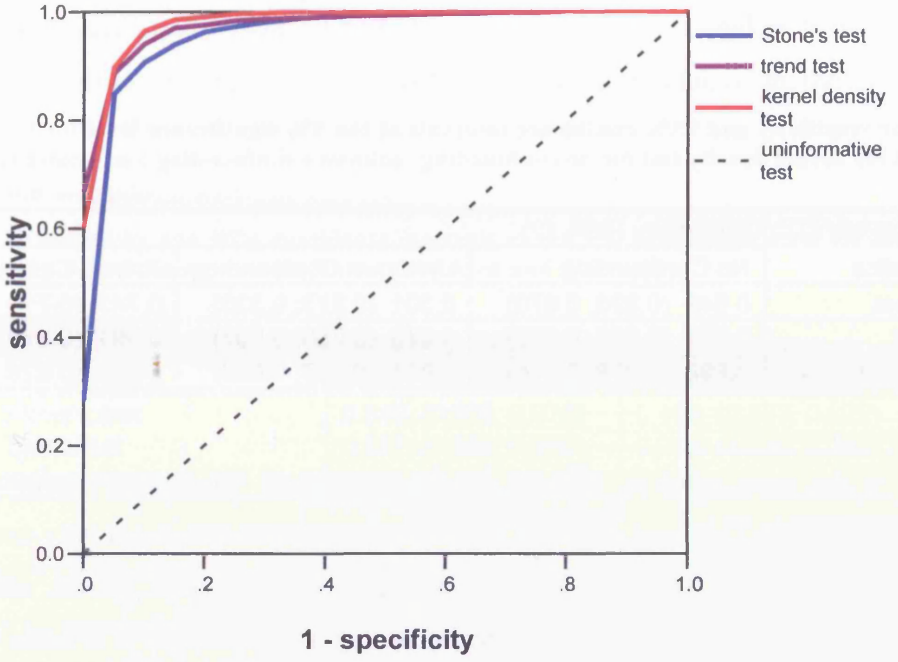


Table 3.15: the sensitivity and 95% confidence intervals at the 5% significance level for Stone's test, the trend test and the kernel density test

Test	Sensitivity (95% CI)
Stone's test	0.849 (0.826, 0.870)
Trend test	0.887 (0.866, 0.905)
Kernel density test	0.898 (0.878, 0.915)

In Sections 2.4.2 and 3.3.4 a number of factors that influence the power of all three testing procedures are discussed. These factors include confounding, effect size, population size, misclassified exposure and the spatial distribution of the risk factor of interest. Consequently, in this section the efficiency of the tests under each different situation is revisited to explore further the relative performance of each test. The ROC curve comparisons under each situation are given in Section A1.11.

In the main assessment of the performance of the tests no confounding was assumed and consequently no adjustment was required. If there is known confounding factors then each of the testing procedures must make adjustments for these factors and as previously discussed adjusting for confounders influences the performance of the testing procedure. In many situations there is unknown confounding which means that an adjustment cannot be made

within the testing procedure. The sensitivities at the 5% significance level for the situations where there is no confounding, known confounding and unknown confounding are given in Table 3.16.

Table 3.16: the sensitivity and 95% confidence intervals at the 5% significance level for Stone's test, the trend test and the kernel density test for no confounding, unknown confounding and known confounding

	Sensitivity (95% CI)		
Test statistics	No Confounding	Unknown Confounding	Known Confounding
Stone's test	0.849 (0.826, 0.870)	0.301 (0.273, 0.330)	0.747 (0.719, 0.773)
Trend test	0.887 (0.866, 0.905)	0.878 (0.856, 0.897)	0.891 (0.870, 0.909)
Kernel density test	0.898 (0.878, 0.915)	0.831 (0.807, 0.853)	0.851 (0.828, 0.872)

Stone's test is the most severely affected by the adjustment for known confounding or by unknown confounding. This loss of power indicates that Stone's test is not very useful in most situations where there is likely to be some kind of confounding either known, unknown or a combination of the two. Therefore, Stone's test is not a good test to use in relation to the other testing procedures. Both the trend test for spatial association and the kernel density test still perform well when there is known or unknown confounding. The power is reduced to a greater extent in both tests when there is unknown confounding as expected but not to such an extent that the tests are no longer informative. The trend test for spatial association performs slightly better than the kernel density test when confounding is involved but the differences are small.

The effect size affects the performance of the testing procedures with a decline in the power of the test with a smaller effect size. The sensitivities at the 5% significance level are given in Table 3.17 for a weaker and stronger effect size. It is clear that when there is a small effect size the kernel density test is by far the most efficient testing procedure with a power of 63.0% at the 5% significance level.

Table 3.17: the sensitivity and 95% confidence intervals at the 5% significance level for Stone's test, the trend test and the kernel density test with a weaker, original and stronger effect size

	Sensitivity (95% CI)		
Test statistics	Weaker Effect Size	Original Effect Size	Stronger Effect Size
Stone's test	0.339 (0.310, 0.369)	0.849 (0.826, 0.870)	0.986 (0.977, 0.992)
Trend test	0.440 (0.410, 0.471)	0.887 (0.866, 0.905)	1.000 (0.996, 1.000)
Kernel density test	0.630 (0.600, 0.659)	0.898 (0.878, 0.915)	1.000 (0.996, 1.000)

If the study population size is smaller then the testing procedures are less efficient. The sensitivities at the 5% significance level for Nant y Gwyddon and Abernant are given in Table 3.18. When considering Abernant, where the population size is smaller than it is in Nant y Gwyddon, the kernel density test is the most efficient test relative to the two parametric approaches.

Table 3.18: the sensitivity and 95% confidence intervals at the 5% significance level for Stone's test, the trend test and the kernel density test for Nant y Gwyddon and Abernant

	Sensitivity (95% CI)	
Test statistics	Nant y Gwyddon	Abernant
Stone's test	0.849 (0.826, 0.870)	0.196 (0.173, 0.222)
Trend test	0.887 (0.866, 0.905)	0.252 (0.226, 0.280)
Kernel density test	0.898 (0.878, 0.915)	0.390 (0.360, 0.421)

In the majority of situations in which these three testing procedures can be used there will be some degree of misclassification in the exposure surface and despite slight misclassifications it is still desirable to be able to detect a spatial association between the risk factor and the health outcome. All three testing procedures experience a reduction in power when there is misclassified exposure, as expected, and the sensitivities for each of the testing procedures at the 5% significance level are roughly equal, see Table 3.19. Clearly if there is serious doubt surrounding the accuracy of the exposure surface then the testing procedures should not be implemented.

Table 3.19: the sensitivity and 95% confidence intervals at the 5% significance level for Stone's test, the trend test and the kernel density test with true exposure and misclassified exposure

	Sensitivity (95% CI)	
Test statistics	True Exposure	Misclassified Exposure
Stone's test	0.849 (0.826, 0.870)	0.835 (0.811, 0.857)
Trend test	0.887 (0.866, 0.905)	0.846 (0.822, 0.867)
Kernel density test	0.898 (0.878, 0.915)	0.851 (0.828, 0.872)

If the spatial distribution of the risk factor is more homogeneous then the testing procedures become less powerful. In the situation where there is only one exposure source as opposed to two, i.e. exposure to the risk factor is more homogenous, the kernel density test is the most powerful test at the 5% significance level, see Table 3.20.

Table 3.20: the sensitivity and 95% confidence intervals at the 5% significance level for Stone's test, the trend test and the kernel density test with 2 exposure sources and 1 exposure source

	Sensitivity (95% CI)	
Test statistics	2 Exposure Sources	1 Exposure Source
Stone's test	0.849 (0.826, 0.870)	0.664 (0.634, 0.693)
Trend test	0.887 (0.866, 0.905)	0.749 (0.721, 0.775)
Kernel density test	0.898 (0.878, 0.915)	0.800 (0.774, 0.824)

In summary Stone's test is the least powerful test in all situations considered above and consequently should not be used when individual level data is available. Clearly if data is only available in an aggregated form then Stone's test is the only one of the three testing procedures discussed that can be used. If there is some kind of confounding whether it is known or unknown then the trend test for spatial association is the most efficient testing procedure. However, the kernel density test is not that different in terms of performance when there is confounding. If there is a slight misclassification in the exposure surface then the kernel density test is the most powerful testing procedure although the power is not that different for the other two tests. Finally, in situations where there is a smaller effect size, smaller population size or a more homogeneous exposure surface the kernel density test far surpasses the other two testing procedures in terms of performance.

Consequently, the kernel density test for spatial association is the only test that is consistently efficient and therefore is the recommended approach for testing for a spatial association between the health outcome and the risk factor with individual level data. The additional advantage to using the kernel density approach is the inbuilt technique for mapping risk which is superior to mapping summary statistics for the health outcome in aggregated regions.

3.4 Approaches to Spatial Analysis

The first step in spatially analysing a health outcome is to map the risk of the outcome in the study region to seek out any patterns in risk with particular emphasis on finding clusters or regions of elevated risk. This aim can only be achieved by mapping the SER's if the health data is given at area level. However, if individual level data is provided, as it has been for both congenital malformations and sex ratios, then it is unnecessary to aggregate the data in order to plot the SER's. The individual level data can be used to create relative risk contours which have a number of advantages. When considering plotting SER's with individual level data the study region must be partitioned into smaller sub-regions the choice of which is arbitrary yet

the boundaries used will influence the image obtained. However, when plotting the relative risk contours the individual level data is used directly and no aggregation is required. Therefore, the image seen is only influenced by the data itself and not by predetermined boundaries. In addition, the use of regions in the SER approach means that any pattern that occurs within that region will not necessarily be detected as the risk will be mapped as uniform throughout the area, whereas the relative risk contours are not constrained in this way. Therefore, the relative risk contours are a much better method of mapping risk. When plotting relative risk contours it is recommended that regions of low population density are filtered out so that any regions of elevated risk that occur as a result of a low denominator do not appear.

If the relative risk contours reveal that there are regions of elevated risk of the health outcome then their statistical significance can be assessed by applying the global significance test. If the three testing procedures involved in the global significance test are all significant then there is evidence that the hotspot is not a result of random variation and will require further investigation to determine what may have caused it. In Chapter 2 this investigation of spatial patterns was achieved with the likelihood based test for homogeneity of risk. However, the likelihood based test for homogeneity of risk requires the same aggregation as mapping SER's and, therefore, is subject to the same problems. Therefore, the global significance test is recommended for individual level data. However, if the health data is only available at the area level then there is no option but to use the likelihood based test for homogeneity of risk. The added benefit of the global significance test is that it is non-parametric and, therefore, unlike in the likelihood based test for homogeneity no assumptions are made surrounding the distribution of the health outcome.

If there are statistically significant spatial patterns in the risk of the health outcome then there is likely to be a risk factor that influences the health outcome. Therefore, possible risk factors can be considered and a spatial association between exposure to the specified risk factor and the health outcome of interest can be assessed via hypothesis testing. In Chapter 2 the recommendation was to use Stone's test on aggregated data and the trend test for spatial association when the health data is given at the individual level. However, the kernel density test for spatial association is an alternative method that can also be applied to individual level data. The kernel density test and trend test for spatial association are equally efficient in a number of situations but when the effect size is small, the population size is small or the

exposure risk is not very variable over the study region then the kernel density test for spatial association performs better. This improved performance coupled with the advantages of using relative risk contours as a mapping tool mean that the recommendation in this thesis is that the kernel density test should be used with individual level data.

When applying the kernel density method for assessing the spatial association between the risk factor and the health outcome of interest all three test statistics should be used. The smoothing parameters used to calculate the relative risks and in the discriminant analysis should be similar to those that result from the normal optimal smoothing method of obtaining optimal smoothing parameters. In addition, the dataset should not be filtered by population density so all individuals in the population are included in the analysis. Problems with small effect and population sizes can be overcome to some extent by performing multiple site studies.

In conclusion, if the health data is given at an area level then the obvious option is to map the SER's in each area, test for any spatial patterns using the test for homogeneity of risk and finally to apply Stone's test to assess the existence of a spatial association between the health outcome and any possible risk factors. However, if the health data is given at the individual level then this same process is achieved through plotting the relative risk contours, applying the global significance test and, subsequently, using the kernel density test for spatial association to investigate relationships between the health outcome and possible risk factors. The congenital malformation and sex ratio data has been given with geographical grid coordinates for each individual in the study population. Therefore, in Chapters 4 and 5 the main methodologies used will be those discussed in this Chapter.

Chapter Four – Application of Spatial Methodology to Birth Defects

There has been growing concern that exposure to contaminants released from landfill sites increases the risk of congenital malformations. Investigations into the link between exposure to toxicants released from landfill sites and birth defects are discussed in detail in Section 1.2.6. Some of these studies are multiple site studies and the remainder are single site studies.

A major multiple site investigation was performed in Europe by Dolk et al. on 21 landfill sites in 5 European countries and found a statistically significant increase in the risk of congenital anomalies for mothers residing in close proximity to the landfill sites [64]. In fact, those mothers residing within 3km of any of the landfill sites were found to be 33% more likely to have a congenitally malformed birth than those residing between 3km and 7km of the sites. In a similar study using 23 European landfill sites it was found that births to mothers residing within 3km of the centre of any of the sites were 41% more likely to have a chromosomal anomaly than those to mothers residing between 3km and 7km from the sites [65].

Another major multiple site study was carried out by Elliott et al. on 19,196 British landfill sites [68]. In contrast to the European study, the results indicated that there was a statistically significant increased risk of only 1% associated with living within 2km of the centre of any of the sites. A similar study of only Scottish landfill sites found no significant increase in the risk of congenital anomalies associated with proximity to the landfill sites [70].

An independent multiple site study focused on 20 Welsh landfill sites that opened between 1983 and 1997 [73]. The relative risk of congenital anomalies within 2km of the centre of the sites was 0.87 (95% CI: 0.75-1.00) before opening and 1.21 (95% CI: 1.04-1.40) after opening with adjustment for maternal age, year of birth, hospital of birth, gender and socioeconomic deprivation. The standardized risk ratio was 1.39 (95% CI: 1.12-1.72) which suggests that there is a statistically significant increase in the risk of birth defects within 2km of the sites after their opening. However, it should be noted that there was a low relative risk before the opening of the site and the high standardized risk ratio may be a response to that and not necessarily be because of a high relative risk after opening.

The disparities between the results of the studies could be caused by the different definitions used for the exposed and unexposed regions. However, the differences are likely to be a

response to the definitions used for a landfill site. For example, in the British study by Elliott et al. all landfill sites are considered, many of which will be benign, and consequently any causal link is more difficult to identify [69]. This highlights the major problem with multiple site studies as discussed by James et al. [10]. Individual landfill sites having different characteristics in terms of the quantity and type of waste and also with respect to leachate and air dispersion control. Additionally, the topography and meteorology in the local area surrounding the landfill site influences the distribution of the population and pollution and is site-specific. Consequently, multiple site studies are subject to problems that arise from these variations. Furthermore, the protection of the health of the public is a local issue and should be supported by studies that assess public health in a local setting. Consequently, the focus of this Chapter lies in a single landfill site, Nant y Gwyddon. The results of two single site studies have been published about the risk of congenital anomalies around the Nant y Gwyddon landfill site.

These two single site studies focused on the risk of congenital anomalies in relation to the Nant y Gwyddon landfill site from 1983 to 1997. In the earliest study the exposed region comprised five wards from which complaints had been received regarding the site and that were within 3km of the centre of the site. The unexposed region included 22 wards with similar socioeconomic level to those wards in the exposed region [9]. The investigation revealed that the incidence of all congenital anomalies was raised in the exposed wards relative to the unexposed wards after the opening of the site, but there is also an increased incidence in the exposed wards that predates the opening of the site. Hence, the results of this study do not support the hypothesis that the opening of the landfill site affected the risk of congenital anomalies.

In a later study by James et al. contours for the relative risk of congenital anomalies were created in the 20km square region centred over the Nant y Gwyddon landfill site [10]. The relative risk surface was adjusted for maternal age, year of birth, gender, hospital of birth and socioeconomic status. The adjusted relative risk contours revealed two regions over which the relative risk surface was elevated above 2: one to the north east of the landfill site within 3km and a second to the northwest at about 7km from the site. The existence of elevations in the relative risk near to the Nant y Gwyddon landfill site suggests that exposure to substances released from the landfill site may influence the risk of congenital anomalies.

In the thesis written by the same author relative risk contours are generated comparing the density of congenital malformations before the opening of the site with that after the opening of the site [272]. The resultant before and after relative risk contours indicated that there were two regions over which the density of birth defects after opening was twice that before opening within 3km of the landfill site. This indicates that the opening of the Nant y Gwyddon landfill site could have lead to an increased risk of congenital malformations in regions that are in close proximity to the site itself. Thus, this conclusion is in contrast to that drawn by Fielder et al. which may be a consequence of using different methodologies in the analysis.

In James et al.'s concluding remarks it is indicated that there is a need to develop kernel density methodology to generate a formal test for spatial association between the risk of a specified health outcome and exposure to environmental pollution [10]. Additionally, they highlight the problem of multiple comparisons associated with using significance contours representing point differences as a tool in identifying regions of excess density of a health outcome. The desire to overcome this difficulty calls for the development of methodology that enables the calculation of the global statistical significance of regions of excess density of a given health outcome. The methodology to achieve both these objectives for future investigation has been developed in Chapter 3.

A number of applications of the methods described in Chapter 3 are considered in this Chapter and these include:

- exploring the spatial distribution of the risk of congenital anomalies around Nant y Gwyddon; and
- examining the spatial distribution of congenital malformations around Nant y Gwyddon using more recent data to assess whether the elevations in relative risk are still present.

Unfortunately, access to good quality exposure data for Nant y Gwyddon was not forthcoming and consequently it was not possible to test for the spatial association between exposure to chemicals released from the landfill site and birth defects. However, the methodology developed can be used to test for the spatial association between two relative risk/density surfaces. Therefore, the kernel density test for spatial association was applied to

- checking the adequacy of the adjustment in the relative risk surface for confounding; and

- investigating the differences between chromosomal and non-chromosomal anomalies.

In this Chapter, the risk of congenital anomalies around the Nant y Gwyddon landfill site is initially considered. The relative risk of congenital anomalies around the Nant y Gwyddon landfill site is mapped for 1983 to 1997 and the global significance of regions of elevated relative risk is assessed. Possible risk factors for the reported rate of congenital anomalies are investigated including maternal age, gender, year of birth, hospital of birth and socioeconomic deprivation. These risk factors are adjusted for in the relative risk contours and the global significance test is performed on regions of elevated adjusted relative risk.

The method of adjusting relative risk surfaces for known confounding factors has never been questioned. The procedure for testing for a spatial association between a health outcome and exposure can be used to assess the adjustment for confounding used in generating relative risk contours. The comparison of the adjusted relative risk contours for all congenital anomalies with that of male births, for example, should reveal that there is no relationship between the two if the adjustment for gender is correct. Consequently, the second part of this Chapter is an assessment of the method used to adjust the relative risk surface for confounding factors using the adjusted relative risk contours for congenital anomalies around Nant y Gwyddon between 1983 and 1997.

The third application of the methodology to congenital malformations focuses on demonstrating how the kernel density test for spatial association between a health outcome and environmental exposure can be utilised to test for the spatial association between two health outcomes. There are numerous different types of congenital malformations, many of which have differing aetiologies. One of the ways in which congenital malformations can be categorised is by chromosomal and non-chromosomal anomalies. Therefore, in order to illustrate the test for spatial association between two health outcomes, chromosomal and non-chromosomal anomalies are considered.

The demonstration of the methodology using chromosomal and non-chromosomal defects is important in terms of public health because different types of congenital anomalies are typically grouped together in studies on the risk of birth defects around landfill sites. The reason for grouping all birth defects together for spatial analysis is that congenital malformations are rare events with only around 2-3% of births affected [268]. If only one type

of congenital malformation is considered then cases become even rarer, for example, 1 in 1,000 births have a neural tube defect and gastroschisis occurs in only 1 in 10,000 births [23]. Consequently considering only one type of congenital malformation significantly reduces the power of any statistical analysis. Thus to achieve higher statistical power most investigators have to group different types of birth defects together for their study [37]. However, most types of congenital malformations probably have unique aetiologies and hence grouping can prevent real effects related to exposure to certain risk factors from being identified. If there is a difference in the aetiologies of chromosomal and non-chromosomal anomalies this would be reflected in a lack of spatial association between the two types of birth defects.

The final application relates to the age of the dataset. The most recent data used to fit the relative risk contours around the Nant y Gwyddon landfill site dates back to 1997 and thus it is not clear whether the issue of elevated relative risk around the landfill site is still a current problem. Data on congenital anomalies and related risk factors for 1998 to 2004 are considered to assess whether the regions of elevated relative risk are still present using this more recent data. The relationships between the risk of congenital anomalies and the five risk factors that were investigated for 1983 to 1997 are reviewed for 1998 to 2004. Relative risk contours are then produced for this more recent time period with adjustment for confounding factors. Direct comparisons between the two time periods are not possible because the introduction of the Congenital Anomaly Register and Information Service for Wales (CARIS) in 1998 has influenced the reporting rates of congenital malformations. Prior to undergoing all this detailed analysis the spatial distribution of congenital anomalies around the Nant y Gwyddon landfill site is considered.

4.1 Risk of Congenital Malformations around Nant y Gwyddon 1983-1997

The collection of information on congenital anomalies on a national basis was proposed by the Chief Medical Officer in 1963 following the Thalidomide epidemic [277]. The National Congenital Anomaly System (NCAS) records information on congenital anomalies for live and still births in England and Wales and is run by the Office of National Statistics (ONS). Conditions are coded to the ICD-10 international statistical classification of diseases. Information is provided on a voluntary basis by local community trusts or health authorities using SD56 forms and if there are no congenital anomalies in a given month then a nil return should be filled out on one of these forms and returned [268].

Denominator data was provided by the Child Health System (CHS) at the Office of National Statistics which was linked to data from NCAS. Information provided by this merging of data included the location of the residence at birth (based on the postcode), the date of birth, the gender of the birth, birth weight, gestational age, the age of the mother, the hospital of birth, Townsend score, whether or not the birth had at least one congenital anomaly and if so whether the anomaly was of a chromosomal nature. The postcode for each birth was converted to east and north grid coordinates using the British National Grid reference system.

The grid coordinates for Nant y Gwyddon were provided by the Environment Agency as were those for the other 31 landfill sites. The landfill sites considered contained a large volume of waste, were in close proximity to residential areas and had either opened or undergone significant changes in use during the study period between 1983 and 1997. The locations of all 32 landfill sites were required to fit the logistic regression model used to adjust the relative risk contours for confounding factors discussed in detail later on. The study population of interest included all births between 1983 and 1997 to mothers who resided within the 20km square region centred over the Nant y Gwyddon landfill site at the time of birth. This part of the investigation of congenital anomalies focuses on their spatial distribution in relation to the location of the Nant y Gwyddon landfill site. Information on known risk factors is used to determine to what extent these factors explain the spatial patterns observed and whether there is still some unexplained spatial heterogeneity after adjustment for these factors that may be attributable to exposure to contaminants released from the landfill site.

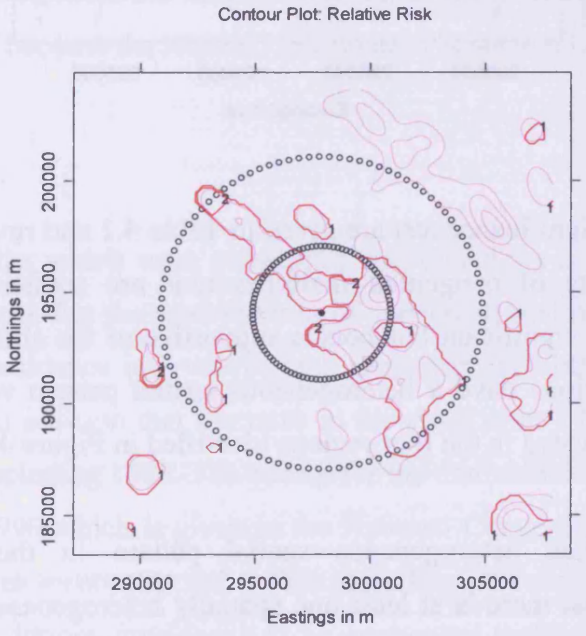
4.1.1 Crude Risk of Congenital Malformations

In the 20km square region centred over the Nant y Gwyddon landfill site the recorded rate of birth defects from 1983 to 1997 was 1.23%, which is higher than the rate for the whole of Wales of 1.13% for that period. This rate is low compared to the 2-3% rate for major congenital anomalies reported by Boyd et al. [268]. Boyd et al. indicate that given the voluntary nature of reporting, the ascertainment rate for NCAS was only approximately 40% between 1991 and 1999, which would explain the relatively low rate obtained from the dataset for 1983 to 1997.

The concentric circle approach as discussed in Chapter 2 reveals that the relative risk of a congenital anomaly for those exposed to the landfill site, i.e. those within 3km of the centre of the site, is 1.303 (95% CI 1.031, 1.646) where the unexposed group was treated as those residing between 3km and 7km from the site. This result suggests that there is a statistically significant increase in risk for births to mothers who resided within 3km of the centre of Nant y Gwyddon at the time of the birth. However the use of distance as a way of predetermining the exposed and unexposed regions is not accurate because exposure to contaminants released from the landfill site will be influenced by other additional factors such as topographical and meteorological variables. Therefore a better initial step in the approach to investigating the influence of the landfill site on the risk of congenital malformations is to map contours of the relative risk of birth defects in the study region.

The relative risk contours of all birth defects without adjustment for confounding in the 20km square region centred over Nant y Gwyddon are given in Figure 4.1. The lighter contours represent the probability contours for all births over the same region indicating the regions over which the density of all births is high.

Figure 4.1: relative risk contours of congenital malformations in the 20km square region centred over Nant y Gwyddon 1983-1997



There are four regions over which the relative risk is greater than 2 and these are considered to be hotspots of excess density of congenital anomalies and are presented in Figure 4.2. The

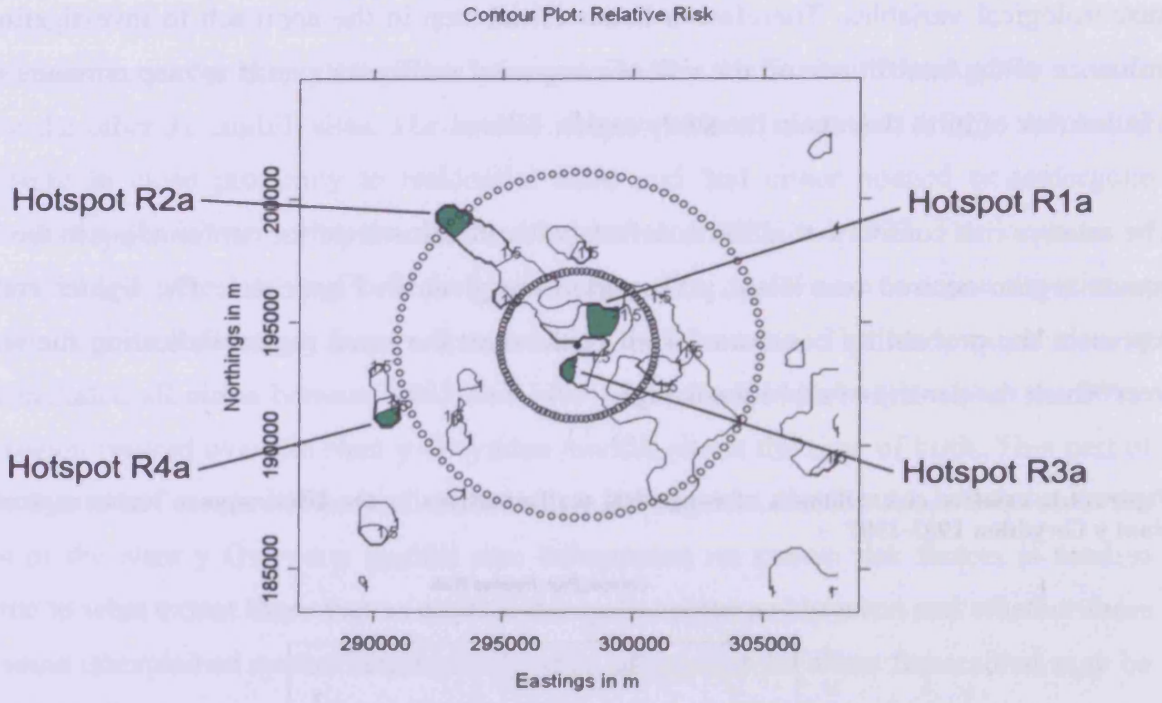
global significance test has been used to assess the statistical significance of these elevations in the relative risk contours. The hypotheses for the global significance test are as follows:

H_0 : the spatial patterns in the relative risk contours are homogeneous

and

H_1 : the spatial patterns in the relative risk contours are heterogeneous.

Figure 4.2: hotspots of excess density of congenital malformations in the 20km square region centred over Nant y Gwyddon 1983-1997



The results of the global significance test are given in Table 4.1 and reveal that all four of the hotspots of excess density of congenital malformations are statistically significant. The existence of 4 statistically significant hotspots is supportive of the alternative hypothesis i.e. that congenital malformations have a heterogeneous spatial pattern where the relative risk surface is significantly elevated in the four regions identified in Figure 4.2.

The statistically significant heterogeneous spatial pattern in the risk of congenital malformations suggests that there is at least one spatially heterogeneous risk factor for birth defects. The dataset produced by merging NCAS and the CHS contains information on maternal age, socioeconomic deprivation, year of birth, hospital of birth and gender for each of the births recorded. Consequently, the relationship between each of these risk factors and

the risk of congenital malformations is investigated below to see if any could be associated with the regions of high relative risk.

Table 4.1: results of the global significance test on hotspots of excess density of congenital malformations in the 20km square region centred over Nant y Gwyddon 1983-1997

Hotspot		R1a	R2a	R3a	R4a
Number of births		425	602	269	346
Number of congenital malformations		16	18	10	15
Observed mean relative risk		2.479	2.251	2.408	2.355
p-value	test 1	<0.001	<0.001	<0.001	<0.001
	test 2	<0.001	<0.001	0.003	<0.001
	test 3	<0.001	<0.001	0.001	<0.001

4.1.2 Possible Risk Factors

The relationships between the risk of congenital malformations and five possible risk factors are investigated using all births in Wales between 1983 and 1997. These five risk factors include year of birth, gender, maternal age, socioeconomic deprivation and hospital of birth which are factors that were used in adjustments of risk made for confounding in two studies of Welsh landfill sites [10, 73]. There are many other possible risk factors, for example, maternal drug use, viral infections, maternal health disorders, environmental agents and radiation that cannot be investigated because the required information is not available.

Year of birth

The proportion of births which were congenital malformations is given by year of birth in Figure 4.3 which reveals that the reported rate of congenital malformations has a decreasing trend by year. The confidence intervals for the proportion of birth defects in each year are given in Table 4.2 and indicate that the rates in the years after 1989 are significantly lower than those up to and including 1989. The reason for the dramatic decline is the introduction of the exclusion list in 1990 which is given in the National Congenital Anomaly System report [277]. The exclusion list meant that rather than recording all congenital malformations, some birth defects were no longer considered to be congenital malformations and consequently were not recorded at all, for example, cerebral palsy, congenital deafness and inter-uterine growth retardation. Furthermore, some congenital malformations were considered to be minor

anomalies and were only recorded if they occurred in combination with other anomalies, for example, clicking hip. The exclusion of these birth defects from the National Congenital Anomalies System meant that the proportion of congenital anomalies dropped significantly.

Figure 4.3: the proportion of births in Wales with a congenital malformation by year of birth 1983-1997

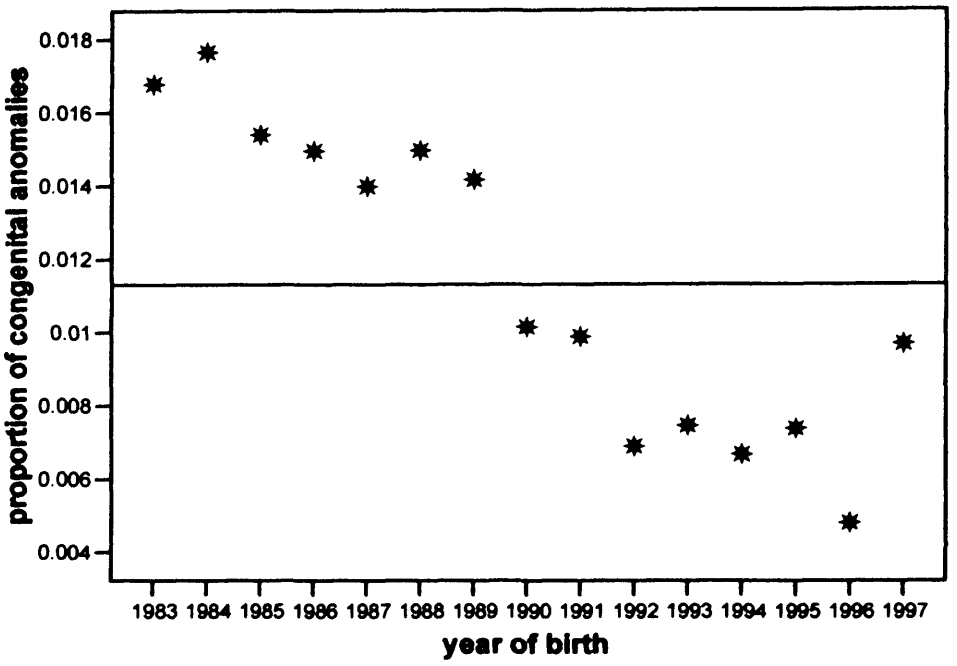


Table 4.2: the proportion of births in Wales with a congenital malformation with 95% confidence intervals by year of birth 1983-1997

Year of Birth	Prop. of Births with Defects (95% CI)	Total No. of Births
1983	0.0168 (0.0154, 0.0183)	31598
1984	0.0177 (0.0163, 0.0192)	32634
1985	0.0154 (0.0141, 0.0168)	33979
1986	0.0149 (0.0137, 0.0162)	35125
1987	0.0140 (0.0129, 0.0153)	36707
1988	0.0150 (0.0138, 0.0163)	37901
1989	0.0142 (0.0131, 0.0155)	37509
1990	0.0101 (0.0092, 0.0122)	38466
1991	0.0099 (0.0090, 0.0110)	37761
1992	0.0069 (0.0061, 0.0078)	37331
1993	0.0074 (0.0066, 0.0083)	36380
1994	0.0067 (0.0059, 0.0076)	35241
1995	0.0074 (0.0065, 0.0084)	34319
1996	0.0048 (0.0041, 0.0056)	34784
1997	0.0097 (0.0087, 0.0108)	34274

A further change to the reporting of congenital anomalies occurred in 1995 where rather than only recording birth defects that were identified within 10 days of birth, NCAS could be notified of a congenital anomaly whenever they were ascertained [278]. This change in the

reporting could explain the increase in the proportion of congenital malformations in 1997. 1995 and 1996 did not have higher reporting rates of birth defects but this may be because it would take a while for the longer reporting window to influence the annual rate.

The small fluctuations in the annual rates of congenital malformations could be the result of changes in the diagnostic procedures and different levels of under reporting in a given time period [278]. The changes in the proportion of birth defects between years provide the motivation for including year of birth in the logistic regression model for predicting the probability of a congenital anomaly for each birth.

Gender

The proportion of births with a congenital anomaly and the confidence intervals for males and females are given in Table 4.3. The proportion of females with a congenital anomaly is significantly lower than the proportion of males with a birth defect suggesting that gender should be included in the adjustment made for confounding risk factors.

Table 4.3: the proportion of births in Wales with a congenital malformation with 95% confidence intervals by gender 1983-1997

Gender	Prop. of Births with Defects (95% CI)	Total No. of Births
Male	0.0129 (0.0125, 0.0133)	273587
Female	0.0096 (0.0092, 0.0100)	260422

Maternal age

Figure 4.4 and Table 4.4 reveal that the rate of congenital malformations is fairly constant by maternal age except in the over 40 category where the proportion of birth defects is higher. The increase in proportion is not statistically significant but there are relatively few mothers giving birth at these older ages (χ^2 test for linear association: p-value = 0.601). Dolk indicates that older maternal age is associated with an increased risk of chromosomal aneuploidies, for example Down's syndrome, which is concurrent with the results seen here [23]. However, in contrast increased risk of gastroschisis is associated with younger maternal age. Dolk et al. indicate that 8% of the registered cases of congenital anomalies suffer from Down's syndrome [36] and that equates to 1 in 1000 births compared to 1 in 10,000 births

having gastroschisis [23]. Therefore, adjusting for maternal age is an adjustment for an overall effect that might not be appropriate for an individual condition.

Figure 4.4: the proportion of births in Wales with a congenital malformation by maternal age 1983-1997

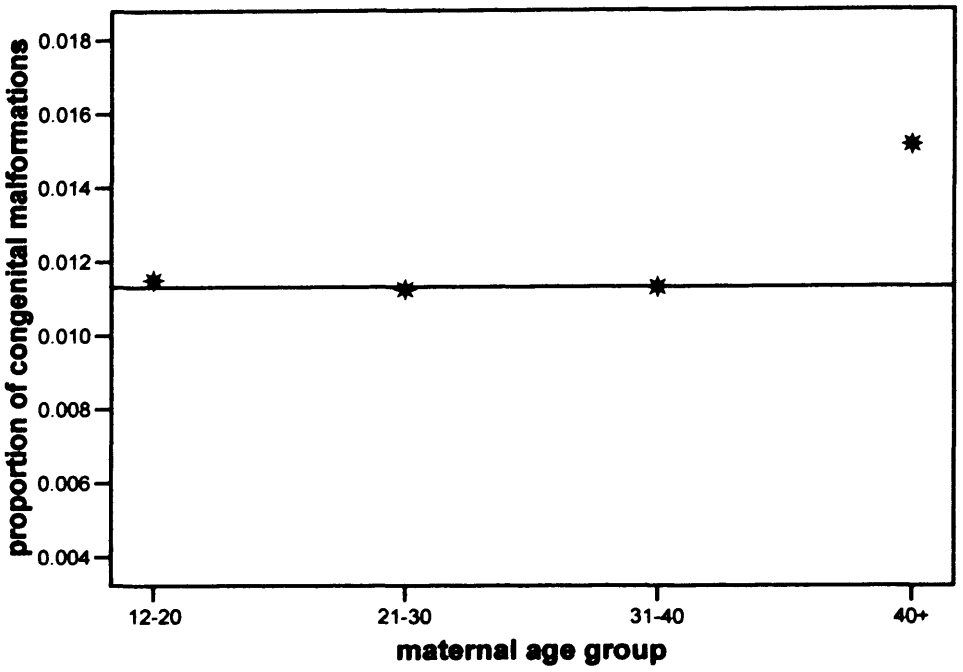


Table 4.4: the proportion of births in Wales with a congenital malformation with 95% confidence intervals by maternal age 1983-1997

Maternal Age Group	Prop. of Births with Defects (95% CI)	Total No. of Births
12-20	0.0115 (0.0107, 0.0124)	63078
21-30	0.0112 (0.0108, 0.0116)	333811
31-40	0.0113 (0.0107, 0.0119)	131897
>40	0.0151 (0.0122, 0.0188)	5223

Socioeconomic deprivation

Socioeconomic deprivation is determined using the Townsend score which is derived from the following predictors:

1. the percentage of private households with no car;
2. the percentage of private households not in owner-occupied accommodation;
3. the percentage of private households with more than one person per room; and
4. unemployed persons as a percentage of the economically active.

A high Townsend score indicates a higher level of socioeconomic deprivation. The Townsend score is used to split the population into five categories of socioeconomic deprivation referred

to here as Townsend quintiles: most deprived, next deprived, median, next affluent and most affluent. The Townsend score for each mother was allocated using the postcode of the residence of the mother at the time of birth

The proportion of congenital malformations by Townsend quintile is given in Figure 4.5 and Table 4.5. There does seem to be a slightly higher risk amongst the more deprived groups although the variation amongst groups is not large and is not statistically significant (χ^2 test: p-value = 0.818). A common scale has been used for Figure 4.3 to Figure 4.6 to aid comparison.

Figure 4.5: the proportion of births in Wales with a congenital anomaly by Townsend quintile 1983-1997

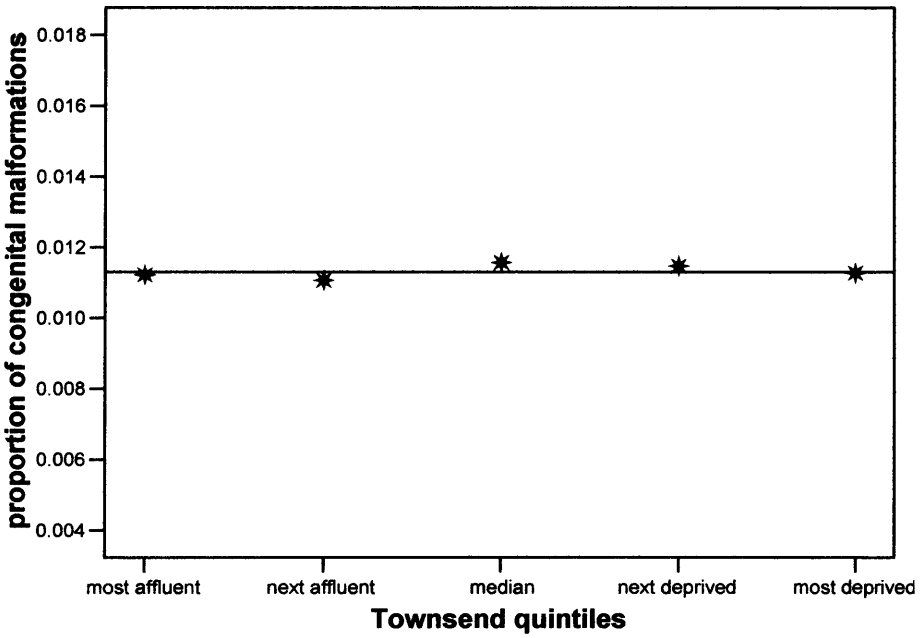


Table 4.5: the proportion of births in Wales with a congenital malformation with 95% confidence intervals by Townsend quintile 1983-1997

Townsend Quintiles	Prop. of Births with Defects (95% CI)	Total No. of Births
Most Affluent	0.0112 (0.0106, 0.0118)	106718
Next Affluent	0.0111 (0.0105, 0.0117)	106836
Median	0.0116 (0.0110, 0.0123)	106928
Next Deprived	0.0115 (0.0108, 0.0121)	106754
Most Deprived	0.0113 (0.0106, 0.0119)	106773

Hospital of birth

The proportion of congenital malformations by hospital of birth is given in Figure 4.6 and the confidence intervals are provided in Table 4.6. The reporting rate of birth defects differs by hospital with Nevill Hall having the lowest reporting rate of 0.59% and County Hospital Hereford having the highest at 1.68%. The reporting rate varies between hospitals because it depends on the diagnostic processes adopted by the staff and also on the diligence of the midwives in notifying NCAS of cases of birth defects [278]. Some midwives may have chosen not to report some minor anomalies whereas others will have reported every congenital malformation prior to the introduction of the exclusion list, thus extenuating the hospital effect. Boyd et al. suggest that ascertainment rates vary by hospital because some prioritise notification more than others [268]. Additionally, mothers with a high risk of congenital malformations are often referred to hospitals that provide specialist services which would increase the rate for those hospitals to which they are referred [23]. The disparities in the reporting and notification of congenital malformations between hospitals motivated the inclusion of hospital of birth in the logistic regression model.

Figure 4.6: the proportion of births in Wales with a congenital malformation by hospital of birth 1983-1997

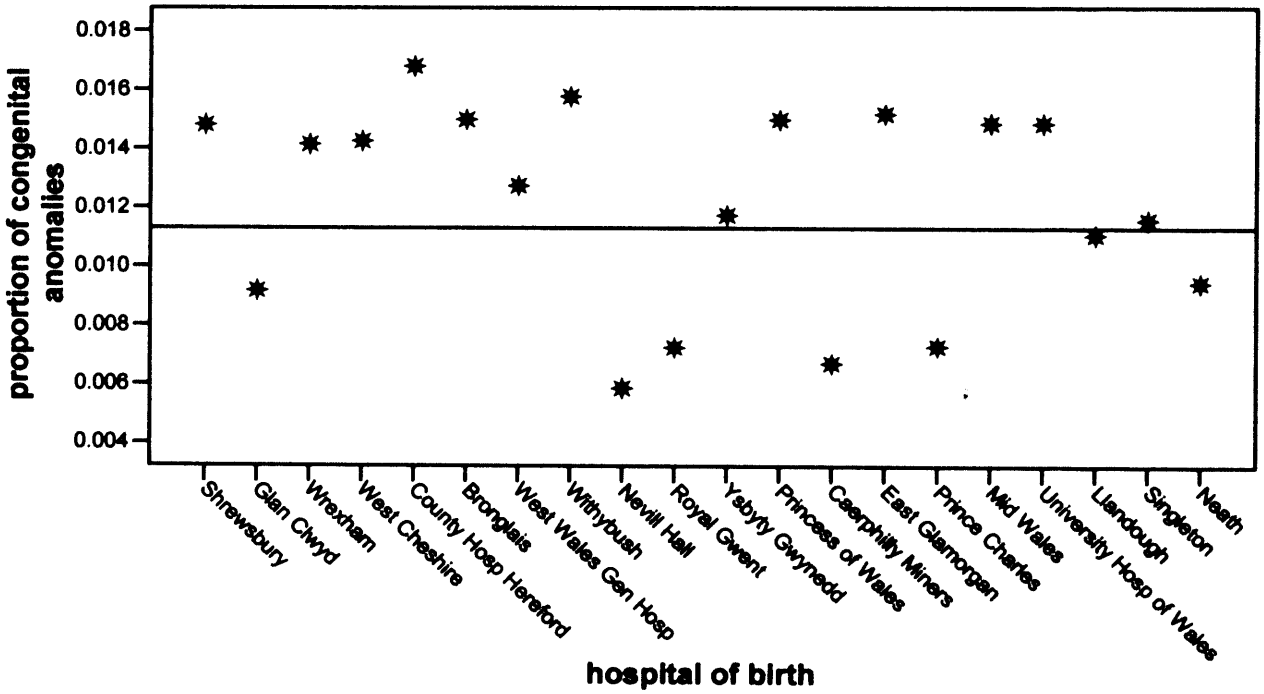


Table 4.6: the proportion of births in Wales with a congenital malformation with 95% confidence intervals by hospital of birth 1983-1997

Hospital of Birth	Prop. of Births with Defects (95% CI)	Total No. of Births
Shrewsbury	0.0148 (0.0121, 0.0181)	6349
Glen Clwyd	0.0092 (0.0083, 0.0102)	38037
Wrexham	0.0141 (0.0129, 0.0154)	34512
West Cheshire	0.0143 (0.0120, 0.0167)	8914
County Hosp Here	0.0168 (0.0125, 0.0225)	2621
Bronglais	0.0150 (0.0127, 0.0177)	8878
West Wales Gen	0.0127 (0.0111, 0.0144)	17511
Withybush	0.0158 (0.0142, 0.0176)	21322
Nevill Hall	0.0059 (0.0051, 0.0069)	26955
Royal Gwent	0.0072 (0.0066, 0.0079)	62274
Ysbyty Gwynedd	0.0117 (0.0106, 0.0130)	30929
Princess of Wales	0.0150 (0.0137, 0.0165)	27981
Caerphilly Miners	0.0067 (0.0056, 0.0080)	19089
East Glamorgan	0.0152 (0.0139, 0.0166)	33550
Prince Charles	0.0073 (0.0063, 0.0084)	26765
Mid Wales	0.0149 (0.0121, 0.0182)	6260
Uni Hosp Wales	0.0149 (0.0138, 0.0161)	39321
Llandough	0.0111 (0.0101, 0.0121)	42541
Singleton	0.0116 (0.0107, 0.0125)	56578
Neath	0.0094 (0.0082, 0.0107)	23622

After the investigation of the above risk factors, the next step is to adjust the relative risk contours for congenital malformations around the Nant y Gwyddon landfill site. Any resultant patterns of excess risk in the adjusted relative risk contours will therefore be the result of other risk factors that have not yet been considered. This adjustment applied to the relative risk contours is discussed below.

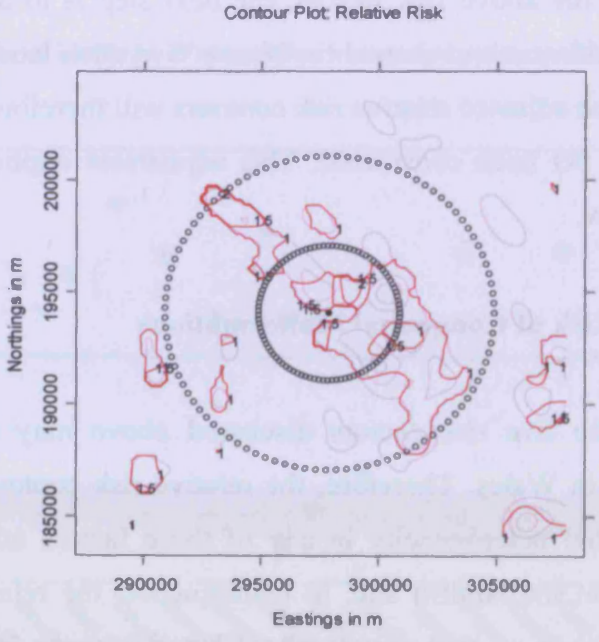
4.1.3 Adjusted Relative Risk of Congenital Malformations

There is evidence that the five risk factors discussed above may influence the risk of congenital malformations in Wales. Therefore, the relative risk contours seen in Figure 4.1 may be the result of spatial heterogeneity in any of these factors and not of exposure to contaminants released from the landfill site. In consequence, the relative risk contours are adjusted for these five factors so that any resultant heterogeneous spatial patterns are the outcome of other risk factors that have not yet been considered.

In order to adjust the relative risk contours for confounding factors, a logistic regression model is required. In the logistic regression model the response variable is whether a birth has a congenital anomaly and the explanatory variables are the five risk factors discussed above.

The inclusion of all five risk factors in the logistic regression model was statistically significant and thus they were all adjusted for, see Appendix 4. This logistic model is fitted using births that reside between 3km and 7km from the centre of any of the 32 landfill sites identified in Wales by the Environment Agency. Those within 3km of any of the landfill sites were excluded from the fitting of the model, because these mothers could have additional risk attributable to residing in close proximity to a landfill site. Mothers living further than 7km from the centre of the site could be exposed to another landfill site or may not be representative of a control region so they too are omitted from the model fit. The resultant logistic model is then used to generate predicted probabilities for all births in the 20km square region centred over Nant y Gwyddon and these predicted probabilities are used as weights in the generation of the expected contours using all births. The adjusted relative risk surface is then derived by dividing the observed by the expected density of congenital anomalies as discussed in Chapter 3.

Figure 4.7: relative risk contours of congenital malformations with adjustment for confounding in the 20km square region centred over Nant y Gwyddon 1983-1997



The relative risk contours for all congenital anomalies with adjustment for confounding are presented in Figure 4.7. There are two regions in the 20km square region centred over Nant y Gwyddon for which the relative risk is elevated over 2 even after adjustment for confounding. These elevations are either random or have occurred because there are other risk factors for

congenital malformations that have not been included in the logistic regression model such as proximity to the landfill site.

The randomness of the elevations can be assessed by testing for homogeneity of the relative risk surface using the global significance test. The hypotheses are:

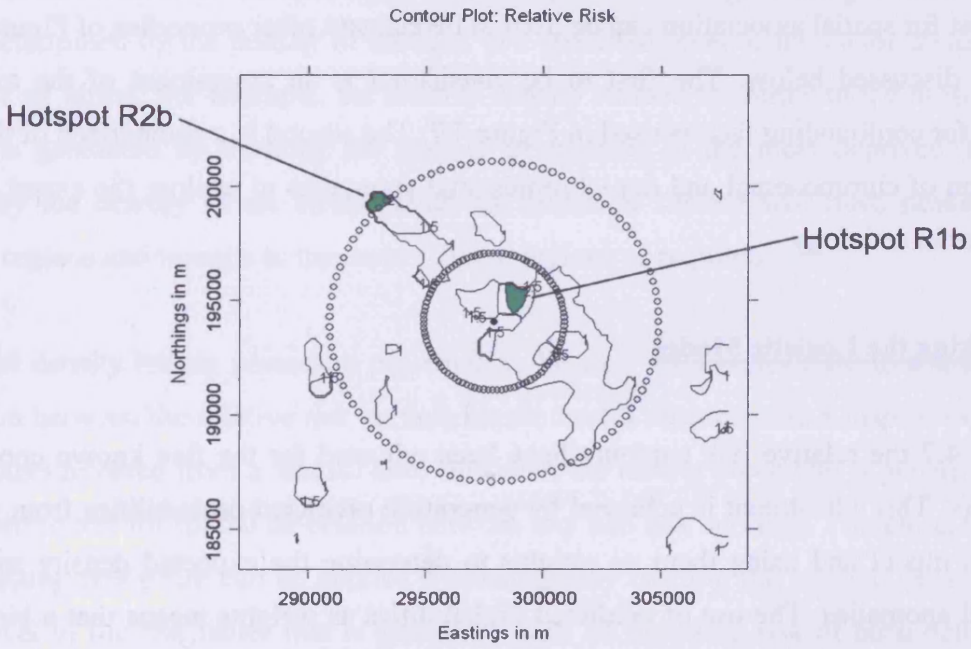
H_0: the spatial patterns in the relative risk contours for congenital malformations are homogeneous after adjustment for known confounding factors

and

H_1: the spatial patterns in the relative risk contours for congenital anomalies are heterogeneous after adjustment for known confounding factors.

The hotspots tested in the global significance test are displayed in Figure 4.8. Adjustment for confounding has reduced the magnitude of the elevation seen for hotspots R3a and R4a in Figure 4.2 to such an extent that they are no longer defined as hotspots. However, hotspots of excess birth defects still remain in the location of hotspots R1a and R2a in Figure 4.2.

Figure 4.8: hotspots of excess density of congenital malformations after adjustment for confounding in the 20km square region centred over Nant y Gwyddon 1983-1997



The results of the global significance test on the two hotspots of excess density of congenital anomalies after adjustment for the five confounding factors are given in Table 4.7. The results

indicate that both hotspots are statistically significant which provides evidence for the alternative hypothesis suggesting that the spatial pattern in the adjusted relative risk surface is heterogeneous. Therefore, in conclusion there could be other risk factors that have not yet been considered which influence the spatial distribution of congenital malformations. One of these risk factors could be exposure to contaminants released from the Nant y Gwyddon landfill site.

Table 4.7: results of the global significance test on hotspots of excess density of congenital anomalies after adjustment for confounding in the 20km square region centred over Nant y Gwyddon 1983-1997

Hotspot		R1b	R2b
Number of births		359	310
Number of congenital malformations		16	11
Observed mean adjusted relative risk		2.222	2.100
p-value	test 1	<0.001	0.001
	test 2	0.001	0.001
	test 3	<0.001	0.006

The next step in this investigation would be to use the kernel density test for spatial association between the adjusted relative risk contours in Figure 4.7 and mapped exposure to contaminants released from the Nant y Gwyddon landfill site. Unfortunately, such mapped exposure was not made available therefore this cannot be presented here. However, the kernel density test for spatial association can be used to investigate other properties of Figure 4.7 and these are discussed below. The first to be considered is an assessment of the method of adjusting for confounding factors used in Figure 4.7. The second is a comparison of the spatial distribution of chromosomal and non-chromosomal anomalies to explore the extent to which these differ in aetiology.

4.2 Checking the Logistic Model

In Figure 4.7 the relative risk contours have been adjusted for the five known confounding risk factors. This adjustment is achieved by generating predicted probabilities from a logistic regression model and using them as weights to determine the expected density surface for congenital anomalies. The use of predicted probabilities as weights means that a birth that is more likely to have a congenital anomaly based on information on the five risk factors contributes more to the expected density surface because at the location of that birth the density of congenital anomalies is expected to be higher. The observed density is then divided

by the expected density generated with the weights to derive the relative risk surface. However, it is unlikely that the logistic regression model explains the relationship between the health outcome and the risk factors perfectly and so the quality of the adjustment process was investigated.

If the adjustment of the relative risk for all birth defects for a given factor is successful then the resultant relative risk contours should not be spatially associated with the risk factor. Therefore, the success of the adjustment for a risk factor can be assessed by testing for spatial association between the health outcome with adjustment for confounding and the risk factor.

The hypotheses are

H_0: the adjusted relative risk contours for all congenital anomalies and a specified risk factor that has been included in the adjustment are not spatially associated

and

H_1: the adjusted relative risk contours for all congenital anomalies and a specified risk factor that has been included in the adjustment are spatially associated.

The spatial distribution of the risk factor can be represented by generating a relative density surface determined by the density of mothers in a specified level of the factor divided by the density of all births. For example, the relative density surface for births to the most deprived mothers is generated by dividing the density of mothers in the most deprived Townsend quintile by the density of all births. Thus, the resultant surface will have peaks in more deprived regions and troughs in the most affluent regions as required.

The kernel density testing procedure presented in Section 3.3.3 is designed to test for spatial association between the relative risk surface for the health outcome and a map of exposure to contaminants released from a landfill site. However, the testing procedure is generic in that it can be used to test for spatial association between any two risk surfaces. Therefore, the kernel density testing procedure can be applied by analogously treating the relative density surface for the level of the risk factor that is associated with an increased risk of birth defects as an 'exposure' surface. Therefore, the scatterplot that forms the basis of the kernel density test is plotted between the observed relative density of the risk factor of interest for each individual derived from the relative density surface (x-axis) and the corresponding relative risk of

congenital malformations with adjustment for confounding (y-axis). Here, the definition of those in the background region are those individuals that do not have an excess risk of births in the specified level of the risk factor of interest i.e. those with a relative density less than or equal to 1.

In the kernel density test for spatial association, the three test statistics are as follows:

1. the mean relative risk in the background region;
2. the proportion of relative risks above 1 in the exposed region; and
3. the gradient of the linear regression line fitted to the scatterplot in the exposed region.

The kernel density test, as discussed in Section 3.3.3, is performed by simulating health outcome events under the null and alternative hypotheses, calculating the three test statistics for each simulation and then discriminant analysis is used to determine which of the hypotheses are most likely to be correct. The kernel density estimator for the joint distribution of the test statistics is produced under the null and the alternative hypotheses separately using the values of the test statistics derived in the simulations. These two kernel density estimators under the null and the alternative hypotheses can be used to calculate the 'height ratio' which for the column vector of test statistics $\mathbf{x}$ is defined as

$$\frac{\text{the height of the kernel density for } \mathbf{x} \text{ under the alternative hypothesis}}{\text{the height of the kernel density for } \mathbf{x} \text{ under the null hypothesis}}.$$

The 'height ratio' is calculated for all simulations under the null hypothesis and also for the observed situation.

The p-value for the test is given by the proportion of simulations under the null hypothesis for which the 'height ratio' is greater than the 'height ratio' for the observed situation. In order to calculate the 'height ratios' simulations must be performed under the alternative as well as under the null hypothesis. The probability of birth i having a birth defect, p_i, is equal to the predicted probability from the logistic regression model for congenital anomalies, denoted α_i, under the null hypothesis. In contrast, under the alternative hypothesis p_i is equal to α_i multiplied by the height of the relative density surface for the risk factor, denoted ξ_i.

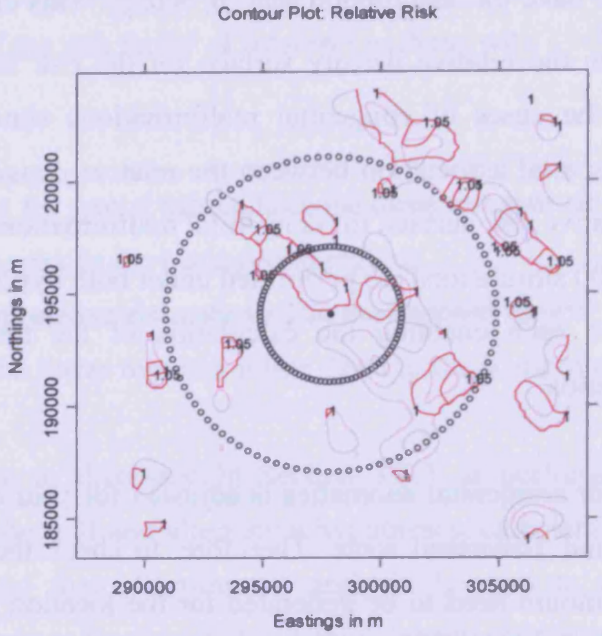
Therefore the simulations under the null are performed under the assumption that $p_i = \alpha_i$ and those under the alternative have the assumption that $p_i = \alpha_i \xi_i$. This ensures that there is no spatial association between the relative density surface for the risk factor and the adjusted relative risk surface for the cases of congenital malformations simulated under the null hypothesis. Additionally, spatial association between the relative density surface for the risk factor and the adjusted relative risk surface for congenital malformations under the alternative hypothesis is enforced. 1000 simulations are performed under both hypotheses and the relative risk surfaces are fitted to each enabling the calculation of the three test statistics and subsequently the 'height ratios'.

The relative risk surface for congenital anomalies is adjusted for year of birth, maternal age, gender, hospital of birth and Townsend score. Therefore, to check the adjustment for these factors relative density contours need to be generated for the location of births that fall into the levels of each of these factors that are associated with an increased risk of congenital malformations. The kernel density testing procedure can then be applied to these to assess the spatial association. If there is no spatial association then the adjustment is sufficient and if there is spatial association then the adjustment is not as good as it should be.

4.2.1 Gender

Table 4.3 indicates that male births are at a greater risk of having a congenital anomaly. Therefore, if the adjustment for gender was not adequate then there would still be elevated relative risk of birth defects in regions where there is an excess density of male births. The relative density contours for male births are given in Figure 4.9. The contours are drawn where the relative density surface has a height of 1 and 1.05 which are of a much lower magnitude to the values at which contours have been drawn in previous figures. The reason for this is the comparatively flat relative density surface for gender.

Figure 4.9: relative density contours for male births in the 20km square region centred over Nant y Gwyddon 1983-1997



The scatterplot between the relative risk of congenital anomalies with adjustment for confounding and the relative density of male births is given in Figure 4.10. The majority of the relative density values are close to 1 which is expected from the fairly flat relative density surface represented in Figure 4.9. The observed values for the three test statistics are determined from Figure 4.10. The mean relative risk value in the background region would be close to 1, the slope in the exposed region around 0 and the proportion of relative risk values above 1 in the exposed region would be about 0.5 under the null hypothesis. On the other hand, under the alternative hypothesis it is expected that the background mean would be lower than 1, the slope in the exposed region would be positive and the proportion of relative risk values above 1 in the exposed region would be higher than 0.5.

In Figure 4.10 the proportion of relative risks greater than 1 in the exposed region is 0.470, the slope in the exposed region is -0.830 and the mean relative risk in the background region is 1.010. The kernel density test for spatial association between the relative density for excess males and the adjusted relative risk contours of congenital anomalies gives a p-value of 0.680, a clear indication supporting the null hypothesis. Therefore, the adjustment for gender in the relative risk contours for congenital anomalies is adequate.

Figure 4.10: the scatterplot between the relative density of male births and the adjusted relative risk of congenital anomalies

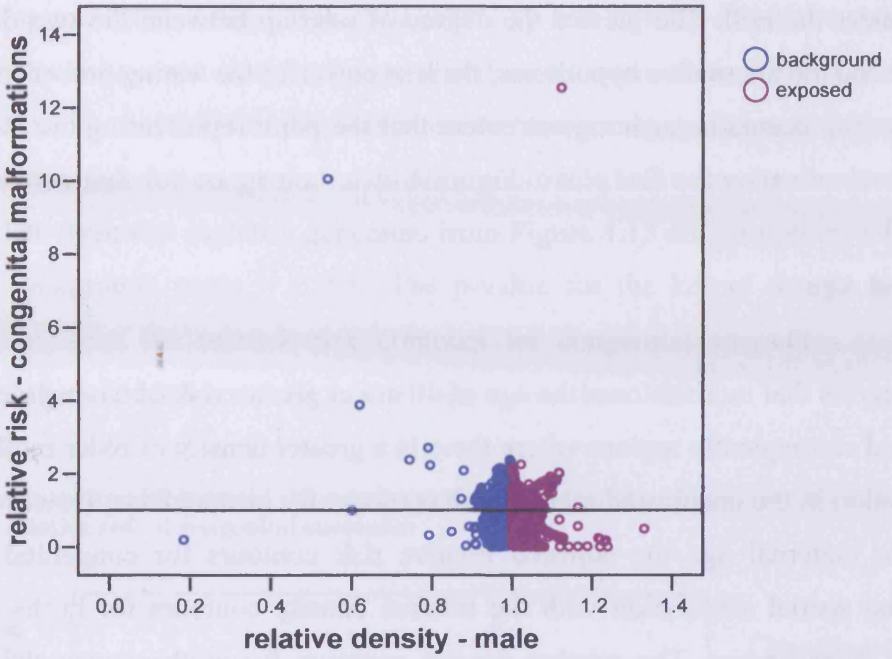


Figure 4.11: the distribution of the three test statistics for spatial association between the adjusted relative risk of congenital malformations and the relative density of male births under the null and alternative hypotheses

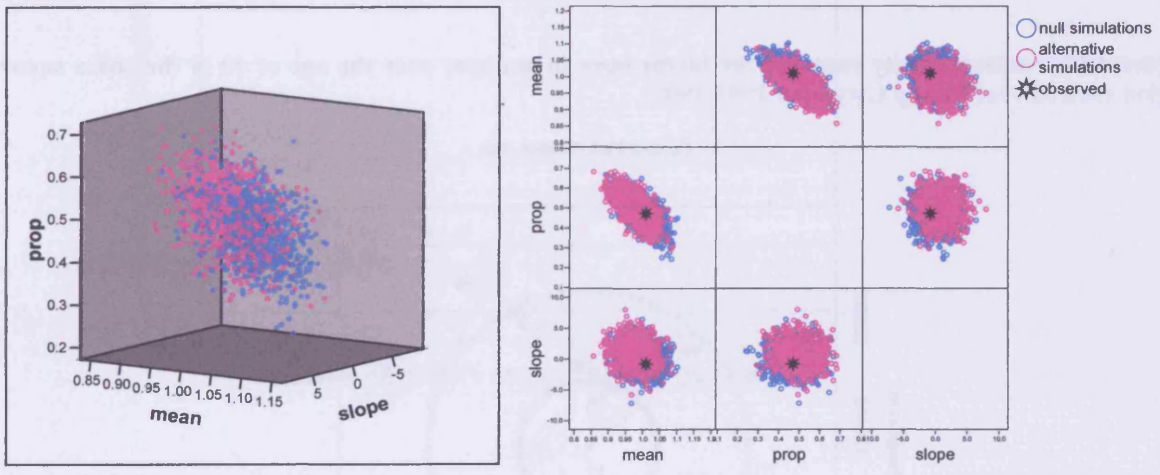


Figure 4.11 presents the distribution of the three test statistics under the null and alternative hypotheses and also the observed values of the three test statistics. The location of the point representing the observed values for the test statistics relative to the distribution of the test statistics under the null and alternative hypotheses indicates which hypothesis is more likely to be correct. In Figure 4.11 the two distributions of the test statistics under the two different hypotheses overlap to a high degree. The overlap of points under the null and alternative

hypotheses occurs because of the comparatively flat relative density surface for male births which meant that simulations under the alternative did not differ hugely from those simulations under the null. The greater the degree of overlap between the two distributions under the null and the alternative hypotheses, the less powerful the testing procedure becomes. In fact, the overlap occurs to such a great extent that the point representing the observed test statistics cannot be seen on the first plot in Figure 4.11.

4.2.2 Maternal Age

Figure 4.4 suggests that mothers over the age of 40 are at greater risk of having a child with a birth defect and consequently regions where there is a greater density of older mothers should lead to a elevation in the unadjusted relative risk contours for birth defects. However, with the adjustment for maternal age the adjusted relative risk contours for congenital anomalies should have no spatial association with the relative density contours for births to mothers older than 40 years of age. The relative density contours for mothers over the age of 40 around Nant y Gwyddon are given in Figure 4.12. The contours are drawn where the relative density is equal to 1, 1.5 and 2 indicating that there are regions with a higher density of older mothers.

Figure 4.12: relative density contours for births born to mothers over the age of 40 in the 20km square region centred over Nant y Gwyddon 1983-1997

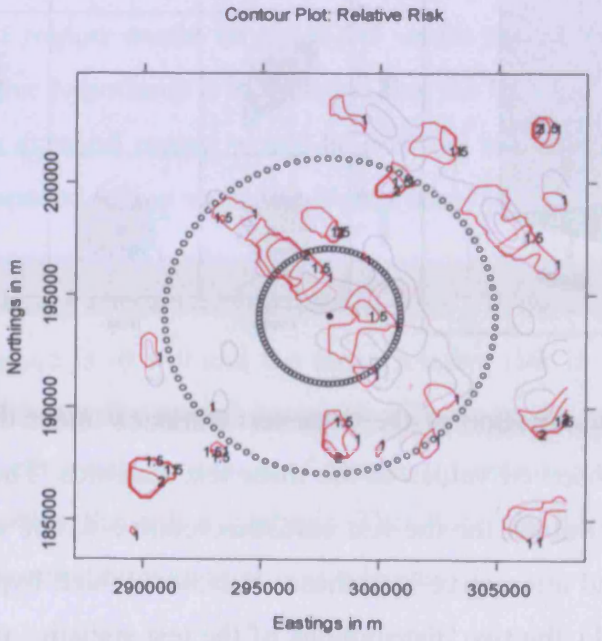
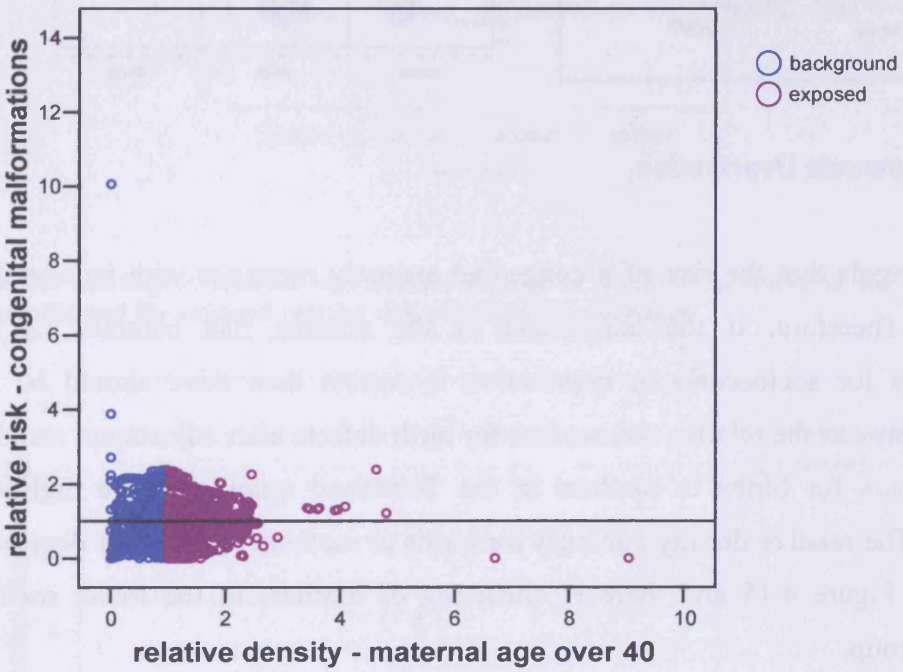


Figure 4.13 indicates that the relative density surface for mothers over 40 has a higher variation than when considering gender. Some of the relative density values are close to 0 corresponding to regions where there is a relatively low density of older mothers and some relative density values are close to 2 corresponding to areas where there is a relatively high density of older mothers. There does not seem to be a strong spatial association between the adjusted relative risk for congenital anomalies and the relative density for mothers over the age of 40. The three test statistics generated from Figure 4.13 are proportion = 0.520, slope = 0.230 and background mean = 0.950. The p-value for the kernel density test for spatial association between the relative risk contours for congenital anomalies and the relative density of births to mothers over 40 years of age is less than 0.001.

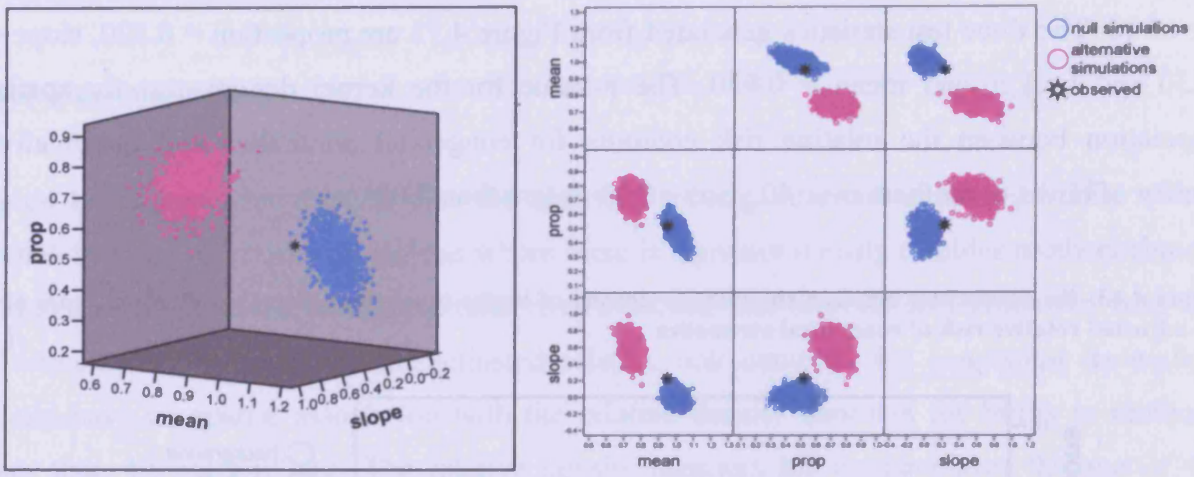
Figure 4.13: the scatterplot between the relative density of births to mothers over the age of 40 years and the adjusted relative risk of congenital anomalies



Plotting the test statistics under the two hypotheses in Figure 4.14 reveals that there is a greater degree of separation between the distributions of the test statistics under the two hypotheses. The reason for the increased degree of separation is that the relative density surface for mothers over the age of 40 is not as flat as that of male births and consequently the simulations under the alternative hypothesis deviate from those under the null hypothesis to a greater degree. The plotted point for the observed test statistics is on the periphery of those points derived under the null hypothesis explaining the low p-value. Therefore, the result

would be deemed significant however it is clear that of the two hypotheses it is more supportive of the null. Hence, the adjustment seems to have worked although there is room for improvement.

Figure 4.14: the distribution of the three test statistics for spatial association between the adjusted relative risk of congenital malformations and the relative density of births to mothers over the age of 40 years under the null and alternative hypotheses



4.2.3 Socioeconomic Deprivation

Figure 4.5 reveals that the risk of a congenital anomaly increases with increased levels of deprivation. Therefore, if the adjustment of the relative risk contours of congenital malformations for socioeconomic deprivation is correct then there should be no spatial association between the relative risk surface for birth defects after adjustment and the relative density contours for births to mothers in the Townsend quintile of the highest level of deprivation. The relative density contours for births to mothers in the most deprived quintile are given in Figure 4.15 and there is clustering of mothers in the lower socioeconomic deprivation group.

The scatterplot between the adjusted relative risk of birth defects and the relative density of high deprivation presented in Figure 4.16 seems to indicate that there is a lack of spatial association. It can also be seen from Figure 4.16 that the relative density for high deprivation can be as high as 4 indicating that there is clustering in the socioeconomic deprivation of mothers as is expected with more deprived and more affluent geographical areas. The three test statistics are proportion = 0.630, slope = -0.069 and background mean = 0.910. The p-value for testing for the spatial association between this relative density surface and the

adjusted relative risk surface for congenital malformations is less than 0.001 as was the case for maternal age.

Figure 4.15: relative density contours for births to mothers in the most deprived Townsend quintile in the 20km square region centred over Nant y Gwyddon 1983-1997

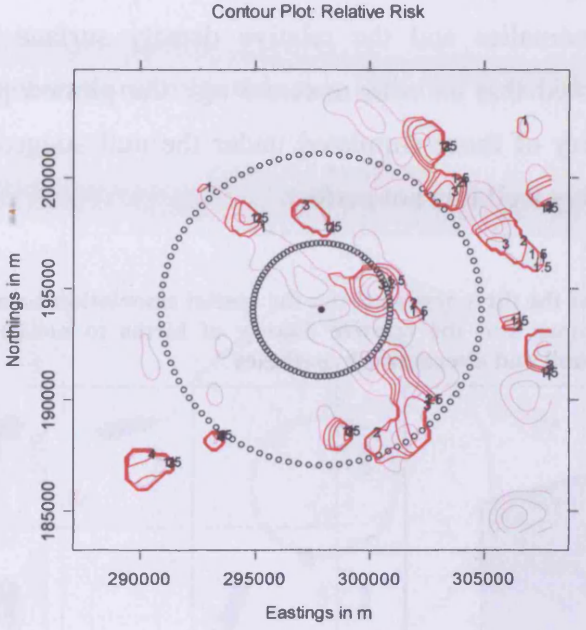
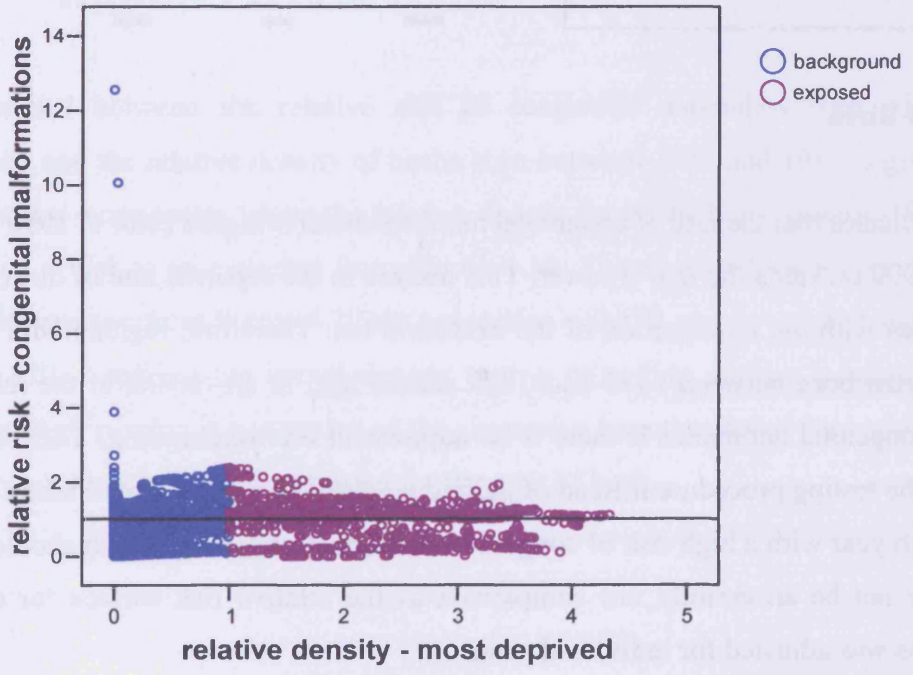
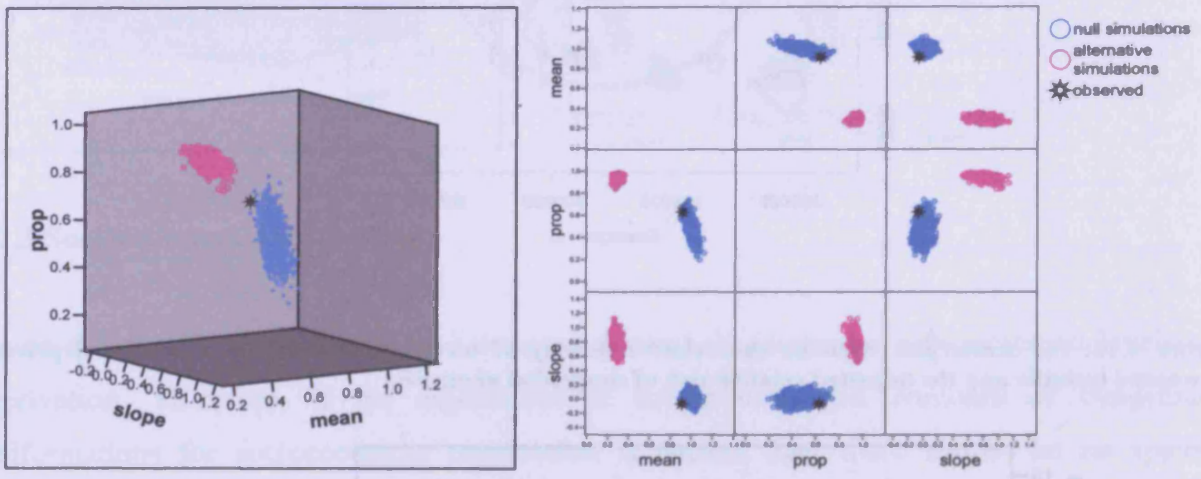


Figure 4.16: the scatterplot between the relative density of births to mothers in the most deprived Townsend quintile and the adjusted relative risk of congenital anomalies



The low p-value is supportive of the alternative hypothesis of spatial association. However, the plotted point for the observed test statistics is more consistent with those from simulations under the null hypothesis, which can be seen in Figure 4.17. Therefore, the result is more supportive of the null hypothesis of no spatial association between the adjusted relative risk surface for congenital anomalies and the relative density surface for high deprivation. However, it should be noted that as with maternal age the plotted point for the observed values lies on the periphery of those simulated under the null suggesting that although the adjustment works reasonably well it is not perfect.

Figure 4.17: the distribution of the three test statistics for spatial association between the adjusted relative risk of congenital malformations and the relative density of births to mothers in the most deprived Townsend quintile under the null and alternative hypotheses

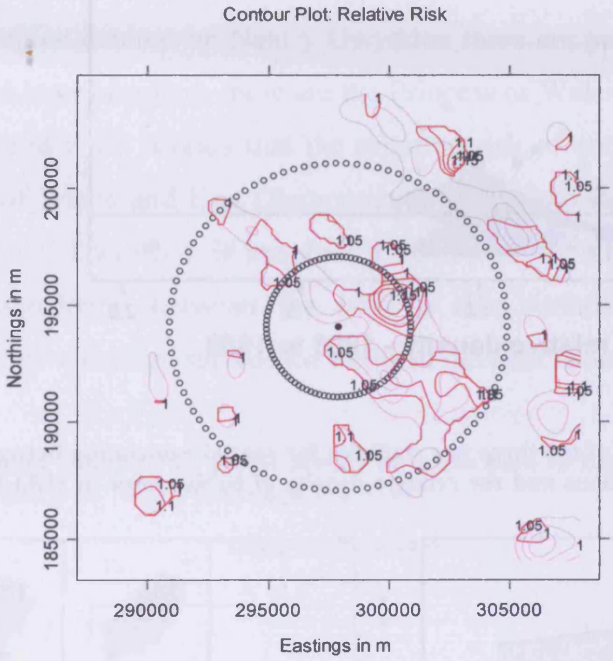


4.2.4 Year of Birth

Figure 4.3 indicates that the rate of congenital malformations is higher prior to 1990 and in the years from 1990 onwards the rate is lower. This decline in the reported rate of birth defects in 1990 coincides with the introduction of the exclusion list. Therefore, regions with an excess density of births born between 1983 and 1989 should lead to elevations in the relative risk surface for congenital anomalies if there is no adjustment for confounding. This time period was used in the testing procedure instead of individual years because it would take a long time to look at each year with a high risk of congenital malformations. However, it should be noted that this may not be an entirely fair comparison as the relative risk surface for congenital malformations was adjusted for individual years.

Figure 4.18 displays the relative density contours for excess density of births born in the years 1983 to 1989. The contours are drawn where the relative density surface has a height of 1, 1.05, 1.1, 1.15 and 1.2 reflecting the comparatively flat relative density surfaces that are produced when considering the spatial location of births born in a specific time period because there is no reason to believe that this would have a spatial pattern.

Figure 4.18: relative density contours for births born in 1983-1989 in the 20km square region centred over Nant y Gwyddon



The scatterplot between the relative risk of congenital anomalies with adjustment for confounding and the relative density of births born between 1983 and 1989 is given in Figure 4.19. There are more points where the relative density for births born between 1983 and 1989 is close to 1 which is as expected from Figure 4.18. The observed values for the three test statistics determined from Figure 4.19 are proportion = 0.670, slope = -0.500 and background mean = 0.820. Applying the kernel density testing procedure gives a p-value of less than 0.001. The small p-value is supportive of the alternative hypothesis that the logistic regression model may not have adequately adjusted for the year of birth.

Figure 4.19: the scatterplot between the relative density of births born in 1983-1989 and the adjusted relative risk of congenital anomalies

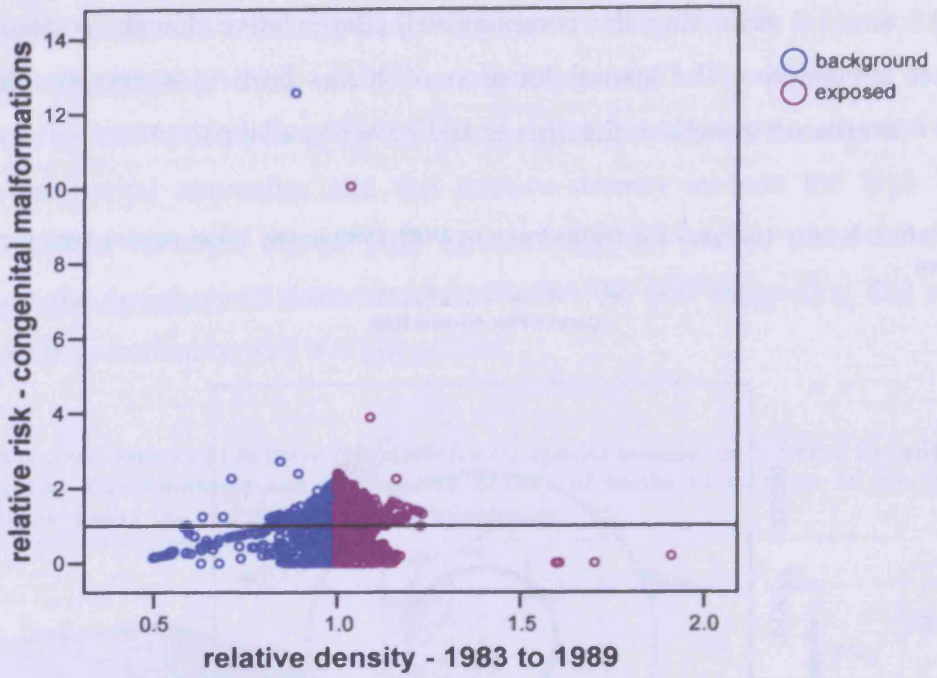
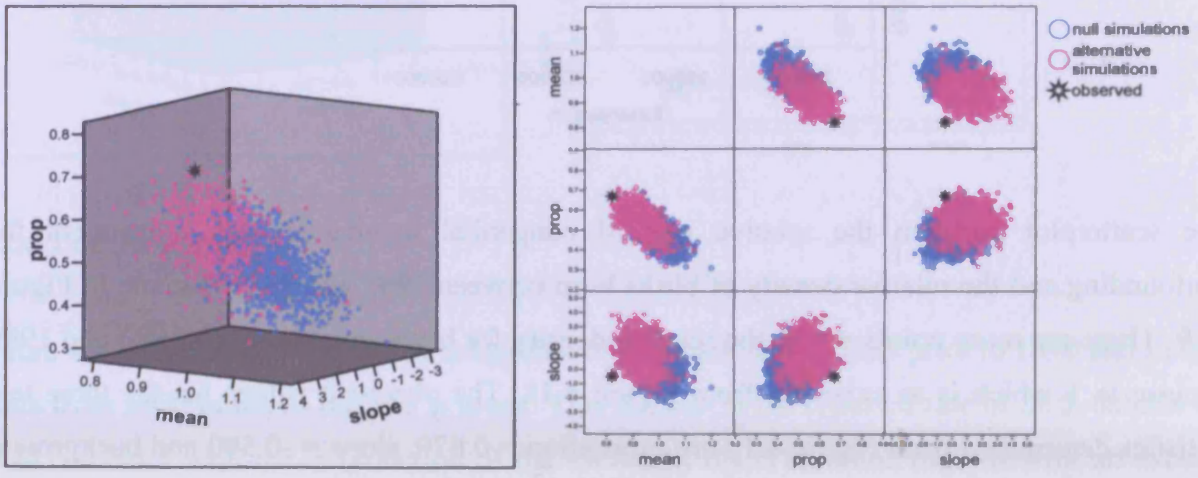


Figure 4.20: the distribution of the three test statistics for spatial association between the adjusted relative risk of congenital malformations and the relative density of babies born in 1983-1989 under the null and alternative hypotheses



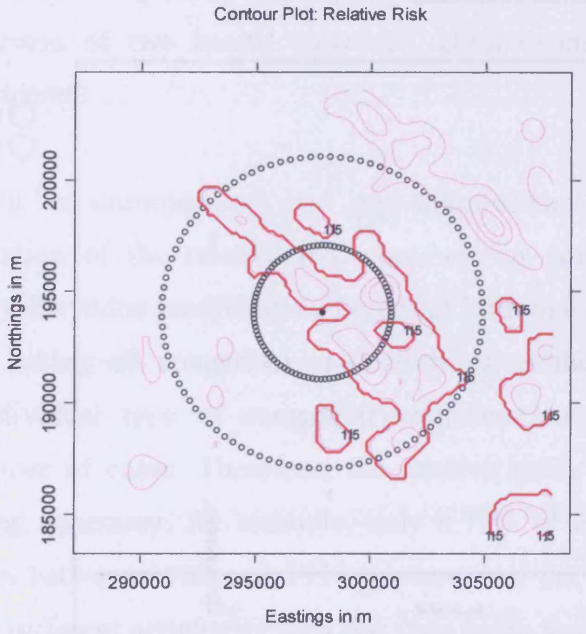
The small p-value is partially explained by Figure 4.20 which presents the distribution of the three test statistics under the null and alternative hypotheses and also the observed values of the three test statistics. The position of the plotted point for the observed test statistics is marginally closer to those for simulations under the alternative than those under the null. However, the two distributions of the test statistics under the two different hypotheses substantially overlap. The overlap of points under the null and alternative hypotheses occurs

because of the comparatively flat relative density surface for births born in 1983 to 1989 and results in a less powerful testing procedure. The overlap means that the observed point is not totally implausible under the null hypothesis. Moreover, the logistic regression model adjusts for year of birth by each year separately and not as a dichotomised variable used here so this may not be an absolutely fair comparison.

4.2.5 Hospital of Birth

In the 20km square region centred on Nant y Gwyddon there are only 4 possible hospitals at which the births could have occurred, these are the Princess of Wales, East Glamorgan, Prince Charles and Neath. Figure 4.6 reveals that the reported risk of congenital malformations is high in the Princess of Wales and East Glamorgan hospitals and that it is low in Neath and Prince Charles hospitals. Therefore, if hospital of birth was not adjusted for correctly there should be spatial association between the relative risk surface for birth defects with adjustment and the relative density surface for births in the East Glamorgan hospital, say.

Figure 4.21: relative density contours for excess density of births born in East Glamorgan hospital in the 20km square region centred over Nant y Gwyddon 1983-1997

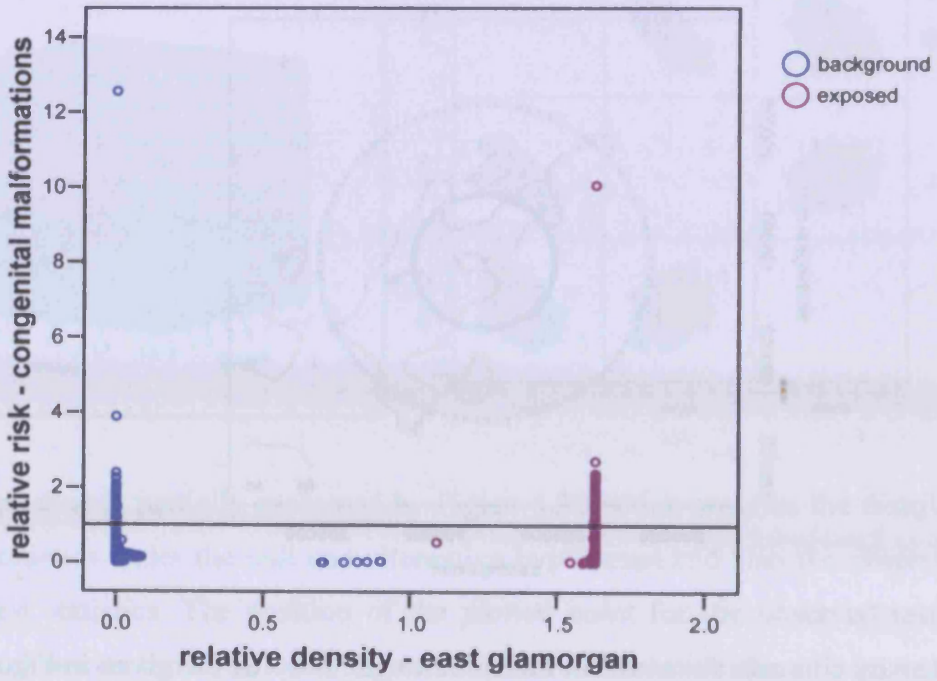


The relative density contours for births in East Glamorgan hospital are given in Figure 4.21. If the adjustment for hospital of birth was successful then there should be no spatial association

between this set of relative density contours and the adjusted relative risk contours for birth defects.

The scatterplot between the adjusted relative risk for congenital malformations and the relative density for births born in East Glamorgan hospital is given in Figure 4.22. The majority of relative density values for births born in East Glamorgan hospital are either close to 0 where there are very few or no births in East Glamorgan hospital or they are close to about 1.63 where the majority of births are born in East Glamorgan hospital. There are only a few points for which the relative density lies between 0 and 1.63. Any point in the exposed region that is not close to 1.63 has a huge influence on the slope and the relatively small number of such points means that the slope is heavily reliant on only a few relative density values. Therefore, the slope is unstable and thus the testing procedure cannot be performed for hospital of birth. Fortunately, the majority of births in the study region were born in one hospital and therefore the adequacy of the adjustment for hospital of birth is relatively unimportant when plotting the relative risk contours for congenital malformations around the Nant y Gwyddon landfill site.

Figure 4.22: the scatterplot between the relative density of births in East Glamorgan hospital and the adjusted relative risk of congenital anomalies



In conclusion, the adjustment made to the relative risk contours for confounding is not perfect with some factors having a spatial association with the adjusted contours. However, it is likely that no method of adjusting the relative risk contours would be perfect, and since the methodology does adjust the relative risk contours in the right direction with the adjusted contours typically being much closer to having no association than having spatial association with the risk factor, this method of using predicted probabilities from a logistic regression model to adjust the relative risk contours is thought to be reasonable and is utilised in the remainder of this thesis.

4.3 The Spatial Comparison of Two Health Outcomes

The method of comparing two sets of spatial contours can be adopted in the comparison of the spatial distributions of two health outcomes. There are a large number of different types of congenital anomalies, many of which have completely different aetiologies. One of the major disparities between certain birth defects is that some are the result of errors that occur in the chromosomes and the remainder are a result of the unsuccessful development of part of the foetus. The former are referred to as chromosomal anomalies and the latter are non-chromosomal anomalies. Consequently, in order to illustrate the use of the testing procedure in the spatial comparison of two health outcomes, chromosomal and non-chromosomal abnormalities are considered.

The spatial comparison of chromosomal and non-chromosomal anomalies is of interest because in the generation of the relative risk contours for congenital malformations in proximity to the Nant y Gwyddon landfill site all types of birth defects were grouped together. The main reason for taking all congenital malformations as the outcome as opposed to focusing on each individual type of congenital malformation separately was to get a sufficiently large number of cases. Therefore, the relative rarity of certain types of birth defects makes grouping necessary, for example, only 6.77% of congenital anomalies were chromosomal in Wales between 1983 and 1997. However, if the congenital malformations grouped together have different aetiologies then this may mean that the impact of certain risk factors could be missed. In the dataset from the CHS and NCAS chromosomal anomalies are flagged and thus in this section chromosomal and non-chromosomal anomalies are considered separately to assess the extent to which congenital malformations can differ aetiologically even though the numbers are small.

4.3.1 Possible Risk Factors for Chromosomal and Non-Chromosomal Anomalies

There were five variables considered for congenital malformations in Section 4.1.2. The relationships between these risk factors and the probability of each of the two types of congenital anomalies are summarised in this subsection and the corresponding tables and plots are provided in Appendix 4. If any of these relationships are different for chromosomal and non-chromosomal anomalies then this emphasises the problem of grouping congenital anomalies together in analyses. These relationships are analysed using all births in Wales between 1983 and 1997. The relationships between each of the risk factors and non-chromosomal anomalies are similar to those for all birth defects because the relatively large numbers of non-chromosomal defects means that they dominate the patterns seen for all congenital malformations. Therefore, the emphasis lies in the relationships between chromosomal anomalies and each of the risk factors.

Year of birth

The temporal changes seen in the rate of reporting for all congenital malformations is not observed for chromosomal anomalies. The reasons for the relatively stable annual rates are that no chromosomal anomalies appear on the exclusion list and the majority of chromosomal anomalies are easy to diagnose and are likely to be diagnosed within the first 10 days of birth. Therefore, the changes in reporting procedures that were made would not affect chromosomal anomalies. This fairly constant rate of chromosomal anomalies through time is in contrast to claims by Rankin et al. and Dolk et al. that there is a statistically significant increase in the rate of chromosomal anomalies through time [36, 279].

The proportion of non-chromosomal anomalies varies temporally in a similar way to that of all congenital malformations because it is the changes in reporting of non-chromosomal anomalies that are responsible for the patterns seen in that of all birth defects.

Gender

The proportion of chromosomal anomalies is only slightly higher amongst males than it is for females and the difference is not statistically significant. Consequently, it is clear that the

disparity between the rates of birth defects for males and females are largely a result of the different risks of non-chromosomal anomalies between the sexes which is similar to that of all birth defects.

Maternal age

The proportion of chromosomal anomalies increases with increasing maternal age and the rate of chromosomal anomalies is over 4 times greater for mothers over 40 than for mothers between 31 and 40 years, a statistically significant difference. The positive relationship between the risk of chromosomal anomalies and maternal age is consistent with the knowledge that Down's syndrome is more probable amongst women over 35 years old [36].

In contrast to chromosomal anomalies, the risk of non-chromosomal abnormalities falls with maternal age with mothers over 40 being at a lower risk. However, this decreasing trend is not statistically significant. There is independent evidence that the risk of gastroschisis increases with decreasing maternal age and thus the maternal age effect for gastroschisis may account for the increased risk of non-chromosomal anomalies amongst younger mothers [23]. The inclusion of all other types of non-chromosomal anomalies in the analysis and the relative rarity of gastroschisis could explain the lack of statistical significance.

The increased risk for all birth defects combined by maternal age despite the decreasing risk of non-chromosomal anomalies with increased maternal age highlights the enormity of the maternal age effect for chromosomal anomalies, particularly with their relative rarity. The opposing influence of maternal age on chromosomal anomalies compared to non-chromosomal abnormalities indicates the extent to which types of defects can differ in terms of aetiology. In addition, it is clear that by combining all birth defects in the analysis the true relationship between non-chromosomal anomalies and maternal age was not detected.

Socioeconomic deprivation

The risk of chromosomal anomalies increases with affluence which is likely to be the result of more affluent women tending to have children at an older age and consequently the pattern seen here is the result of the maternal age effect. The relationship between socioeconomic deprivation and non-chromosomal anomalies is similar to that of all birth defects.

Hospital of birth

The reporting rate for chromosomal anomalies varies by hospital with Nevill Hall still reporting the lowest rate and Wrexham has the highest reporting rate. The results for non-chromosomal anomalies are similar to all congenital malformations combined.

The disparities in the influence of these risk factors on chromosomal and non-chromosomal defects indicate the extent to which the aetiologies can differ for certain birth defects. Consequently, the impact of certain risk factors can be missed if birth defects are grouped together. In the analysis of the risk of congenital malformations near the Nant y Gwyddon landfill site, all birth defects were grouped together and one logistic regression model was used to generate the weights for each of the births. However, it is clear from this analysis that the impact of risk factors vary for chromosomal and non-chromosomal anomalies so the logistic regression models will be different for each type of congenital malformation. Therefore, in the next section the adjusted relative risk contours are produced for chromosomal and non-chromosomal abnormalities separately where the logistic regression models are fitted for only that type of birth defect.

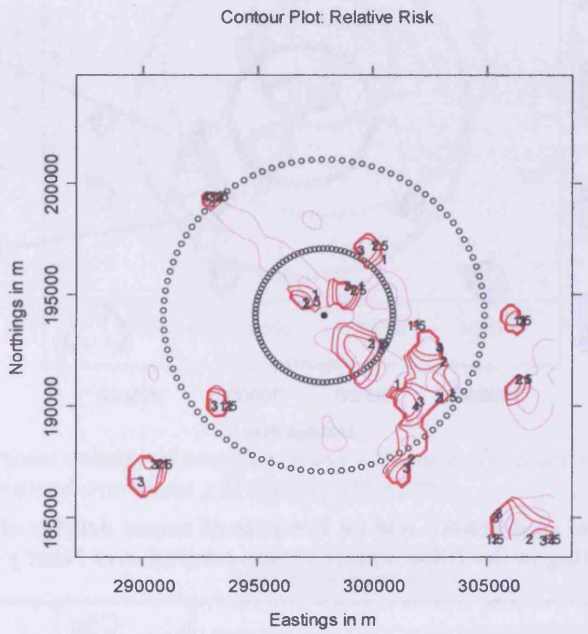
4.3.2 Adjusted Relative Risk of Chromosomal and Non-Chromosomal Anomalies

The logistic regression model for congenital malformations used to adjust the relative risk contours for all congenital malformations combined included year of birth, maternal age, gender, hospital of birth and Townsend quintile. Therefore these were the factors considered for chromosomal and non-chromosomal anomalies however not all these confounding factors are significant when considering only chromosomal defects or only non-chromosomal anomalies as indicated in the logistic regression results in Appendix 4. In fact, a forward stepwise approach results in a logistic regression model with maternal age and hospital of birth as the only statistically significant risk factors for chromosomal anomalies. The model for non-chromosomal abnormalities includes year of birth, gender and hospital of birth using the forward stepwise method.

A traditional statistical approach would result in fitting the logistic regression model with only the statistically significant risk factors and using the resultant predicted probabilities to adjust

the relative risk surfaces for known confounding. However, the lack of statistical significance is not the same as the absence of a causal link and it may be that the population size is not large enough to achieve statistical significance. Therefore, the relative risk contours for both chromosomal and non-chromosomal anomalies have been adjusted for all five confounding factors regardless of statistical significance. The corresponding results for when only statistically significant risk factors are adjusted for are presented in Appendix 4 and they do not differ greatly to those for when all five are used in the adjustment.

Figure 4.23: relative risk contours for chromosomal anomalies adjusted for confounding in the 20km square region centred over Nant y Gwyddon 1983-1997



The relative risk contours for chromosomal anomalies adjusted for the five confounding factors are presented in Figure 4.23. Visual comparison of these contours with those for all congenital malformations highlights some similarities but also some contrasts. The most obvious difference is higher relative risk values in the peaks in the surface for chromosomal anomalies compared to those of congenital malformations. The reason for this disparity is the relatively smaller number of chromosomal defects leading to a greater variance in the relative risk surface and consequently greater maximum relative risk values. The greater variance is the motivation for using a relative risk of 4 and above to define hotspots of chromosomal anomalies instead of the usual value of 2. The criterion that the hotspots must have a relative risk of chromosomal abnormalities that is greater than 4 ensures that 3.7% of the total

population of births lie within the boundaries of a hotspot which is comparable to the 3.9% of births that lie within hotspots of excess relative risk of congenital anomalies when 2 is used as the threshold.

Figure 4.24: hotspots of excess density of chromosomal anomalies after adjustment for confounding in the 20km square region centred over Nant y Gwyddon 1983-1997

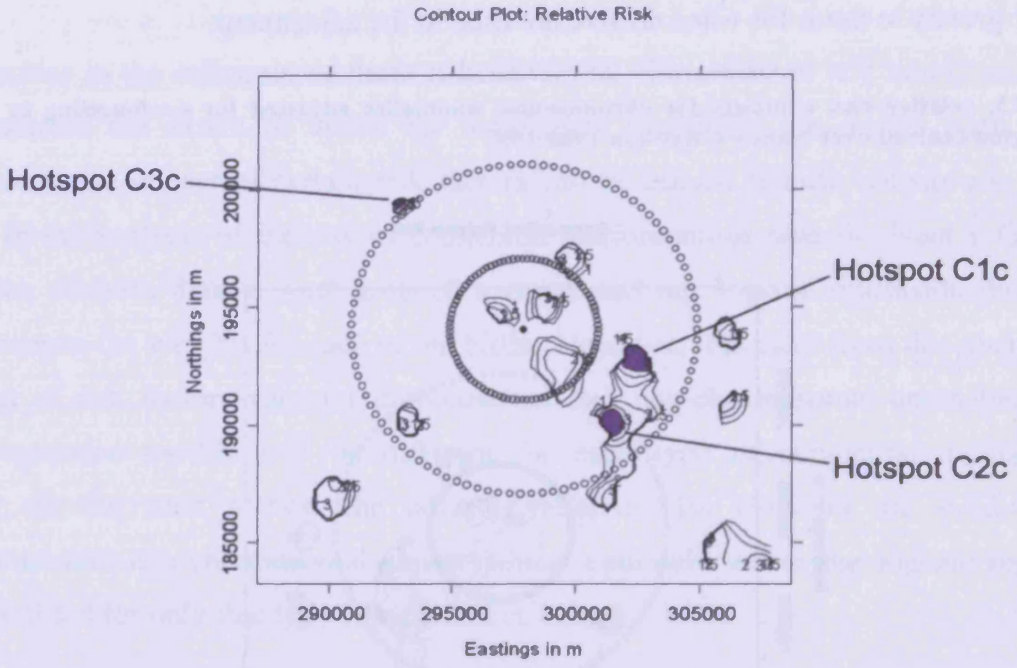


Table 4.8: results of the global significance test on hotspots of excess density of chromosomal anomalies after adjustment for confounding in the 20km square region centred over Nant y Gwyddon 1983-1997

Hotspot		C1c	C2c	C3c
Number of births		447	243	550
Number of chromosomal anomalies		2	2	2
Observed mean adjusted relative risk		5.947	5.010	8.853
p-value	test 1	0.004	0.012	0.002
	test 2	0.007	0.013	0.017
	test 3	0.053	0.014	0.019

There are three regions over which the adjusted relative risk for chromosomal anomalies exceeds 4 and these are highlighted in Figure 4.24. It is apparent from Table 4.8 that two of these hotspots are statistically significant, hotspots C2c and C3c. Moreover hotspot C1c is close to being statistically significant and if an overall p-value was calculated for all three tests then it is likely that it will be below 0.05. Therefore, there is evidence that the spatial

distribution of chromosomal anomalies is heterogeneous after adjustment for all five confounders.

Figure 4.25: relative risk contours for non-chromosomal anomalies adjusted for confounding in the 20km square region centred over Nant y Gwyddon 1983-1997

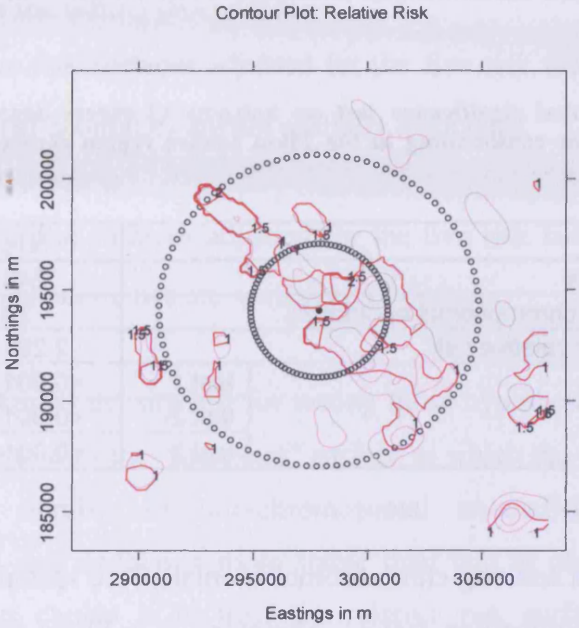
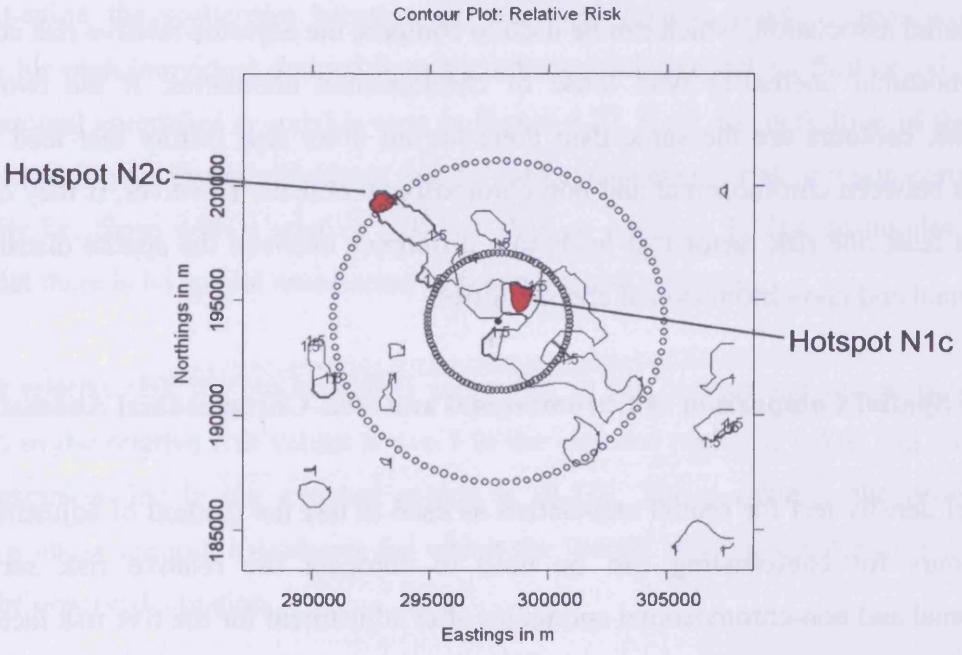


Figure 4.26: hotspots of excess density of non-chromosomal anomalies after adjustment for confounding in the 20km square region centred over Nant y Gwyddon 1983-1997



The relative risk contours for non-chromosomal anomalies after adjustment for confounding using all five confounding factors are shown in Figure 4.25. The results of the global significance test given in Table 4.9 indicate that the two hotspots identified in Figure 4.26, N1c and N2c, are statistically significant. Therefore, there is evidence that the spatial distribution of the risk of non-chromosomal anomalies is heterogeneous.

Table 4.9: results of the global significance test on hotspots of excess density of non-chromosomal anomalies after adjustment for confounding in the 20km square region centred over Nant y Gwyddon 1983-1997

Hotspot		N1c	N2c
Number of births		543	284
Number of non-chromosomal anomalies		20	13
Observed mean relative risk		2.291	2.215
p-value	test 1	<0.001	<0.001
	test 2	<0.001	<0.001
	test 3	<0.001	<0.001

Clearly, both chromosomal and non-chromosomal anomalies are spatially heterogeneous after adjustment for the five confounding factors. Therefore, it is likely that there are other risk factors that have not yet been considered for both types of congenital malformation. However, it is not clear as to whether or not these unknown risk factors are the same for both chromosomal and non-chromosomal anomalies. This is determined using the kernel density test for spatial association, which can be used to compare the adjusted relative risk contours of non-chromosomal anomalies with those of chromosomal anomalies. If the two adjusted relative risk contours are the same then there are no other risk factors that lead to spatial disparities between chromosomal and non-chromosomal defects. However, if they differ then there is at least one risk factor that leads to a difference between the spatial distribution for chromosomal and non-chromosomal abnormalities.

4.3.3 The Spatial Comparison of Chromosomal and Non-Chromosomal Anomalies

The kernel density test for spatial association as used to test the method of adjusting relative risk contours for confounding can be used to compare the relative risk surfaces for chromosomal and non-chromosomal anomalies after adjustment for the five risk factors. If the two relative risk surfaces are spatially associated then those five risk factors are the only ones

that lead to disparities in the relative risk surfaces for chromosomal and non-chromosomal defects, any other risk factors influence both types of birth defects in the same way. If the relative risk surfaces are not correlated then there are more risk factors that influence either only one of the types of birth defects or that affects them in different ways.

The null hypothesis for the testing procedure is

H_0: the relative risk surfaces adjusted for the five risk factors for chromosomal and non-chromosomal anomalies are not spatially associated

and the alternative hypothesis is

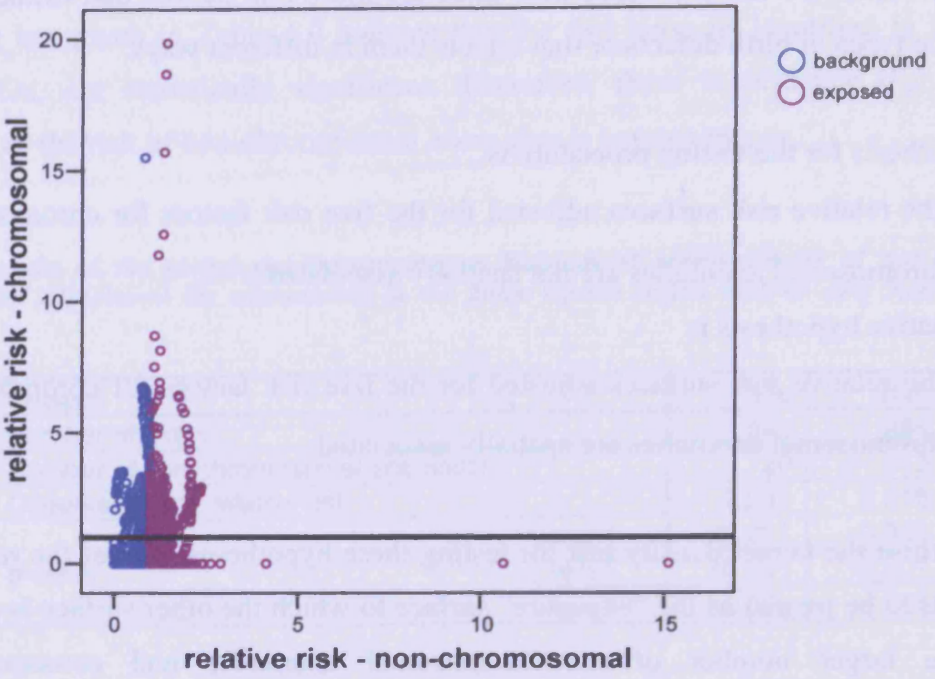
H_1: the relative risk surfaces adjusted for the five risk factors for chromosomal and non-chromosomal anomalies are spatially associated.

In order to utilise the kernel density test for testing these hypotheses, one of the relative risk surfaces needs to be treated as the "exposure" surface to which the other surface is compared. There are a larger number of non-chromosomal anomalies and consequently the corresponding relative risk surface is more stable than that of chromosomal abnormalities. Therefore, the obvious choice is to treat the relative risk surface for non-chromosomal anomalies as the "exposure" surface and the simulations under the null and alternative hypotheses are simulations of cases of chromosomal defects.

In this situation the scatterplot between the observed relative risks of non-chromosomal anomalies for each individual derived from the adjusted relative risk surface (x-axis) and that of chromosomal anomalies (y-axis) is seen in Figure 4.27. Here, the definition of those in the background region are those individuals that do not have an excess risk of a non-chromosomal abnormality i.e. those with a relative risk less than or equal to 1. The scatterplot seems to indicate that there is no spatial association between the two surfaces.

The mean relative risk for chromosomal anomalies in the background region is 0.754, the proportion of the relative risk values above 1 in the exposed region is 0.416 and the gradient of the regression line in the exposed region is -0.139. The p-value is the proportion of simulations under the null hypothesis for which the 'height ratio' is greater than the 'height ratio' of the observed situation.

Figure 4.27: scatterplot between the observed relative risk of non-chromosomal anomalies and the observed relative risk of chromosomal anomalies for each birth in the 20km square region centred over Nant y Gwyddon 1983-1997



1000 simulations of chromosomal anomalies are performed under the null and the alternative hypotheses to carry out the kernel density test for spatial association. In order to simulate, the probability of having a chromosomal anomaly must be assumed for individual i under each hypothesis, this probability is denoted as p_i. The null hypothesis is that there is no spatial association therefore the scatterplot must have a completely random scatter of points. This random scatter is achieved by simulating under the assumption that the probability of a chromosomal anomaly is equal to that of the predicted probability of a chromosomal defect from the logistic regression model, α_i, i.e. $p_i = \alpha_i$ under the null. The alternative hypothesis is that there is spatial association thus the probability of a chromosomal anomaly is the predicted probability α_i multiplied by the magnitude of the relative risk surface for non-chromosomal abnormalities at the location of residence for birth i, denoted ζ_i, i.e. $p_i = \alpha_i \zeta_i$ under the alternative. This ensures that the simulations under the alternative produce adjusted relative risk contours for chromosomal anomalies that are correlated with that of non-chromosomal anomalies.

The 1000 simulations under each of the hypotheses are used to calculate the three test statistics (background mean, proportion and slope) and subsequently generate a kernel density for the joint distribution of the test statistics under each of the hypotheses. These two resultant kernel densities are used in the discriminant analysis part of the testing procedure. The 'height ratio', defined as

$$\frac{\text{the height of the kernel density for } \mathbf{x} \text{ under the alternative hypothesis}}{\text{the height of the kernel density for } \mathbf{x} \text{ under the null hypothesis}},$$

was calculated for the observed situation. The 'height ratios' were also calculated for each of the 1000 simulations under the null hypothesis. The proportion of 'height ratios' for the simulations under the null that were greater than that of the observed was 0.284 which gives the p-value. Consequently, it is not possible to reject the null hypothesis so there is evidence that there are other risk factors other than maternal age, year of birth, hospital of birth, Townsend score and gender that influence chromosomal defects in a different way to non-chromosomal defects.

In conclusion, chromosomal and non-chromosomal anomalies may have different aetiologies and these differences extend beyond the five risk factors discussed in this section. The differences in the aetiologies indicate that grouping all congenital malformations together for analysis may mean that the impact of some risk factors could be overlooked. The problem of grouping extends further than indicated here because there are numerous different types of chromosomal and non-chromosomal anomalies which can also differ in aetiology. Unfortunately, as explained previously, the rarity of a specific type of birth defect reduces the power of any statistical analysis as highlighted in parts of the spatial analysis of chromosomal anomalies hence investigators are forced to group different types of congenital malformations together. Although, grouping is usually unavoidable it is important to be aware of the problems that arise as a result.

4.4 The Risk of Congenital Malformations around Nant y Gwyddon 1998-2004

The data used in all the above analyses covers the period from 1983 to 1997 and was taken from NCAS and matched with denominator data from the CHS. The data reveals that there are regions around the Nant y Gwyddon landfill site for which the risk of congenital malformations is elevated. Additionally, these elevations cannot be explained by year of birth, gender, maternal age, socioeconomic deprivation and hospital of birth. This section focuses on

investigating whether or not these regions of elevated risk are still present when considering more recent data. Therefore, the risk of congenital malformations around the Nant y Gwyddon landfill site is investigated using data for 1998 to 2004. Furthermore, the relationships between the five risk factors previously discussed are reviewed using this more recent data.

Data on birth defects used for this section is from the Congenital Anomaly Register and Information Service for Wales (CARIS) and denominator data on healthy births from the CHS were matched to the CARIS dataset for 1998 to 2004, only births after 24 weeks were included. CARIS is one of the local registers that have been increasingly introduced across the UK from 1998 onwards. These local registers use multiple sources of information to identify cases of birth defects and supply this information to NCAS via electronic transfer.

CARIS was launched on the 1st January, 1998 and collects information on all foetuses and babies that have or are suspected to have a congenital malformation whose mother resided in Wales at the time of the birth [33, 280]. Data is collected for diagnoses between conception and the end of the first year of life. In addition to live and still births, pregnancies that end in termination or spontaneous loss before 24 weeks are recorded, although these are not included in this analysis since these were not recorded in the database for 1983 to 1997. CARIS actively seeks information via reporting forms, warning cards and reports from specialist sources and databases. Reporting forms contain all clinical information collected on any baby or foetus for which there is reasonable evidence of at least one birth defect. Warning cards flag potential cases. Specialist sources usually involve detailed diagnostic data. There are nominated coordinators for each delivery unit who help fill out and supply the warning cards and reporting forms.

The recorded rate of congenital anomalies for Wales taken from the CARIS dataset between 1998 and 2004 was 3.76%. This recorded rate is lower than that of 4.8% as quoted in the CARIS review, however, this rate includes terminations and spontaneous abortions before 24 week's gestation, which are excluded from the analysis presented here since they were excluded from the database for 1983 to 1997 [33]. Furthermore, the recorded rate of birth defects for this more recent time period is over three times that for 1983 to 1997. The increased recording rate does not relate to a true increase in the risk of congenital malformations but relates to the introduction of CARIS in 1998 which led to an increase in the rate of reporting [281]. The ascertainment rate of cases of congenital malformations for

NCAS was estimated by Boyd et al. to be only 40% whereas the local anomaly registers seek information on cases more actively therefore their ascertainment rate would be much higher [268]. Prior to 1998 CARIS did not exist and thus the more fastidious collection of data by CARIS compared to NCAS could account for the apparent increased rate in congenital malformations. In fact, the reported rates of congenital malformations are much higher in Wales than in England as seen in Table 4.10 provided by the ONS [282].

Table 4.10: rates of congenital malformations per 10,000 births in 2005 [282]

Birth Defect	England	Wales
CNS anomalies	5.5	8.5
Cleft lip and palate	7.7	16.2
Other face, ear and neck	2.6	4.9
Heart and circulatory	17.6	95.2
Alimentary	6.5	14.3
Genital organs	9.6	11.0
Urinary System	13.0	24.4
Musculoskeletal	29.9	50.0
Skin and integument	2.8	5.8
Chromosomal	9.8	29.3
Down's syndrome	6.7	13.1
Congenital metabolic disorders	2.6	8.9
Other	6.1	16.2

This greater level of reporting is corroborated by Dave Tucker from CARIS who states that "...the introduction of CARIS had a profound effect on reported rates of congenital anomalies in Wales". Regardless of the reason behind it, this change in the rate of birth defects between the two time periods prevents any formal comparisons. Consequently, the analysis presented in this section comprises the generation of relative risk contours and risk factor investigations as an exploratory analysis of whether the risk of birth defects is still elevated around the Nant y Gwyddon landfill site but no formal testing procedures are considered.

4.4.1 Possible Risk Factors

Any spatial patterns seen in the relative risk surface may be the result of specified risk factors. The five risk factors were considered for the 1998 to 2004 data to assess whether there is agreement between the two time periods regarding the relationship between each of the risk factors and congenital malformations. These relationships are summarised here and the relevant tables and figures are given in Appendix 4.

Year of birth

In general between 1998 and 2002 the recorded rate of birth defects is fairly constant, however, in 2003 and 2004 the rate of birth defects seems to fall. The reason for the drop may lie in late diagnosis or notifications of some congenital anomalies that may not be included in the dataset yet, which was last updated at the beginning of 2006. Disregarding this fall in risk in the last two years, the relatively constant rate through time is indicative of the more consistent collection of data maintained by CARIS compared to that prior to 1998 where changes in the reporting method led to large changes in the rate of birth defects by year of birth. Moreover, unlike between 1983 and 1997, where there were major changes in reporting caused by the introduction of the exclusion list and the extended time period during which a congenital malformation can be reported, there are no such changes for 1998 to 2004.

Gender

The proportion of birth defects is statistically significantly higher amongst male births compared to female births in 1998 to 2004, which is consistent with that reported using the data for 1983 to 1997.

Maternal age

The proportion of congenital malformations is much higher for births to mothers over 40 years of age and this elevation in reported rates is statistically significant. The rate of birth defects is also slightly higher amongst mothers under the age of 19 years although it is not as elevated as it is for mothers over 40 years. The pattern in the proportion of congenital malformations by maternal age seen for 1998 to 2004 mirrors that of 1983 to 1997.

Socioeconomic deprivation

There is a general increase in the proportion of congenital anomalies with increased deprivation in 1998 to 2004. However, the rate of birth defects is high among mothers in the most affluent group, but this is likely to be Down's to the maternal age effect. The rates for each of the Townsend quintiles are statistically significantly different (χ^2 test: p-value <

0.0005). The relationship between Townsend quintile and the rate of birth defects in 1998 to 2004 is different to that of 1983 to 1997. The reason for this disparity may lie in the differences between the Townsend quintiles in the two datasets. In the 1983 to 1997 dataset from NCAS, the Townsend quintiles were divided up using the quintiles of the Townsend scores of all the births in the dataset for Wales in that time period. In contrast, for 1998 to 2004 the Townsend quintiles are those set up using the whole population in the 2001 census. The difference in the way that the Townsend scores were split into quintiles could explain the difference in the relationships between Townsend quintile and birth defects for the two time periods.

Hospital of birth

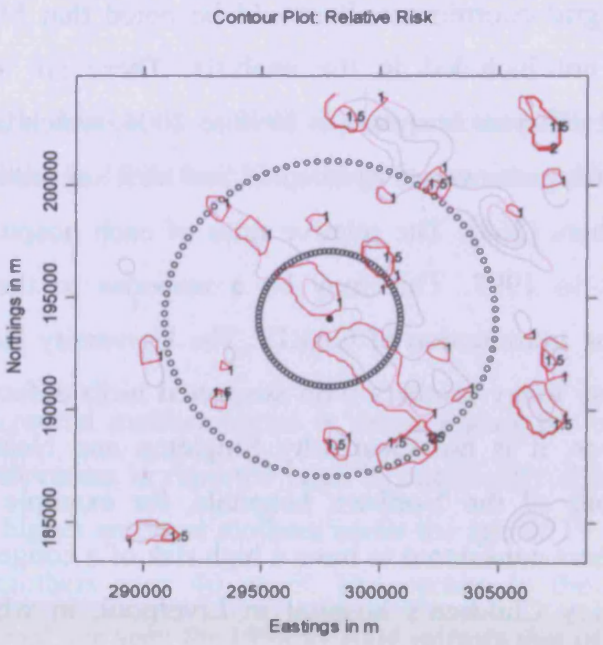
When assigning the hospital of birth for 1998 to 2004, if a birth did not occur at any of the 20 main hospitals with maternity wards then that birth was assigned to the nearest main hospital using the east and north grid coordinates. It should be noted that Mid Wales hospital has closed and is therefore not included in the analysis. There are statistically significant differences in the rates for different hospitals in 1998 to 2004, which is consistent with Boyd et al.'s claim that the reporting rate varies by hospital and trust suggesting that some prioritise notification more than others [268]. The relative rates of each hospital have changed from what they were in 1983 to 1997. This may be a response to the improvement in the ascertainment rate after the introduction of CARIS. The University Hospital of Wales has a high reporting rate because many mothers with suspected birth defects in South Wales are referred to there. However, it is not clear why Singleton and Neath have high rates of congenital anomalies. None of the Northern hospitals, for example West Cheshire, have higher rates, because mothers considered to have a high risk of a congenital malformation are often referred to Alder Hey Children's hospital in Liverpool, in which case, those births would be allocated to the nearest hospital to their residence at the time of the birth.

There have been some changes in the relationships between risk factors and birth defects for the two different time periods. However, to a certain extent all five risk factors still seem to influence the rate of congenital malformations. Consequently, the relative risk contours for congenital anomalies should be adjusted for these factors and this is presented below.

4.4.2 Adjusted Relative Risk Contours for Congenital Anomalies 1998-2004

The relationships between the five risk factors and the rates of congenital anomalies are similar for 1983 to 1997 and 1998 to 2004. The main difference lies in the rates for each of the hospitals with different hospitals reporting the highest and lowest rates for the two time periods. The optimal logistic regression model for all congenital anomalies no longer contains Townsend quintile and maternal age, see Appendix 4. The reason for the loss of significance for the inclusion of some of the variables in the logistic regression model possibly lies in the smaller number of births for 1998 to 2004 or could lie in the better quality of data collected during this period. Only the full models are used so that comparisons can be made between the adjusted relative risk contours for 1983 to 1997 and those for 1998 to 2004.

Figure 4.28: relative risk contours for congenital malformations adjusted for confounding in the 20km square region centred over Nant y Gwyddon 1998-2004



The relative risk contours for congenital malformations with adjustment for confounding is given in Figure 4.28. The four hotspots that were present in 1983 to 1997 after adjustment for confounding are no longer present in 1998 to 2004. Furthermore, there is only one very small region over which the relative risk surface exceeds 2. Consequently, if there was a risk factor that was leading to a spatially heterogeneous pattern in the risk of birth defects then it is no longer active in 1998 to 2004. For example, the dumping of waste at the Nant y Gwyddon

landfill site ceased in March, 2002. Alternatively, the appearance of regions of elevated risk of birth defects around the Nant y Gwyddon landfill site in 1983 to 1997 may be down to the poorer quality of data collected during that period relative to that collected in 1998 to 2004, i.e. the hotspots could have been a consequence of reporting patterns by midwives.

4.5 Conclusion

The results presented in this Chapter provide a more detailed insight into the risk of congenital malformations around the Nant y Gwyddon landfill site. Initially, a number of possible risk factors were investigated. The rate of congenital anomalies was found to vary by year of birth with the introduction of the exclusion list [277], the increase in the length of the period within which a birth defect could be reported [278], deficiencies in reporting ascertainment by NCAS and changes in the diagnostic procedures. The rate of congenital malformations is higher amongst male births compared to female births. There is a maternal age effect with a greater risk of Down's syndrome amongst mothers of an older age [23] and an increased risk of gastroschisis amongst younger mothers [23]. The rate of congenital malformations increases slightly with increased deprivation although there is a slightly higher rate in the most affluent group, which reflects the higher proportion of older mothers in the most affluent quintile coupled with the maternal age effect. Some hospitals have higher reported rates of congenital malformations because mothers with a high risk of a birth defect can be referred to hospitals with specialist services thus increasing the rate at that hospital [23]. Moreover, the diligence of the midwives in notifying NCAS varies [278] and some hospitals prioritise notification to NCAS more than others [268].

There are statistically significant elevations in the relative risk surface for birth defects in close proximity to the landfill site in 1983 to 1997 after adjustment for year of birth, gender, maternal age, socioeconomic deprivation and hospital of birth. These elevations could be the result of inadequate adjustment for the five confounding risk factors and consequently the effects of these factors would still be present in the pattern of the resultant adjusted contours. Therefore, the adequacy of the adjustment of the relative risk surface for confounding was assessed. The results of this assessment indicate that the adjustment for gender was adequate and for maternal age and socioeconomic deprivation is not perfect but is a move in the right direction. The adjustment for hospital of birth could not be assessed and the comparison made for year of birth may not be a completely fair comparison. Ultimately, the adjustment for

confounders was thought to be reasonable since it is unlikely that any method used to adjust for confounding factors would be perfect and this method seemed to be a move in the right direction.

The belief that the adjustment for confounding is adequate means that it is unlikely that inadequate adjustment for the five risk factors is responsible for the statistically significant elevations in the adjusted relative risk surface. Therefore, it is more likely to be a result of risk factors that have not yet been included in the adjustment. One of these risk factors could be exposure to contaminants released from the Nant y Gwyddon landfill site.

However, the hypothesis that exposure to contaminants released from the Nant y Gwyddon landfill site increases the risk of congenital malformations does appear to be refuted by the lack of elevations in the adjusted relative risk contours around the landfill site in 1998 to 2004. This apparent lack of peaks in the relative risk surface in 1998 to 2004 and the existence of a statistically significant peak within 3km of the landfill site in 1983 to 1997 can be explained in a number of ways. For example, the elevations in the relative risk surface for congenital malformations in 1983 to 1997 may be a random event but the statistical significance of the hotspots indicates that this is unlikely to be the case. In which case, the elevations could be down to poor data collection by the registers in that period and in this situation the patterns in the relative risk contours could occur because of inconsistent reporting rates between midwives, hospitals and trusts in 1983 to 1997.

Alternatively, the elevations in the relative risk surface could be the result of a spatially heterogeneous risk factor for congenital anomalies. Therefore, in this situation it is the lack of spatial heterogeneity of risk in 1998 to 2004 that needs to be explained. It may be that the risk factor that caused the elevations in 1983 to 1997 is no longer active in 1998 to 2004. For example, if the risk factor was exposure to contaminants from the landfill site then it may be that the amount of contaminants released from the site has subsided with time. This could be the case because the Nant y Gwyddon landfill site was closed in March 2002.

Another possibility is that the more rigorous data collection performed by CARIS could have lead to the recording of less severe congenital malformations the inclusion of which could dilute the effect of the risk factor. This problem of diluting effects of risk factors is highlighted in the consideration of the impact of maternal age where the increased risk of non-

chromosomal anomalies amongst younger mothers was masked by the inclusion of chromosomal anomalies which have increased risk amongst older mothers. Moreover, a test for spatial association between chromosomal and non-chromosomal anomalies revealed that there were risk factors that influenced the two types of birth defects differently other than the five already considered in this Chapter.

Ultimately, it is not possible to conclude whether or not exposure to contaminants released from the Nant y Gwyddon landfill site influences the risk of congenital malformations without good quality exposure data and consistent quality in recording birth outcomes.

Chapter Five – Application of Spatial Methodology to Sex Ratios

Sex ratios are reportedly changing temporally in many countries [76-79, 93, 94, 102, 103, 134]. In the developing world, the proportion of boys born is increasing, for example, in China and India [79, 93, 94]. This increase may be a result of preferences for male births leading to practices that influence the gender of a child, such as female infanticide. However, some believe that the increase in male births is in fact a product of improvements in maternal well-being reducing the number of foetal deaths [94]. There is thought to be a larger proportion of male foetal deaths and, consequently, in response to the decrease in the number of foetal deaths, the proportion of male births increases. The measures taken by parents in developing countries to influence the gender of their children make it difficult to separate the result of this manipulation of the sex ratio from that of natural factors. Furthermore, there lies an inadequacy in the registration of vital statistics in developing countries, which makes it difficult to obtain accurate secondary sex ratios [75]. Therefore, the focus of many studies is on data from Westernized countries.

In contrast to developing countries, Westernized countries have seen a decline in the proportion of male births post-World War II, which is, as yet, unexplained [76-78, 102, 103]. Several attempts have been made to account for the recent fall in the sex ratio. However as a result of inconsistent or non-significant findings none of these explanations have been conclusive. The main risk factor under consideration is that of environmental contaminants [104, 105, 107, 112-118] but other factors have been mentioned which include timing of insemination [121, 122, 124-126], environmental conditions [128], stress [129-132, 134-136, 138], nutrition [126, 139-143], parental drug use [144, 145, 147], family status [87, 88, 150-153] and ethnicity [154].

The existence of these downward trends in Westernized countries has motivated the research presented in this Chapter. Initially, the temporal pattern was investigated in England and Wales to determine whether England and Wales also encountered a decrease in the sex ratio through time. Information on the number of live and still births of each gender in each year for England and Wales was taken from the Office of National Statistics (ONS) website. All births must be registered by law within 42 days of their occurrence. The records of each live and still birth registration contain information on the time of birth, mother's age, father's age, marital status, social class, multiple maternities and area of usual residence. The records of

live and still birth registration across England and Wales are compiled by the ONS and are presented as counts by year on the ONS website. These yearly counts are used to investigate the temporal pattern in sex ratio for England and Wales.

The presence of temporal patterns in the sex ratio suggests that it is not unreasonable to assume that there could also be a spatial pattern. However, the ONS only present this data in terms of counts and it is not possible to obtain geographical information at an individual level. The absence of this information means that the data from the ONS cannot be used to investigate embedded spatial patterns.

The largest dataset available with geographical information on each birth and the corresponding gender is that of the Child Health System (CHS). The CHS contains information on the date of birth, birth weight, gestational age, maternal age, hospital of birth, parity and Townsend score of the father. This information is available for between 1983 and 1997. Unfortunately, only maternal age, Townsend score and parity are of interest as possible risk factors for the sex ratio. This group of explanatory variables was thought not to be complete enough to explain potential spatial patterns and so in conclusion the CHS is not the best dataset to use. Moreover, the relatively small time period which the CHS covers makes it more difficult to pick up any temporal element in the spatial patterns.

The Cardiff Birth Survey (CBS) provides geographical information on all births between 1965 and 2004 for Cardiff and the Vale of Glamorgan. The geographical information consists of addresses from 1989 onwards and geographical area prior to that. This database also contains information on some possible risk factors that are of interest including maternal age, parity, maternal ethnicity and maternal BMI. Additionally, the geographical information enables the Townsend score to be determined for each birth. Therefore, the CBS contains information necessary to generate high resolution maps that will enable localised variations in the sex ratio to be identified.

Consequently, the spatial patterns in the sex ratio are to be performed on data from the CBS. However, before proceeding the temporal trend in Cardiff needs to be investigated to assess whether or not the temporal trend seen for England and Wales is present in Cardiff too. If it is then the temporal element in the variation of the sex ratio should be considered in the spatial patterns, because if the overall sex ratio varies temporally then the spatial pattern may change

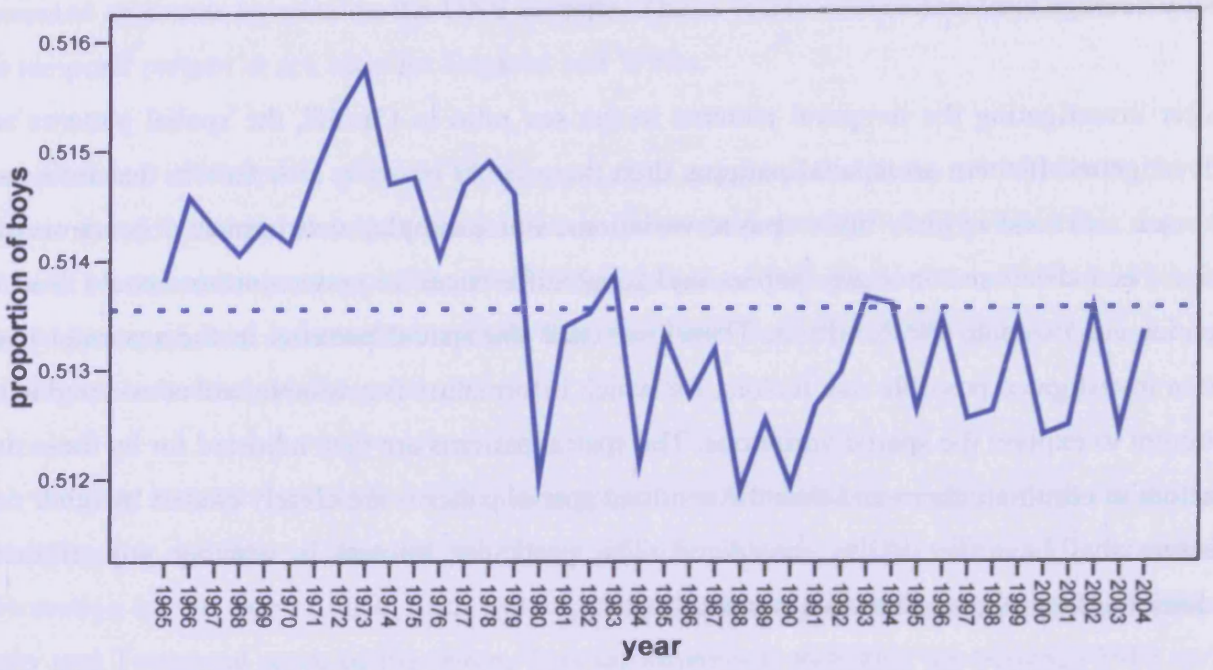
through time too. It should be noted that data for 1987 and 1988 are missing as a result of flood damage to records.

After investigating the temporal patterns in the sex ratio in Cardiff, the spatial patterns are investigated. If there are spatial patterns then there could be some risk factors that influence the sex ratio and explain these spatial variations. For example, environmental contaminants have been discussed in many papers and local differences in concentrations could lead to spatial variations in the sex ratio. Therefore, once the spatial patterns in the sex ratio have been investigated possible risk factors, for which information is available, are considered in an attempt to explain the spatial variations. The spatial patterns are then adjusted for by these risk factors to eliminate them and thus the resultant spatial patterns are clearly caused by other risk factors that have yet to be considered. The particular interest is whether any of these additional factors could be environmental.

5.1 Temporal Trends in Westernized Countries

The overall sex ratio for England and Wales between 1965 and 2004 is 0.5136 which is the proportion of boys calculated from information from the ONS. Figure 5.1 is a temporal plot revealing how the sex ratio varied year by year between 1965 and 2004. The solid line represented the annual sex ratio and the dotted line depicts the overall sex ratio for the period of interest. There is, consistently, a larger proportion of males in all years however these proportions have changed year by year. Despite prominent annual variation in the sex ratio, it is clear that the sex ratio has been falling over the period of study with the proportion of female births increasing. A surprising feature in Figure 5.1 is a substantial drop in the sex ratio between 1979 and 1980. There is no obvious reason for this rapid fall in the proportion of boys. It is important to note the scale on Figure 5.1 revealing that the effect size is very small. However, a small change in the sex ratio equates to a large change in the numbers of each gender born that could lead to changes in social behaviour in years to come. Additionally, slight changes of gender may be a bi-product of changes in social or environmental conditions that could also lead to more adverse effects, for example, an increase in the rate of some form of congenital anomaly.

Figure 5.1: temporal pattern of the sex ratio in England and Wales 1965-2004



The temporal pattern in the sex ratio for England and Wales has been compared to that of other Westernized countries including USA, Canada [76], Germany [134], Denmark [76], Sweden [145], Holland [136] and Spain [78] and these comparisons can be seen in Figure 5.2 to Figure 5.7. The majority of these countries exhibit similar patterns in the annual sex ratio with the proportion of boys appearing to gradually decline and the figures are close to the overall sex ratio for England and Wales of 0.5136. However, USA has a sex ratio that is consistently lower than 0.5136, which could relate to the different racial mix of the population of the USA relative to that of England and Wales. The disparity between the annual sex ratio of the USA and the overall sex ratio of England and Wales is only slight. A more noticeable difference occurs between the temporal pattern of Spain and those of the other countries considered. In Spain the sex ratio increases until 1981 when it eventually begins to drop in accordance with the remaining Westernized countries investigated. The reason given for this increase by Gutierrez-Adan et al. is the improvement in obstetrical care, which led to the decline in foetal deaths, a larger proportion of which are male, and, thus, an increase in male births occurs [78]. However, it is unlikely to be the cause of such a dramatic change and it may in fact be a response to problems in data collection.

Figure 5.2: temporal trend in sex ratios for USA and Canada compared to that of England and Wales [76]

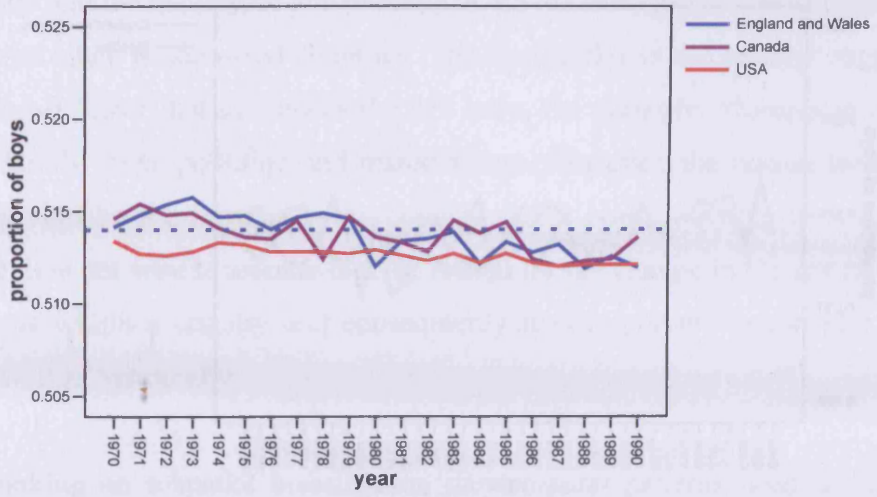


Figure 5.3: temporal trend in sex ratios for Germany compared to that of England and Wales [134]

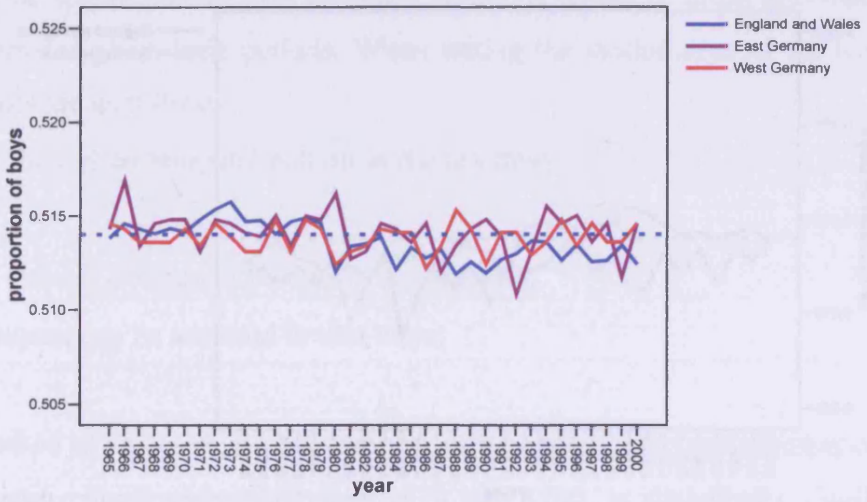


Figure 5.4: temporal trend in sex ratios for Denmark compared to that of England and Wales [76]

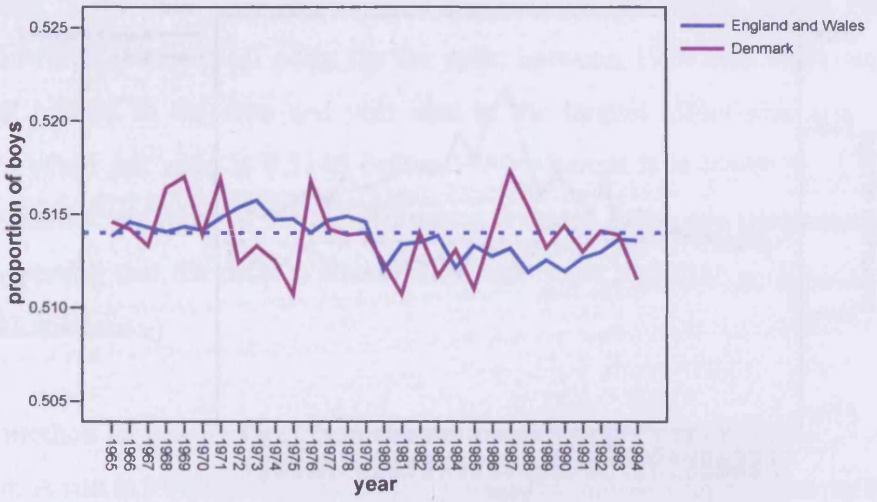


Figure 5.5: temporal trend in sex ratios for Sweden compared to that of England and Wales [145]

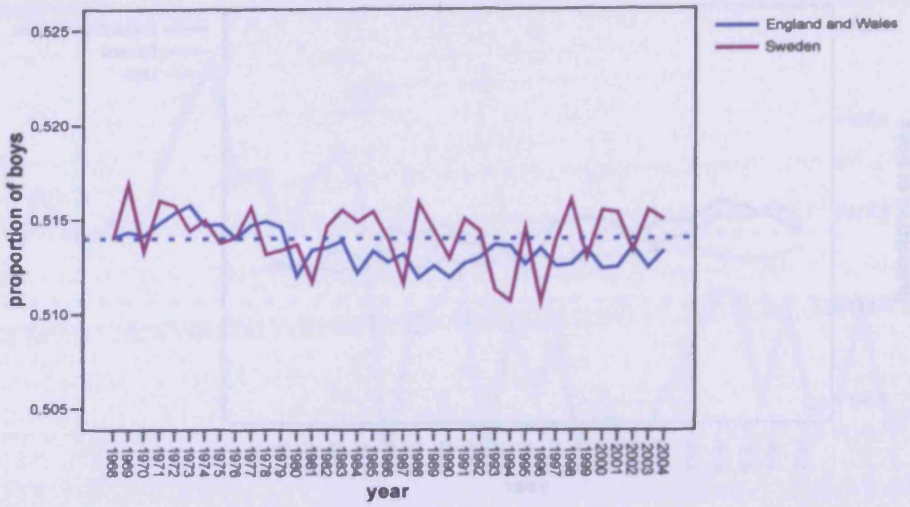


Figure 5.6: temporal trend in sex ratios for the Netherlands compared to that of England and Wales [136]

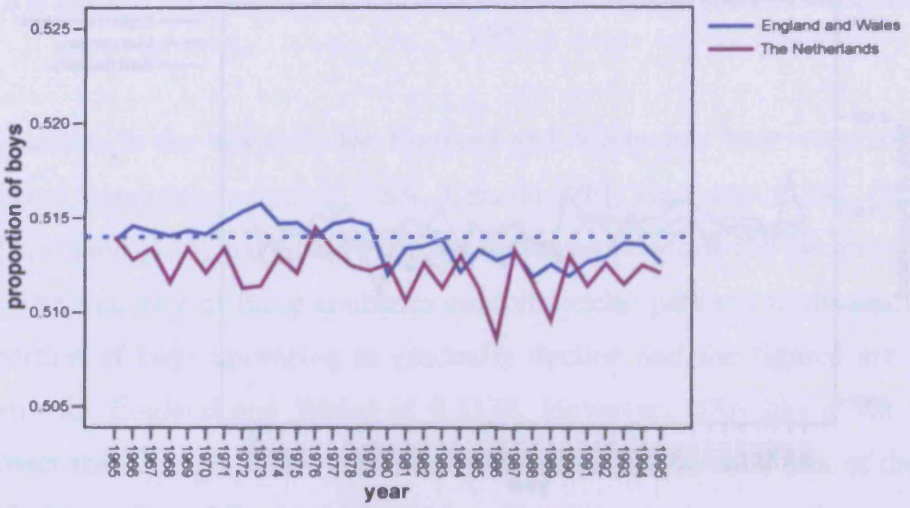
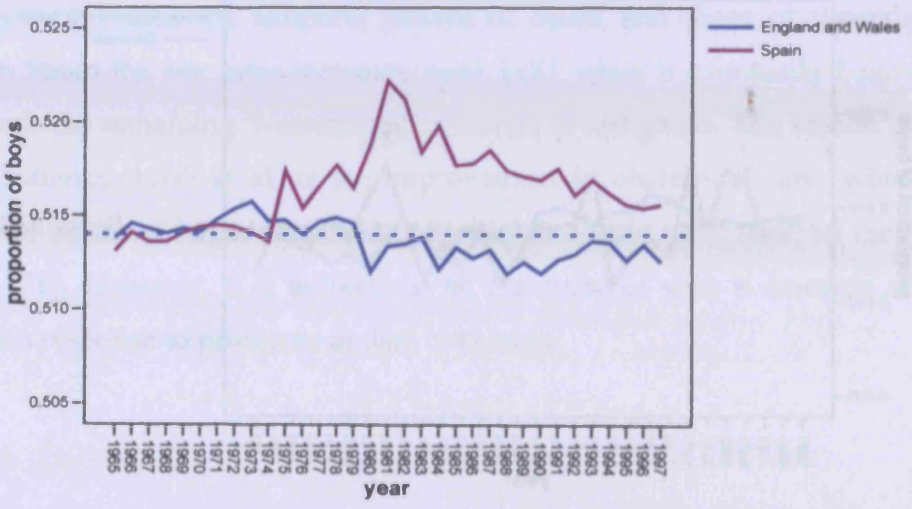


Figure 5.7: temporal trend in sex ratios for Spain compared to that of England and Wales [78]



There does appear to be a temporal pattern in the sex ratio in England and Wales between 1965 and 2004 and the magnitude and patterns in the sex ratio profile are similar to those seen in a number of other Westernized countries. The similarities in the profiles suggest that there may be a global factor that influences the sex ratio, for example, changes in environmental conditions, family sizes, pollution and maternal age. However, the reason for the decline in sex ratio may not be the same in every country and it is not sensible to assume that it is. Furthermore, it is not wise to assume that the reason for the change in the sex ratio is the same for all regions within a country and consequently it is important to consider the sex ratio spatially as well as temporally.

Before embarking on a spatial investigation the temporal patterns need to be investigated further. If there is an apparently non-random temporal effect then it would be unwise to simply plot the spatial pattern over all years, instead it would be more informative to plot the spatial pattern for given time periods. When testing the randomness of the temporal pattern the hypotheses are as follows:

H_0: there is no temporal pattern in the sex ratio

and (H5.1)

H_1: there is a temporal pattern in the sex ratio.

These hypotheses can be assessed in two ways.

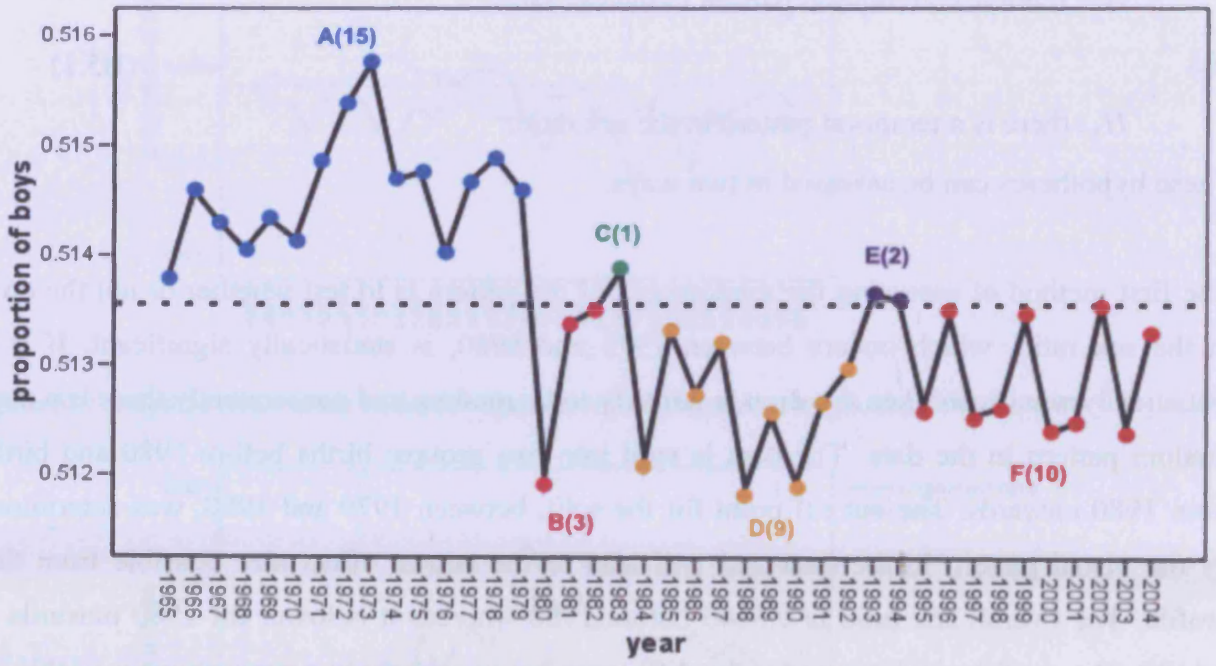
The first method of assessing the randomness of the pattern is to test whether or not the drop in the sex ratio, which occurs between 1979 and 1980, is statistically significant. If it is statistically significant then the drop is unlikely to be random and consequently there is a non-random pattern in the data. The data is split into two groups: births before 1980 and births from 1980 onwards. The cut off point for the split, between 1979 and 1980, was determined by the visual pattern in the data and will lead to the largest effect size possible from this profile. The overall sex ratio is 0.5146 before 1980 whereas it is lower for 1980 onwards at 0.5129. The confidence interval for the difference between these two proportions is (0.00129, 0.00206) suggesting that the drop is statistically significant and thus is unlikely to be a result of random fluctuations.

The second method of testing these hypotheses can be carried out by considering runs in the temporal plot. A run is a string of consecutive points that are all either above or below the line

drawn where the proportion of boys is equal to the overall sex ratio. If the pattern was random then the annual sex ratio would move above and below the reference level more frequently than if there was a risk factor influencing the temporal trend that causes a systematic bias. Therefore, there would be a larger number of runs of smaller length if the temporal pattern was random. Hence, the number of runs or the mean number of runs can be used to assess the probability that the temporal pattern is a random effect. Both the number of runs and the mean run length produce the same results in terms of the p-value so only the number of runs is considered.

The temporal plot for England and Wales is given in Figure 5.8 with the runs labelled by letters A to F and the numbers in the brackets refer to the length of each of these runs. There are 6 runs between 1965 and 2004. The largest run is of length 15 and the smallest run only has length 1.

Figure 5.8: temporal trend in the sex ratio for England and Wales with the runs labelled 1965-2004



The p-value for the testing procedure using runs can be calculated in a simulation exercise. The simulations are performed under the null hypothesis i.e. under the assumption that gender determination is a random event where all births have the same probability of being a boy in each simulation. In a given simulation, the number of boys for a specific year was generated by randomly sampling from a binomial distribution where the event of interest was a male

birth. In the binomial distribution for a specific year the sample size, n, was equal to the number of births in that given year in England and Wales. The probability of a male birth, p, was equal to the overall sex ratio in England and Wales of 0.5136.

Thus, the result of one simulation comprises the number of male births in each year generated from a binomial distribution dependent on the total number of births for that year where the probability of a male birth is 0.5136 for all births. The proportion of boys can then be determined for each year in that simulation as can the overall sex ratio. The number of runs in that simulation is then the number of times that the sex ratio falls below or rises above the corresponding overall sex ratio for that simulation plus 1. This process is repeated 1000 times to generate 1000 counts of runs for simulations performed under the assumption that the gender of the child is random. The p-value is estimated by the proportion of these simulations for which the number of runs is less than the true number of runs in the sex ratio profile for England and Wales which is equal to 6. If the p-value is small, conventionally any value below 0.05, then a statistically significant result is obtained and it is unlikely that the profile in the sex ratio is random.

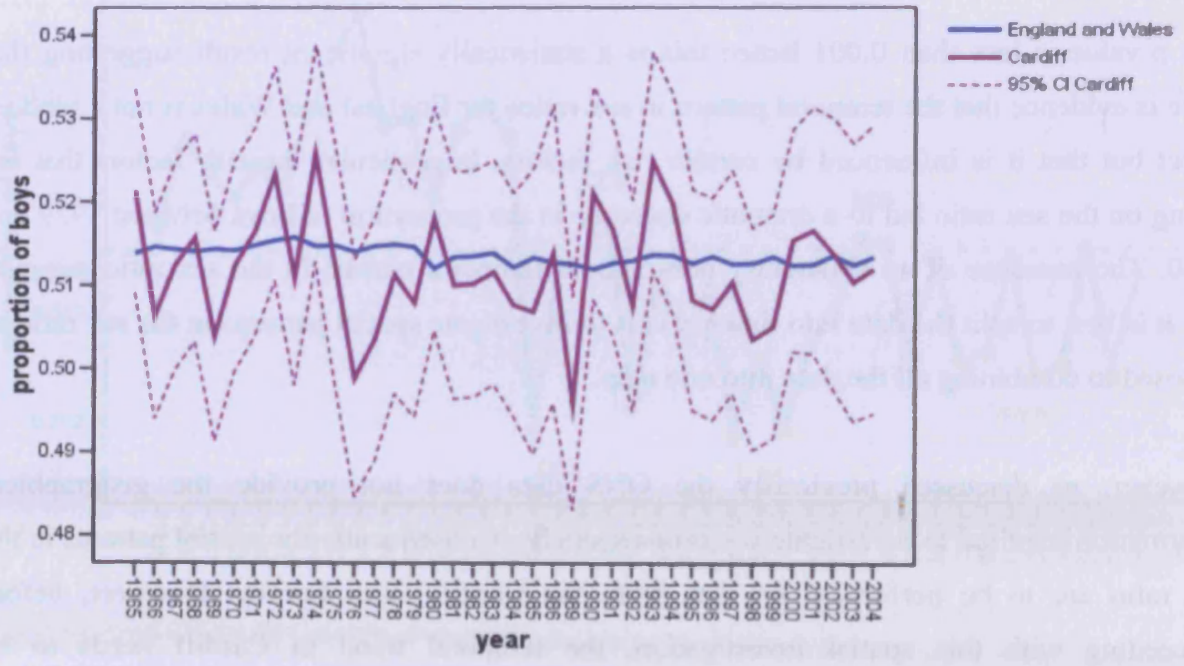
The p-value is less than 0.001 hence this is a statistically significant result suggesting that there is evidence that the temporal pattern in sex ratios for England and Wales is not a random effect but that it is influenced by certain risk factors. In particular the risk factors that are acting on the sex ratio led to a dramatic decrease in the proportion of boys between 1979 and 1980. The presence of an apparently non-random temporal pattern in the sex ratio suggests that it is best to split the data into time periods to investigate spatial patterns in the sex ratio as opposed to combining all the data into one map.

However, as discussed previously the ONS data does not provide the geographical information required to investigate sex ratios spatially. Consequently the spatial patterns in the sex ratio are to be performed on data from the Cardiff Birth Survey. However, before proceeding with this spatial investigation, the temporal trend in Cardiff needs to be investigated to assess whether or not the temporal trend seen for England and Wales is present in Cardiff too.

5.2 Temporal Trends in Cardiff

The overall sex ratio for Cardiff and the Vale between 1965 and 2004 is 0.5120, which suggests that the proportion of female births is slightly higher than at the national level. The temporal pattern of the sex ratio in Cardiff and the Vale is given in Figure 5.9 along with that of England and Wales. The purple dotted lines represent the 95% confidence interval of the sex ratio for Cardiff and the Vale. The national sex ratio lies within this 95% confidence interval for all but two years, 1976 and 1989, where the sex ratio from the Cardiff Birth Survey was significantly lower than that of England and Wales. However, using a 95% confidence interval means that 5% of the time the sex ratio for England and Wales will lie outside the interval. Thus, over 38 years it is expected that the sex ratio for England and Wales will not be inside the interval for approximately 2 of those years, which is what has occurred. Therefore, the temporal sex ratio profile for Cardiff and the Vale of Glamorgan seems to be consistent with that of England and Wales.

Figure 5.9: temporal trend in the sex ratio in Cardiff and the Vale 1965-2004



5.3 Spatial Patterns in the Sex Ratio

Investigating spatial patterns in a given health outcome requires mapping techniques. The CBS contains information on the district council within which each mother resided and thus

initially the sex ratio is mapped by the district council for each five year time period. However, in Chapter 3 it was determined that, with a modest amount of good quality data, the best way to map a health outcome is to use kernel density methodology to create relative density contours. In order to generate relative density contours each individual in the study population requires an east and north grid reference, which pinpoints the location of their residence at the time of the event of interest, in this case the birth. Unfortunately, the addresses for each mother are only available from 1989 and the required pinpointed geographical locations cannot be obtained from the geographical area alone so the kernel density mapping technique has only been applied to data for 1989 to 2004.

Therefore, in this Section the sex ratio is mapped using the geographical area for five year intervals. Then for those time intervals for which addresses are available for each of the births the kernel density method of mapping risk is considered without adjustment. The global significance test for spatial homogeneity is used to test the statistical significance of regions with either an excess density of boys or an excess density of girls. If there are statistically significant regions of excess density of males and/or females then there may be a spatially heterogeneous risk factor that influences the sex ratio.

5.3.1 Mapping Sex Ratios using Area Level Data

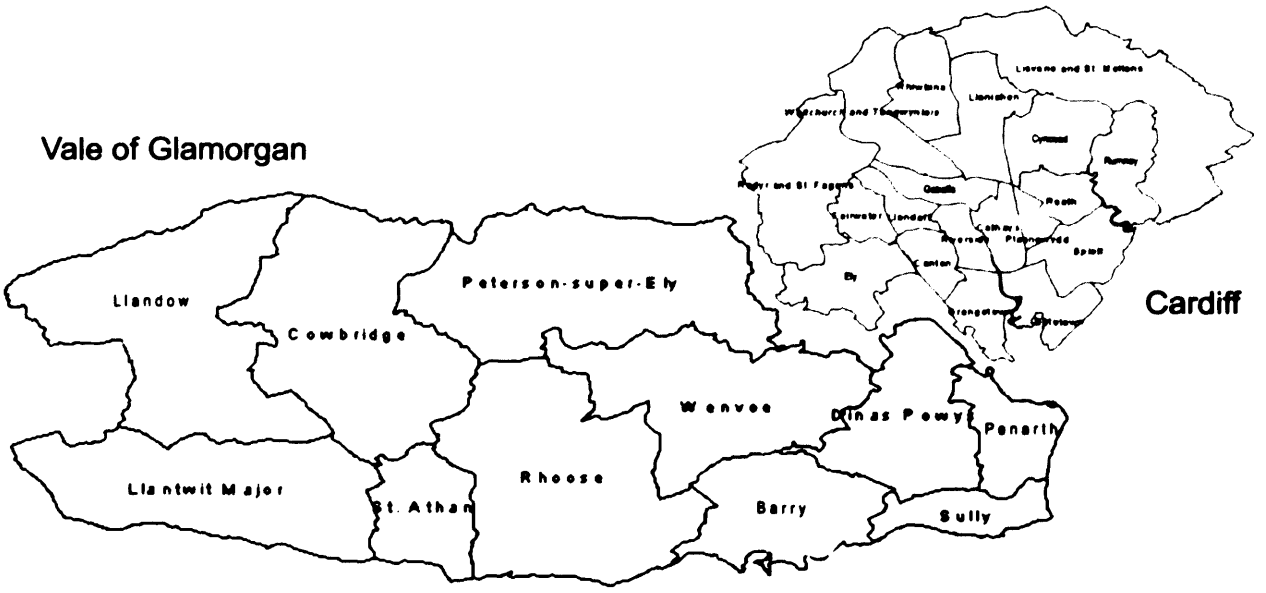
In the Cardiff Birth Survey the geographical area refers to the council district. The council district boundaries are used to define the regions over which the sex ratio is mapped. The database for the Cardiff Birth Survey contains information on the district councils in which each of the mothers resided at the time of the birth. The sex ratio or proportion of boys is calculated for all the mothers who reside in each council district and that sex ratio is then used to determine the colouration used in that council district on the map. There were a number of difficulties that had to be overcome in order to map the sex ratio by council district which are discussed below.

The council district boundaries changed between 1974 and 1975 and also between 1982 and 1983. Prior to 1974 there were only five district councils in Cardiff and the Vale: Cardiff Central Borough, Penarth, Barry, Cardiff Rural District and Cowbridge. The small number of regions over which the mothers are aggregated prevents the generation of high resolution maps which are required to gather detailed information of the spatial variation of the sex ratio.

In addition, data before 1974 is of questionable quality and, therefore, maps have not been created for the sex ratios before 1974.

After 1974 the council districts became smaller which enables generation of maps of a higher resolution. The only difficulty that is faced when mapping data for 1975 onwards is the revision of the boundaries during that period, which fortunately was not as radical as the previous revision. Hence, this problem could be overcome by combining some of the district councils. The finalised regions used for mapping the sex ratio are given in Figure 5.10. Cardiff and the Vale of Glamorgan are separated in the map because the Vale of Glamorgan is rural and consequently has a much lower population density than Cardiff which is a city thus the patterns are more likely to be random for the Vale.

Figure 5.10: boundaries for regions used to map the sex ratio in Cardiff and the Vale of Glamorgan



An additional problem is that the dataset containing the information required for births from 2000 onwards does not include information on the district council. However, grid coordinates have been determined from the address of the mother at the time of the birth which enabled the district council to be determined using ArcGIS.

The annual number of births in each of the council districts can be low particularly in the Vale of Glamorgan, which increases the level of random variation. Therefore, the births were grouped into six five year blocks: 1975 to 1979, 1980 to 1984, 1985 to 1989, 1990 to 1994,

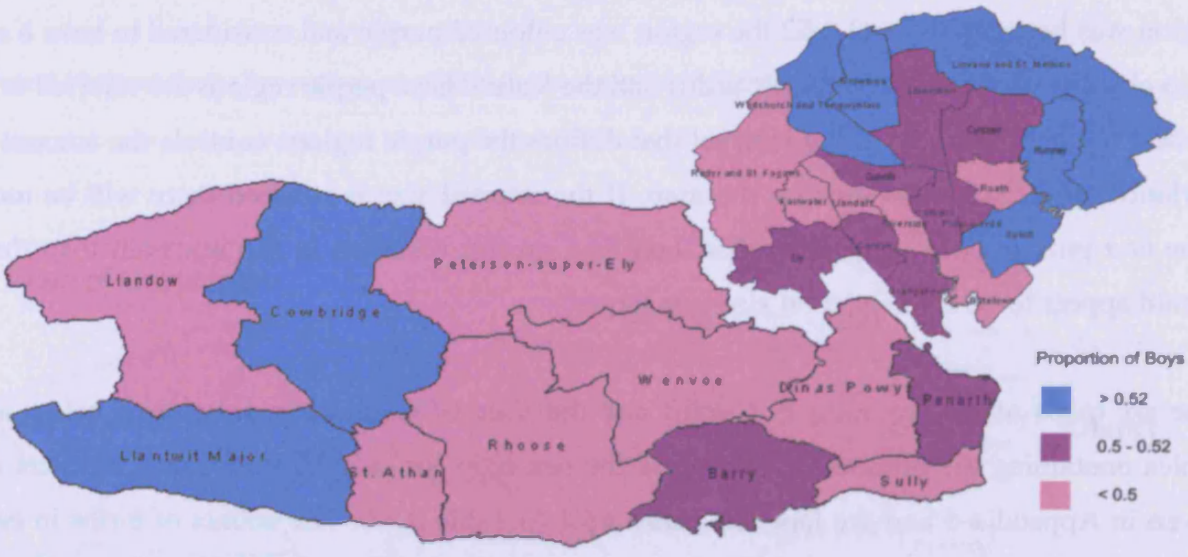
1995 to 1999 and 2000 to 2004. Although, as will be seen later, the number of births in some of the regions is still low when five years are considered together, it does reduce random variation slightly whilst still providing information on how the sex ratio varies temporally as well as spatially.

Figure 5.9 indicates that the annual sex ratio in Cardiff typically falls between 0.5 and 0.52. Therefore, when mapping the sex ratios, regions with a sex ratio of 0.52 were considered to have a higher than expected proportion of boys. These regions are referred to below as male dominated regions and are blue in the maps. Regions with a sex ratio below 0.5 were considered to have a higher than expected proportion of girls. These regions are pink on the maps and are referred to in the text as female dominated regions. When the sex ratio for a region was between 0.5 and 0.52 the region was coloured purple and considered to have a sex ratio close to what is expected for Cardiff and the Vale. These purple regions are referred to as neutral regions. The size of the interval that defines the purple regions controls the amount of variation in the sex ratio seen on the map. If the interval size is reduced there will be more blue and pink regions suggesting that there is a greater variation in the sex ratio than there would appear to be if the interval size was larger.

The six maps of the sex ratio in Cardiff and the Vale of Glamorgan are shown below and tables containing the number of births and the sex ratio for each of the council districts are given in Appendix 5 and are labelled Table A5.1 to Table A5.6. The counts of births in each of the districts and their size reveal how sparsely populated some of the districts in the Vale of Glamorgan are relative to Barry, Penarth and the Cardiff council districts. These sparse populations are reflected in the sex ratios, which have a larger variability in the Vale of Glamorgan. Spatial patterns have emerged from the maps with bands of male dominated regions and some of female dominated regions. These patterns are discussed in more detail below.

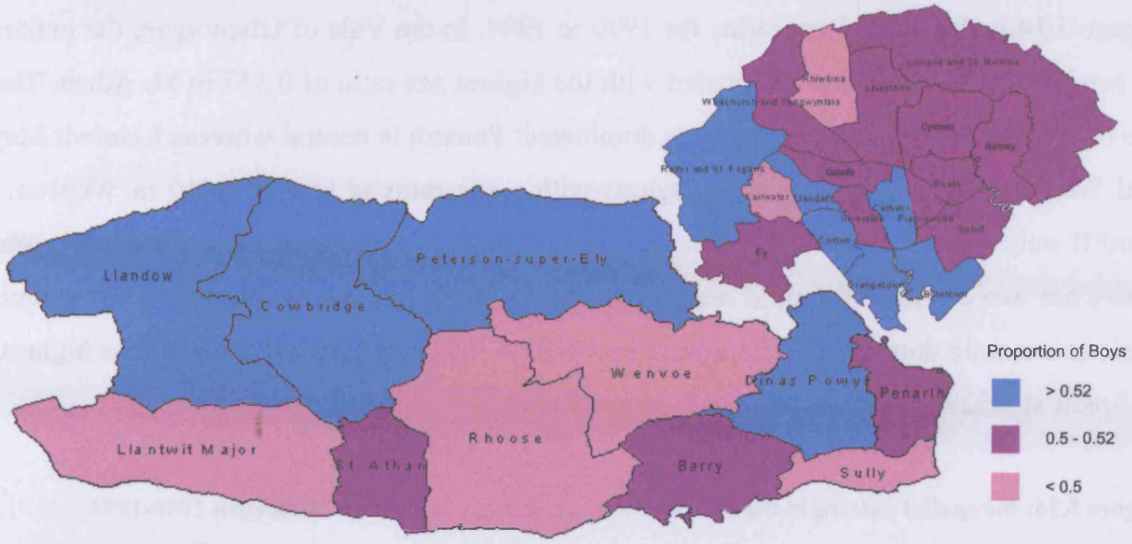
The sex ratio for Cardiff and the Vale in 1975 to 1979 is mapped in Figure 5.11. The majority of the Vale of Glamorgan is female dominated with the lowest sex ratio in Wenvoe of 0.412. There is a small male dominated band, consisting of Llantwit Major and Cowbridge, in the West. However, the two council districts that have the largest populations, Barry and Penarth, are both neutral. In Cardiff there is a band of male dominated regions in the North East with a peak sex ratio in Rhiwbina of 0.558. This band surrounds a central, neutral region and towards the South West of that neutral region there is a band of female dominated regions running from Radyr and St. Fagans down to Butetown. In the far South West the remaining districts are neutral.

Figure 5.11: the spatial pattern in the sex ratio in Cardiff and the Vale of Glamorgan 1975-1979



In 1980 to 1984 the pattern has changed and is shown in Figure 5.12. The North of the Vale of Glamorgan is male dominated where the sex ratio is as high as 0.551 in Peterson-Super-Ely. The South is female dominated with the exception of St. Athan, Barry and Penarth which are neutral and Dinas Powys which is male dominated. The sex ratio is still lowest in Wenvoe at 0.434. The majority of Cardiff is neutral suggesting that the spatial variability in the sex ratio across Cardiff has reduced since the previous time period. However, there is a male dominated band running between Radyr and St.Fagans and Butetown. The sex ratio is highest in Cathays at 0.541, which lies within this band. There are only two female dominated regions, Rhiwbina and Fairwater, and they are non-adjacent. The sex ratio is lowest in Rhiwbina at 0.487 whereas previously it had the highest sex ratio in Cardiff.

Figure 5.12: the spatial pattern in the sex ratio in Cardiff and the Vale of Glamorgan 1980-1984



The map for 1985 to 1989 is presented in Figure 5.13, note that data for 1987 and 1988 are missing. In the Vale of Glamorgan the sex ratio is the highest at 0.594 in Llandow, which is part of a cluster of male dominated regions in the West. This cluster is punctuated by Cowbridge, which is neutral, and St.Athan, a female dominated region. In the East, there is a cluster of female dominated regions where the sex ratio is lowest in Sully at 0.446. Penarth is neutral and, despite its large population size, Barry is female dominated. The North West of Cardiff is male dominated with the highest sex ratio of 0.525 in Lisvane and St. Mellons. There is a wide band of female dominated regions running between Radyr and St. Fagans and Butetown. The lowest sex ratio of 0.451 is in Butetown.

Figure 5.13: the spatial pattern in the sex ratio in Cardiff and the Vale of Glamorgan 1985-1989

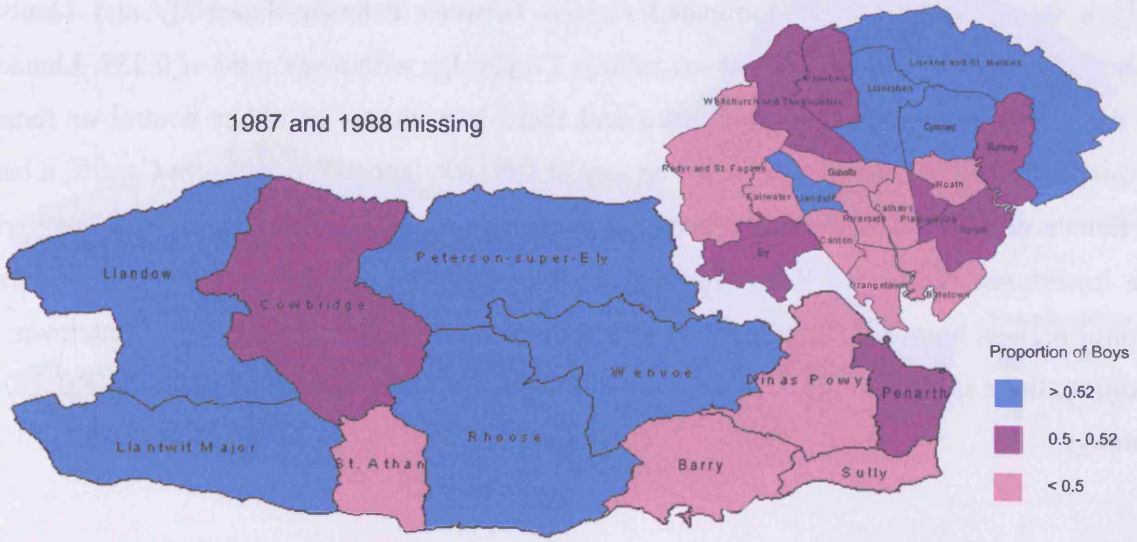
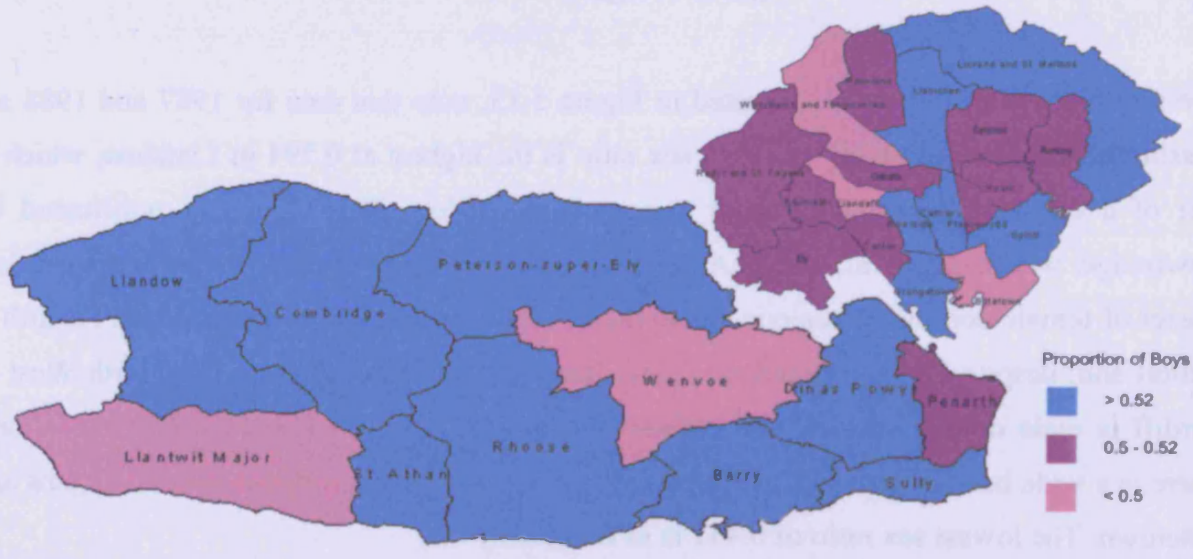


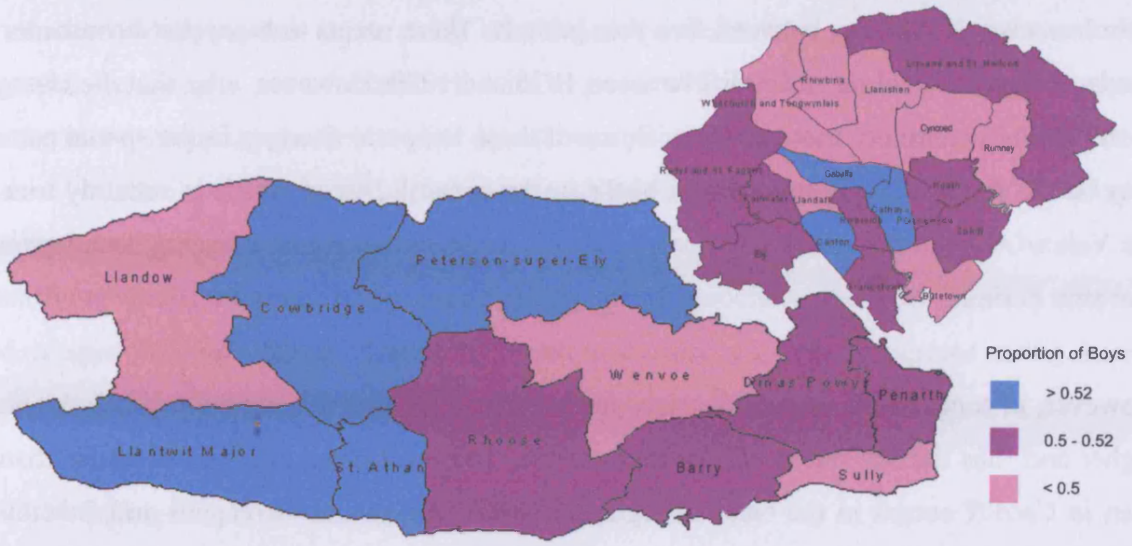
Figure 5.14 is the map of sex ratios for 1990 to 1994. In the Vale of Glamorgan, the majority of council districts are male dominated with the highest sex ratio of 0.557 in St. Athan. There are only three regions that are not male dominated: Penarth is neutral whereas Llantwit Major and Wenvoe are female dominated regions with a sex ratio as low as 0.410 in Wenvoe. In Cardiff only three regions are female dominated, none of which are adjacent to each other. There are two clusters of neutral regions one in the West and one in the East. The remaining regions are male dominated. The lowest sex ratio is 0.464 in Llandaff whereas the highest is in Splott at 0.529.

Figure 5.14: the spatial pattern in the sex ratio in Cardiff and the Vale of Glamorgan 1990-1994



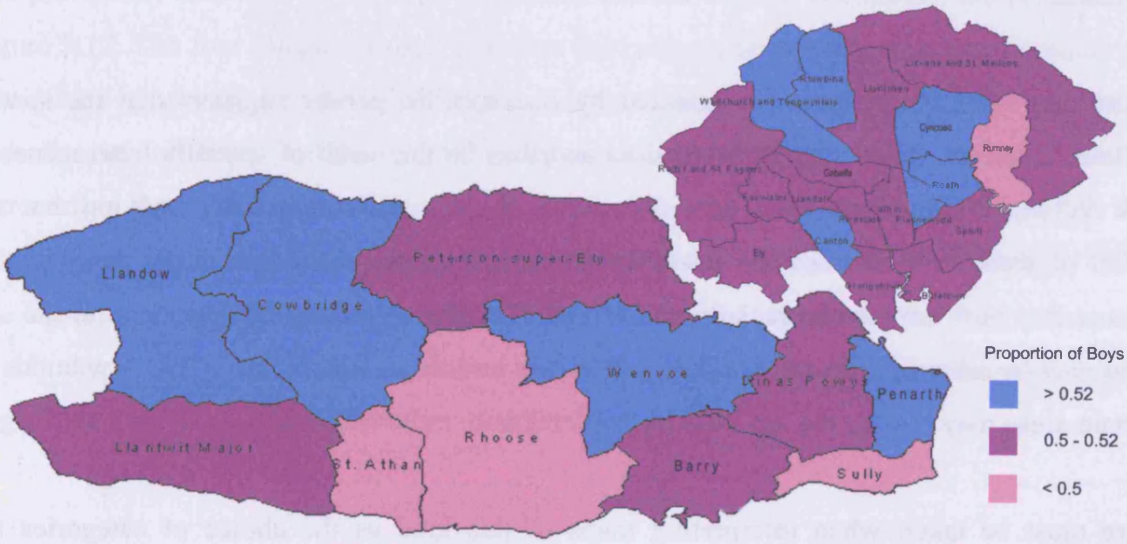
The sex ratios for 1995 to 1999 are mapped in Figure 5.15. In the Vale of Glamorgan there is only a small band of male dominated regions between Peterson-Super-Ely and Llantwit Major. The region with the highest sex ratio is Cowbridge with a sex ratio of 0.533. Llandow in the North West is female dominated and there is a cluster of either neutral or female dominated regions in the South East. The lowest sex ratio is 0.455 in Sully. In Cardiff, a band of female dominated regions runs between Whitchurch and Tongwynlais and Rumney with the lowest sex ratio of 0.482 occurring in Rhiwbina. The South West comprises mainly neutral regions however there are two female dominated regions, Llandaff and Butetown. In addition, there is a small cluster of male dominated regions. The highest sex ratio is 0.530 in Cathays.

Figure 5.15: the spatial pattern in the sex ratio in Cardiff and the Vale of Glamorgan 1995-1999



The map for 2000 to 2004 is given in Figure 5.16. In the Vale of Glamorgan, the North West contains two male dominated regions and to the South East of that there are two female dominated regions. Other than that there are no adjacent regions that are male or female dominated. The highest sex ratio is 0.615 in Wenvoe and the lowest is 0.479 in Sully. The majority of Cardiff is neutral. There is only Rumney which is female dominated with a sex ratio of 0.474. There are two small clusters of male dominated regions: one adjacent to Rumney and one in the North West of Cardiff. The highest sex ratio occurs in Whitchurch and Tongwynlais at 0.553.

Figure 5.16: the spatial pattern in the sex ratio in Cardiff and the Vale of Glamorgan 2000-2004



Overall, clusters of regions that are dominated by one specific gender can occur and the sex ratio in each region varies between five year periods. There seems to be cyclic movements of bands of a specific colour in Cardiff between 1975 and 1989. However, after that the changes in the spatial pattern are more random. Some of these temporal changes in the spatial pattern may be the result of small numbers of births in the council district which is certainly true of the Vale of Glamorgan where there can be as few as 18 births within a district boundary in a five year period.

However, in some of the council districts particularly in Cardiff the number of births is much higher and thus the sex ratios will be more stable. The most prominent of the cyclic change seen in Cardiff occurs in the band of regions between Radyr and St. Fagans and Butetown. This band is initially female dominated, then it becomes male dominated and then returns to being female dominated in 1986-1989. Subsequently, it is a mixture of neutral, male dominated and female dominated regions between 1990 and 1999 then finally in 2000 to 2004 it becomes almost entirely neutral.

Visually, there does appear to be another cyclic change in the North West starting as a male dominated band and becoming neutral, then male dominated for two five year periods, before changing to female dominated and then neutral. However, this pattern may be visually misleading and is likely to be a result of the relatively large size of Lisvane and St. Mellons forming a band in the North West by itself. This visual problem highlights one of the disadvantages of mapping a health outcome by region as opposed to using relative density contours.

Clustering of regions that are dominated by one specific gender suggests that the spatial pattern in the sex ratio may not be random and thus be the result of spatially heterogeneous risk factors. Furthermore, these spatial patterns do change through time implying that the effect of these risk factors also has a temporal component. There are many possible risk factors that may vary both spatially and temporally. For example, the variations in the sex ratio may be influenced by substances used in new buildings, which means that new builds in certain areas may change the sex ratio in that location.

Care must be taken when interpreting maps of this kind as the choice of categories for colouring the maps influences the spatial pattern seen. An additional difficulty faced with

using these maps is that localised spatial patterns that occur within the regions chosen will be overlooked. Mapping sex ratios using kernel density methodology avoids these constraints imposed by using a set of boundaries for regions over which the sex ratio is mapped. Furthermore, mapping the health data by address point means that revisions in the council district boundaries are no longer a problem. In addition, as discussed in Chapter 3, the relative density contours are the initial stages of performing powerful testing procedures for identifying spatial patterns and for investigating spatial association between a health outcome and a specified risk factor. Therefore, relative density contours generated using kernel densities are considered below.

5.3.2 Mapping Sex Ratios using Individual Level Data

Relative density contours, generated using kernel density methodology, require geographical information for each individual that is at least to postcode level. The addresses are available for 1989 onwards so the selected five year periods are 1990 to 1994, 1995 to 1999 and 2000 to 2004. In this chapter, the relative density surface for a given gender is generated by dividing the density of births of that gender by the density of all births. The relative density contours are then drawn to highlight regions over which there is excess density of that gender. There is an alternative option of generating the relative density surface by dividing the density of male births by that of female density or vice-versa and the result is similar, see Figure A5.1. However, this alternative relative density surface cannot be adjusted for confounding factors.

The probability contours for all births in Cardiff and the Vale of Glamorgan are presented in Figure 5.17. The five categories used to define the colouration in the probability contours are determined by quintiles of the height of the density. These indicate that as expected there is a high density of births in the Cardiff area particularly in the council districts directly surrounding the city centre, for example, Ely, Splott, Plasnewydd, Canton, Riverside and Grangetown. The Vale of Glamorgan generally has a lower density of births, although, Barry and Penarth do have a relatively high density. The spatial distribution of births does not hugely vary between each of the five year periods suggesting that migration does not have a large impact on the spatial distribution of births.

Figure 5.17: probability contours for all births in Cardiff and the Vale of Glamorgan a) 1990-1994; b) 1995-1999; c) 2000-2004

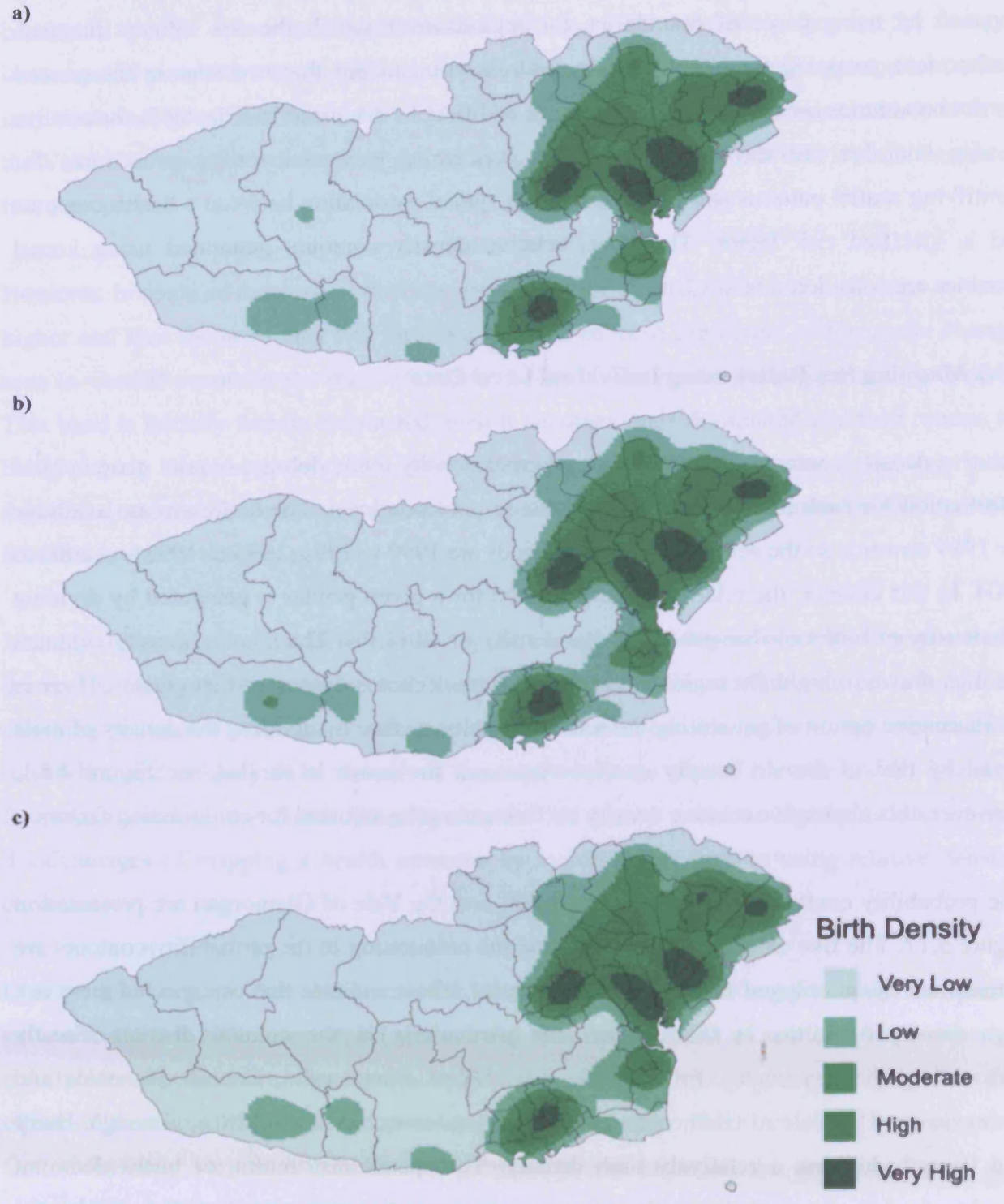
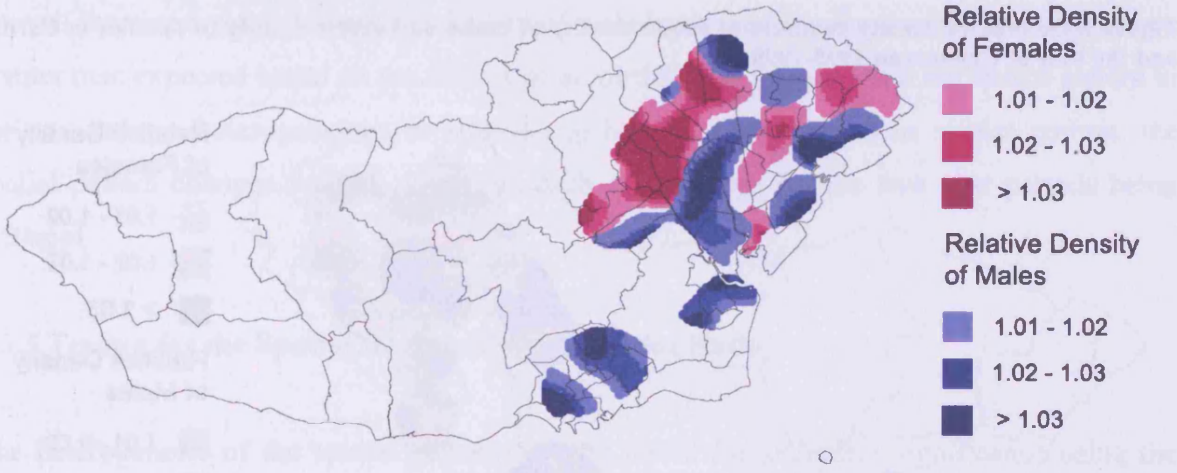


Figure 5.18 reveals the relative density contours indicating regions of excess male births and regions of excess female births in Cardiff and the Vale of Glamorgan in 1990 to 1994. The contours are drawn where the relative density is equal to 1, 1.01, 1.02 and 1.03, which are much lower values than those used for congenital malformations because the effect size is

substantially smaller when considering gender. Contours are only plotted in regions where the population density is high thus reducing problems with elevations in the relative density caused by a low denominator as opposed to a high numerator. Only contours for regions where the population density is at least 20% of the maximum population density are mapped.

Figure 5.18 are the contours for excess density of males and females in Cardiff and the Vale of Glamorgan for 1990 to 1994. In the Vale of Glamorgan, the density of males is greater than expected in Barry and in parts of Penarth. In contrast, Cardiff has both regions where the density of males is greater than expected and regions where the density of females is greater than expected. There are two separate bands of excess female birth density both running in a North East direction: one from Ely to Llanishen and a second from Butetown to the Eastern side of Lisvane and St. Mellons. The main region where there is an elevated density of male births surrounds the second band of excess female density in a horseshoe shape. There is a second region of excess male birth density in the North part of Llanishen.

Figure 5.18: relative density contours of excess density of males and excess density of females in Cardiff and the Vale of Glamorgan 1990-1994



Visual comparison of Figure 5.14 and Figure 5.18 indicates that the spatial patterns in the sex ratio seen in the map created using area level data are mirrored in the contours of excess male and female birth density. For example, in both figures Barry, Grangetown, Riverside, Plasnewydd and Cathays are male dominated regions. Furthermore, Llandaff, Butetown and Whitchurch and Tongwynlais are female dominated regions in both Figures. However, some regions are identified as neutral regions in the aggregated map but in fact have regions of either excess density of male or female births or both within their boundaries. For example,

Penarth is shown to have a sex ratio that is close to that of the overall sex ratio in Figure 5.14. However, from the contours it is clear that there is in fact a large region over which the density of male births is higher than expected. This elevation in the density of male births is not identified when only area level data is used in the mapping procedure. Clearly, the density mapping technique provides a more detailed account of the spatial distribution of health outcomes that is overlooked when using pre-specified boundaries. Consequently, kernel density mapping is a more visually powerful way of mapping a health outcome.

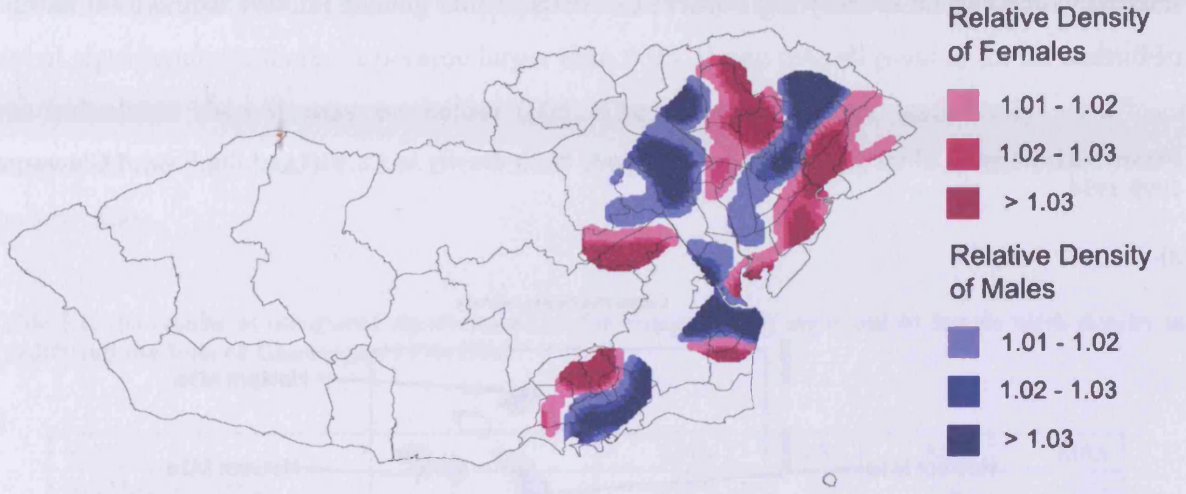
In 1995 to 1999 the picture is different as revealed in Figure 5.19. In the Vale of Glamorgan the East side of Barry still has an excess density of males, however, the West side now exhibits an excess in female density. The parts of Penarth that were previously dominated by excess male density are now predominantly covered by excess female density. In Cardiff, although, there are still two main bands of excess female density these bands have moved: one further to the North East and the other further to the South West. There are two regions of excess male density: one over Ely and Fairwater and the other covers council districts in the centre of Cardiff.

Figure 5.19: relative density contours of excess density of males and excess density of females in Cardiff and the Vale of Glamorgan 1995-1999

Figure 5.20 are the contours of excess male and female birth density in 2000 to 2004. In the Vale of Glamorgan, there is excess male density in the South of Barry and the North of Penarth. Female density is elevated in the North of Barry and the South of Penarth. In Cardiff, there is a band where the female density is elevated running between Grangetown and Lisvane

and St.Mellons. There is also elevated female density in Ely and in a large region over Cyncoed and Llanishen. There is one large area of elevated male density that forms a horseshoe shape running from Rhiwbina to Fairwater, Fairwater to Grangetown and Grangetown to Lisvane and St. Mellons.

Figure 5.20: relative density contours of excess density of males and excess density of females in Cardiff and the Vale of Glamorgan 2000-2004



Overall there are regions where the density of male births and the density of female births are greater than expected based on the density of all births. This suggests that the spatial pattern in the sex ratio is heterogeneous. In addition to having a heterogeneous spatial pattern, the spatial pattern changes through time with each of the maps for the five year periods being different.

5.3.3 Testing for the Spatial Homogeneity of the Sex Ratio

The heterogeneity of the spatial patterns can be tested for statistical significance using the global significance test for spatial homogeneity, see Chapter 3, and if they are statistically significant then there may be some spatially heterogeneous risk factor that influences the sex ratio. The results of the global significance test are presented below.

The hypotheses are

H_0: the spatial patterns are homogeneous

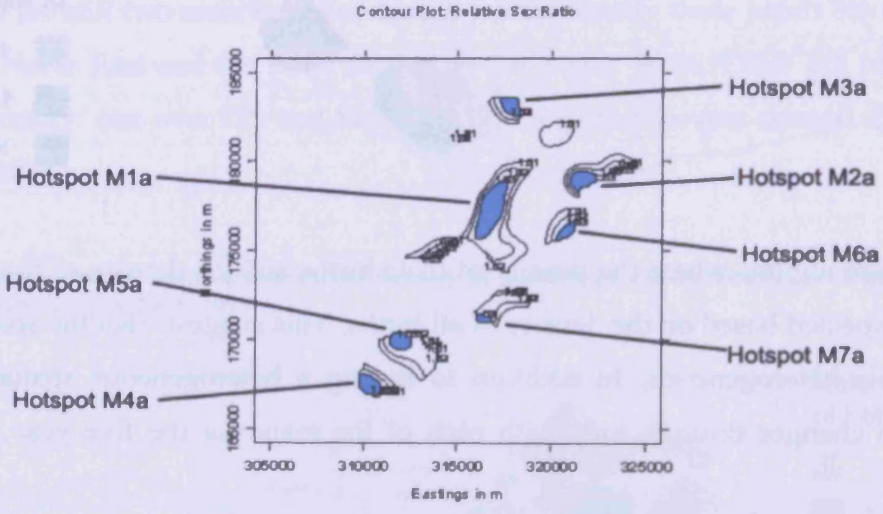
and

H_1: the spatial patterns are heterogeneous.

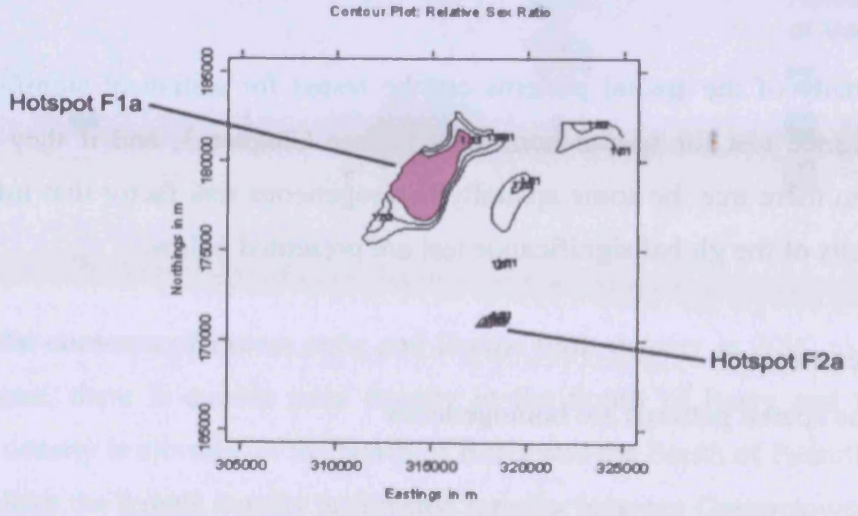
The global significance test assesses the statistical significance of hotspots of excess male birth density or excess female birth density. Hotspots of male births, in this situation, were considered to be regions over which the density of male births divided by the density of all births was greater than 1.03. Similarly, hotspots of female births were defined as regions for which the female birth density over the density of all births was greater than 1.03. The choice of 1.03 to define the hotspots ensured that each hotspot was dominated by the gender of interest which can be seen by the number of births of that gender relative to the total number of births.

Figure 5.21: hotspots of excess a) male and b) female birth density in Cardiff and the Vale of Glamorgan 1990-1994

a)



b)



The hotspots in Cardiff and the Vale of Glamorgan for 1990 to 1994 are given in Figure 5.21 and the results of the global significance test for these hotspots are given in Table 5.1. There are 7 hotspots of excess male density in Cardiff and the Vale in 1990 to 1994 and all of these are statistically significant except for hotspot M6a. However, in contrast there are only 2 hotspots of excess female density for the same period and only 1 of those is statistically significant, hotspot F2a. Hotspot F1a may be statistically significant because only one of the global significance tests has a p-value larger than 0.05. If one overall p-value for all three tests was calculated then it may be below 0.05. The existence of six statistically significant hotspots in the male relative density contours suggests that there is a heterogeneous pattern in the sex ratio.

Table 5.1: the results of the global significance test for hotspots of a) male and b) female birth density in Cardiff and the Vale of Glamorgan 1990-1994

a)

Hotspot		M1a	M2a	M3a	M4a
Number of births		1141	197	243	385
Number of male births		646	344	420	676
Observed mean relative density		1.054	1.047	1.041	1.047
p-value	test 1	<0.001	<0.001	<0.001	<0.001
	test 2	<0.001	0.011	<0.001	0.036
	test 3	<0.001	0.008	0.003	0.001

Hotspot		M5a	M6a	M7a
Number of births		442	23	300
Number of male births		754	44	520
Observed mean relative density		1.068	1.032	1.049
p-value	test 1	<0.001	0.007	<0.001
	test 2	<0.001	0.186	<0.001
	test 3	<0.001	0.388	0.002

b)

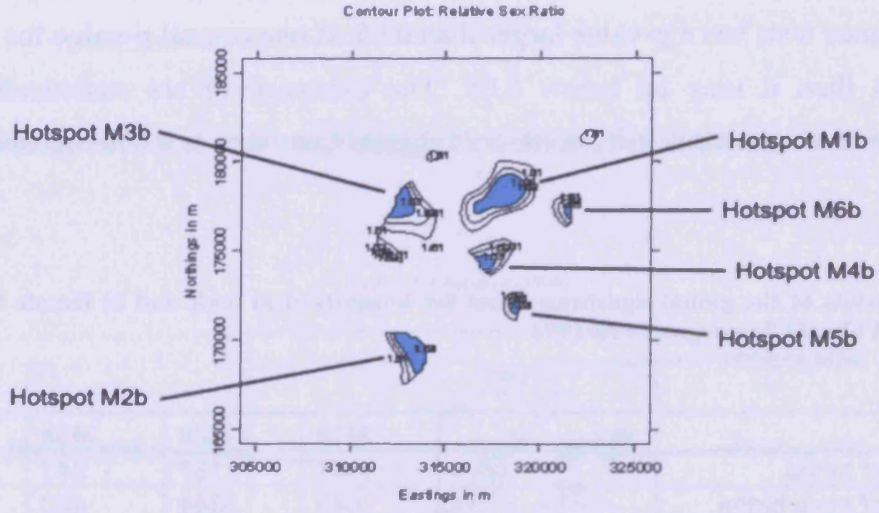
Hotspot		F1a	F2a
Number of births		2735	920
Number of female births		1431	511
Observed mean relative density		1.063	1.113
p-value	test 1	<0.001	<0.001
	test 2	<0.001	<0.001
	test 3	0.111	0.005

The hotspots and global significance results for 1995 to 1999 are presented in Figure 5.22 and Table 5.2. All the hotspots of excess female density are statistically significant. However, two of the hotspots of excess male density may not be statistically significant and these are M4b

and M6b. Therefore, the existence of 6 statistically significant hotspots of excess female density and 4 statistically significant hotspots of excess male density suggest that the spatial pattern in the sex ratio in Cardiff and the Vale is heterogeneous in 1995 to 2000.

Figure 5.22: hotspots of excess a) male and b) female birth density in Cardiff and the Vale of Glamorgan 1995-1999

a)



b)

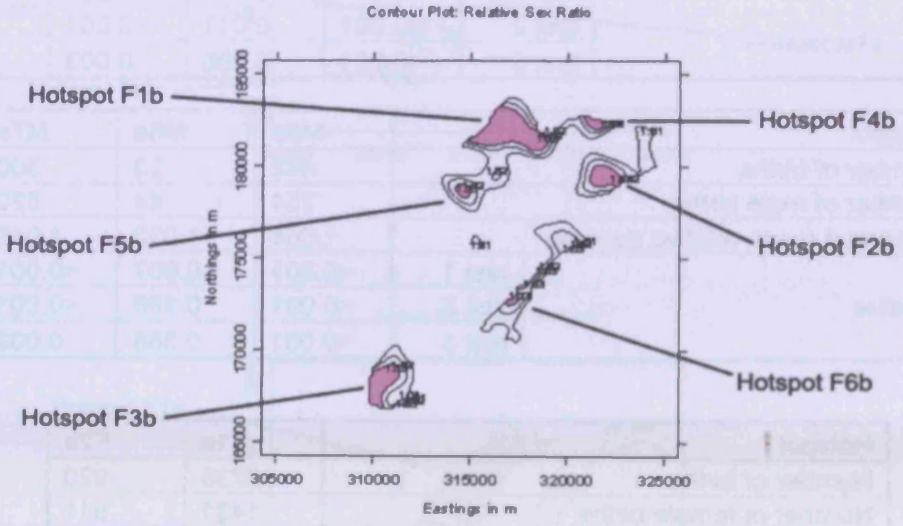


Table 5.2: results of the global significance test for hotspots of a) male and b) female density in Cardiff and the Vale of Glamorgan 1995-1999

a)

Hotspot		M1b	M2b	M3b	M4b
Number of births		840	1173	469	24
Number of male births		459	661	279	14
Observed mean relative density		1.043	1.076	1.078	1.038
p-value	test 1	<0.001	<0.001	<0.001	0.004
	test 2	<0.001	<0.001	<0.001	0.050
	test 3	0.030	0.001	<0.001	0.188

Hotspot		M5b	M6b
Number of births		118	47
Number of male births		70	22
Observed mean relative density		1.077	1.037
p-value	test 1	<0.001	<0.001
	test 2	<0.001	0.137
	test 3	0.021	0.665

b)

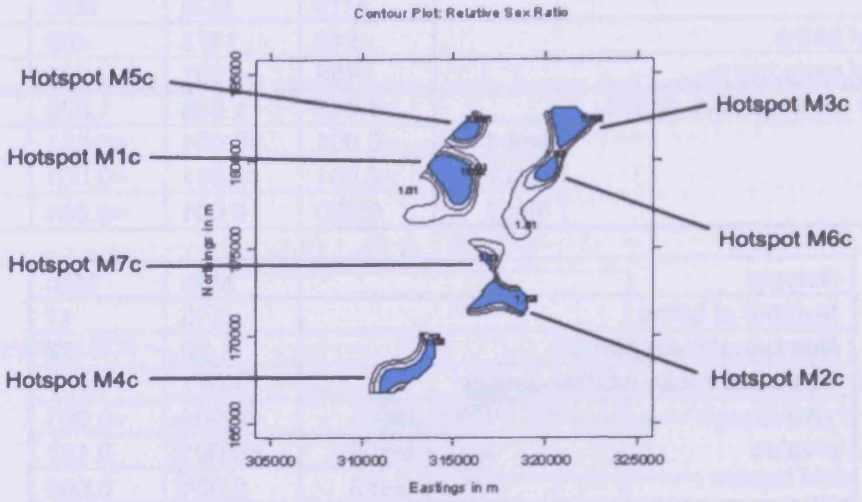
Hotspot		F1b	F2b	F3b	F4b
Number of births		988	370	909	120
Number of female births		532	200	503	70
Observed mean relative density		1.052	1.036	1.052	1.044
p-value	test 1	<0.001	<0.001	<0.001	<0.001
	test 2	<0.001	0.026	0.002	0.021
	test 3	0.001	0.014	<0.001	0.015

Hotspot		F5b	F6b
Number of births		176	205
Number of female births		111	114
Observed mean relative density		1.085	1.050
p-value	test 1	<0.001	<0.001
	test 2	0.003	<0.001
	test 3	<0.001	0.018

The hotspots for Cardiff and the Vale in 2000 to 2004 are given in Figure 5.23 and the results of the global significance test in Table 5.3. All the hotspots of excess male birth density are statistically significant except for hotspots M6c and M7c, both of which have a p-value greater than 0.05 for test 3, an overall p-value for all three tests could be less than 0.05. Hotspot F5c is the only non-statistically significant hotspot of excess female birth density. The existence of 6 statistically significant hotspots for each gender suggests that there is evidence of a heterogeneous spatial pattern in 2000 to 2004.

Figure 5.23: hotspots of excess a) male and b) female birth density in Cardiff and the Vale of Glamorgan 2000-2004

a)



b)

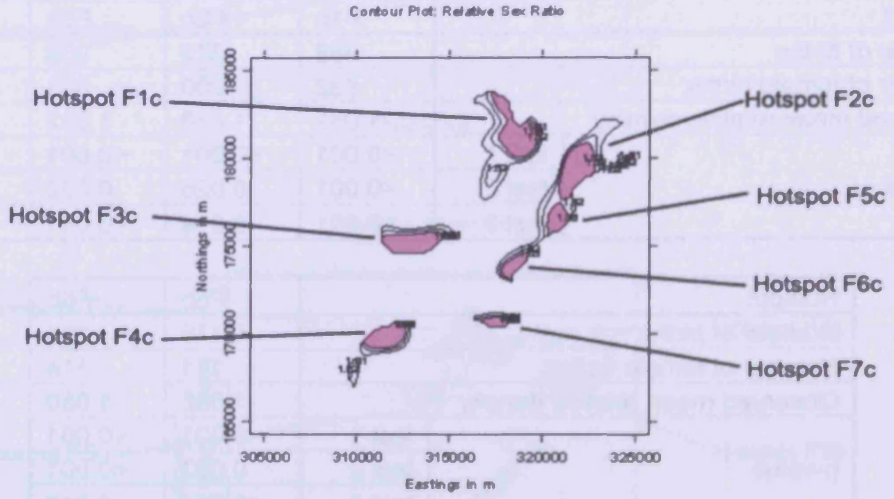


Table 5.3: results of the global significance test for hotspots of a) male and b) female density in Cardiff and the Vale of Glamorgan 2000-2004

a)

Hotspot		M1c	M2c	M3c	M4c
Number of births		720	500	555	602
Number of male births		421	275	305	342
Observed mean relative density		1.079	1.044	1.054	1.052
p-value	test 1	<0.001	<0.001	0.013	<0.001
	test 2	<0.001	<0.001	0.001	<0.001
	test 3	<0.001	0.035	0.034	0.005

Hotspot		M5c	M6c	M7c
Number of births		322	203	30
Number of male births		185	115	18
Observed mean relative density		1.078	1.036	1.060
p-value	test 1	<0.001	0.004	<0.001
	test 2	<0.001	0.004	<0.001
	test 3	0.018	0.059	0.118

b)

Hotspot		F1c	F2c	F3c	F4c
Number of births		531	647	618	875
Number of female births		288	342	356	479
Observed mean relative density		1.050	1.054	1.078	1.079
p-value	test 1	<0.001	<0.001	<0.001	<0.001
	test 2	<0.001	0.002	<0.001	<0.001
	test 3	0.005	0.023	<0.001	0.001

Hotspot		F5c	F6c	F7c
Number of births		298	105	341
Number of female births		148	66	182
Observed mean relative density		1.034	1.038	1.052
p-value	test 1	0.013	<0.001	<0.001
	test 2	0.010	<0.001	<0.001
	test 3	0.371	<0.001	0.032

In conclusion, the results of the global significance test for spatial homogeneity imply that there is a heterogeneous spatial pattern in the sex ratio and thus there is likely to be at least one spatially heterogeneous risk factor that influences the sex ratio. There are a number of possible risk factors for the sex ratio discussed in the literature and information is available on some of these including maternal age, socioeconomic deprivation, maternal ethnicity, maternal BMI and parity. Therefore, these risk factors will be investigated and if there does seem to be a relationship between gender and the specified risk factor then that risk factor will be adjusted for in the contours to see if there is any residual effect.

5.4 Possible Risk Factors

Risk factors that can feasibly be considered using the Cardiff Birth Survey include maternal age, parity, maternal body mass index (BMI), maternal ethnicity and deprivation. These five factors are discussed in detail below. The relationships between each of these factors and the sex ratio are investigated as are the spatial patterns in each of the risk factors. If there are spatial patterns in any of the risk factors and that risk factor influences the sex ratio then that factor may explain the spatial patterns observed in the sex ratio in Cardiff and the Vale of Glamorgan.

5.4.1 Maternal Age

Information is available on maternal age for all years in the Cardiff Birth Survey. The sex ratio for Cardiff and the Vale is given by maternal age group in Figure 5.24 and Table 5.4 which also contains the 95% confidence intervals for the sex ratio. The proportion of boys seems to be slightly higher for mothers under 20 years old and also for those 40 years and over, whereas, mothers between 20 and 39 years of age seem to be less likely to have a male birth. The association between maternal age group and gender is statistically significant (χ^2 test: p-value = 0.028).

Figure 5.24: the sex ratio by maternal age in Cardiff and the Vale of Glamorgan 1965-2004

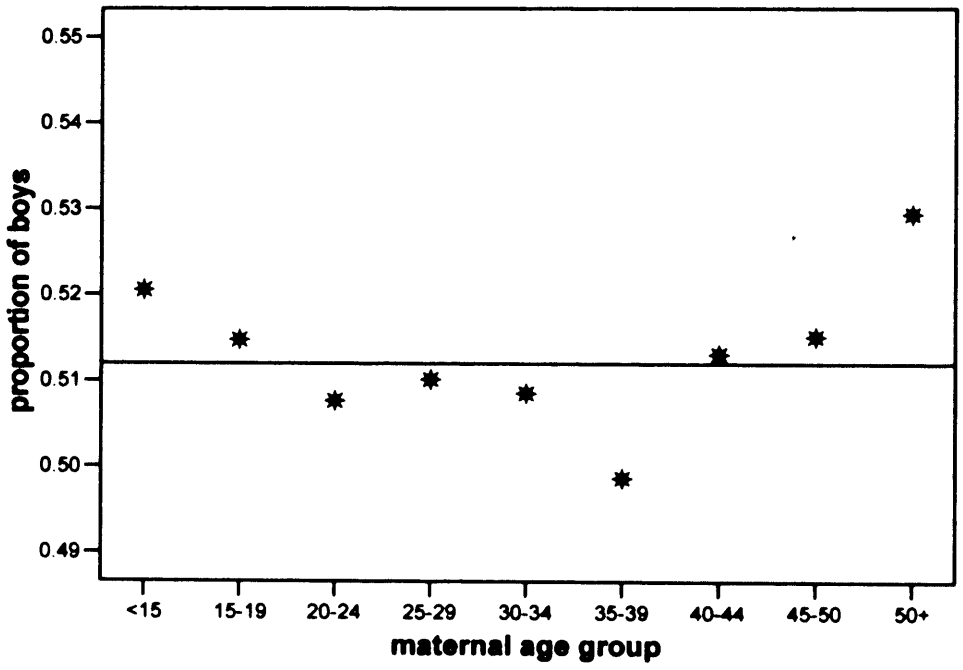


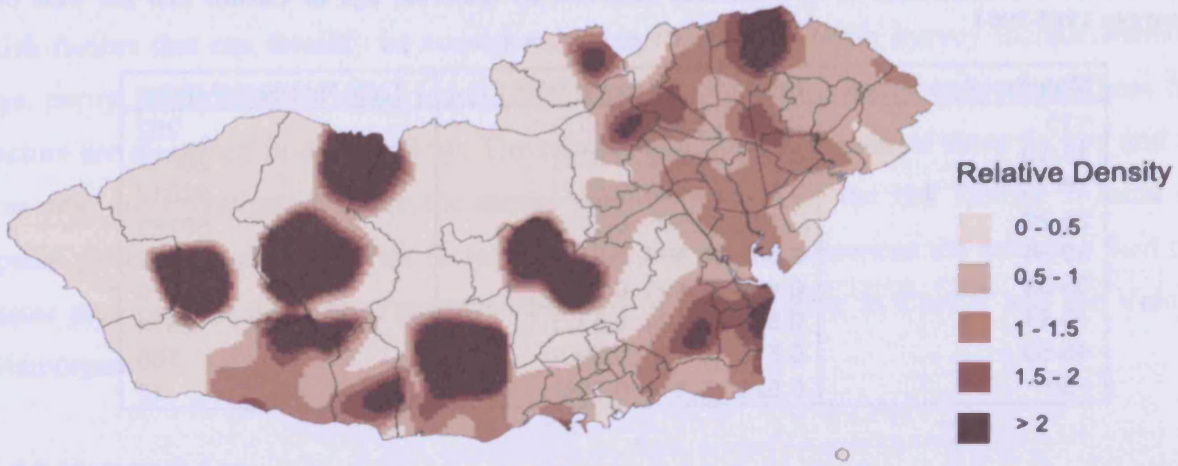
Table 5.4: the sex ratio with 95% confidence intervals by maternal age in Cardiff and the Vale of Glamorgan 1965-2004

Maternal Age Group	Proportion of Boys (95% CI)	Total Number of Births
<15	0.521 (0.463, 0.577)	292
15-19	0.515 (0.508, 0.522)	20052
20-24	0.508 (0.503, 0.512)	52817
25-29	0.510 (0.506, 0.514)	58283
30-34	0.509 (0.504, 0.514)	37442
35-39	0.499 (0.491, 0.507)	14276
40-44	0.513 (0.494, 0.532)	2674
45-50	0.515 (0.439, 0.590)	165
>50	0.529 (0.310, 0.738)	17

The sex ratio for Wales and the USA [283] is also given by maternal age group in Figures A5.2 and A5.3, respectively. These figures reveal that the influence of maternal age on the sex ratio varies according to which population is considered. However, there are other risk factors for gender that may interact with maternal age and therefore explain these variations between populations. For example, the USA has a very different ethnic mix to Wales and ethnicity may influence both the probability of a specific gender and the age at which a mother is most likely to give birth. The patterns for Cardiff and the Vale of Glamorgan and for Wales are fairly similar and the slight differences that have occurred may be down to the relatively small numbers of mothers in the older age groups.

The spatial variation of maternal age is shown in Figure 5.25, which are relative density contours plotted where the relative density of having children after either 40 years of age is equal to 0.5, 1, 1.5 and 2 in 2000 to 2004. Regions where the relative density of maternal age exceeds 2 suggest that there is a high risk that the mother will be over that age when she gives birth therefore the region generally has a higher maternal age than expected. The maternal age is generally higher in the Vale of Glamorgan and parts of North Cardiff such as Lisvane and St. Mellons, Llanishen and Cyncoed. The equivalent plots for all three time periods and for over 30, 35 and 40 years of age are given in Figures A5.4 to A5.6. These regions of generally higher maternal age vary slightly by year but not hugely.

Figure 5.25: relative density contours of births to mothers over 40 years of age in Cardiff and the Vale of Glamorgan in 2000-2004



The regions where the maternal age is generally higher than expected are less economically deprived areas where mothers are more likely develop their career before starting a family. The relationship between socioeconomic deprivation and maternal age is shown in Figure A5.27 where it is clear that the maternal age is higher for those mothers with lower levels of socioeconomic deprivation.

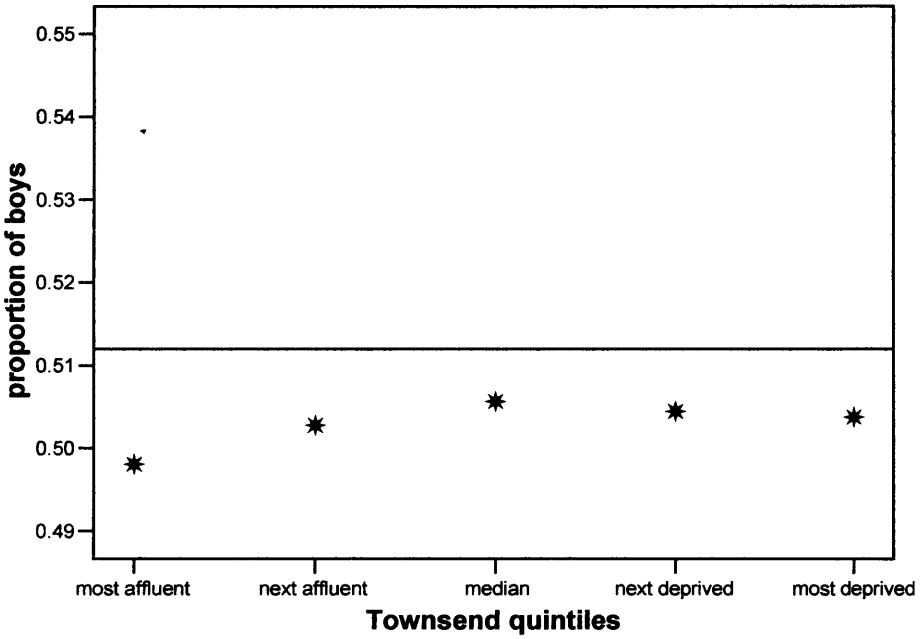
5.4.2 Socioeconomic Deprivation

Socioeconomic deprivation is determined using the Townsend Score which is derived from the percentage of private households with no car, the percentage of private households not in owner-occupied accommodation, the percentage of private households with more than one person per room and unemployed persons as a percentage of the economically active. A high Townsend score indicates a higher level of socioeconomic deprivation. The Townsend score is used to split the population into five categories of socioeconomic deprivation referred to here as Townsend quintiles: most deprived, next deprived, median, next affluent and most affluent.

The Townsend score for each mother was allocated using the postcode of the residence of the mother at the time of birth by Health Solutions Wales. The need for the postcode in the allocation process for the Townsend score consequently means that the Townsend score is only available for 1989 onwards. The sex ratio for each of the Townsend quintiles is displayed in Figure 5.26. The sex ratio generally increases with increasing levels of deprivation although

the sex ratio is fairly similar in the median, next deprived and most deprived groups. The sex ratios for each of the Townsend quintiles are below the line for the overall sex ratio because only those mothers that gave birth from 1989 onwards are included when the sex ratio was lower than in previous years.

Figure 5.26: the sex ratio by socioeconomic deprivation in Cardiff and the Vale of Glamorgan 1989-2004



The 95% confidence intervals for the sex ratio in each of the Townsend quintiles are given in Table 5.5. All the confidence intervals for each of the Townsend quintiles have considerable overlap which could be the result of small sample sizes together with the small effect size. The association between gender and Townsend quintile is not statistically significant (χ^2 test: p-value = 0.760).

Table 5.5: the sex ratio with 95% confidence intervals by Townsend quintile in Cardiff and the Vale of Glamorgan 1989-2004

Townsend Quintiles	Proportion of Boys (95% CI)	Total Number of Births
Most Affluent	0.498 (0.490, 0.506)	13743
Next Affluent	0.503 (0.495, 0.511)	13861
Median	0.506 (0.498, 0.514)	13800
Next Deprived	0.504 (0.496, 0.513)	13357
Most Deprived	0.504 (0.496, 0.512)	13838

Although there are no postcodes available before 1989 and consequently no Townsend scores, the Cardiff Birth Survey did record a social class variable for 1965 to 1986. This social class variable is determined by the occupation of the mother placing the mother into one of a number of categories. Some of these categories are given in Figure 5.27 other categories included single parent, housewife, Her Majesties Services and students. However, these were omitted from the plot because of their ambiguous nature. Figure 5.27 reveals a conflicting result to Figure 5.26 with mothers in the professional category, most likely to coincide with the most affluent Townsend quintile, having the highest sex ratio. Moreover, the group most likely to coincide with the most deprived Townsend quintile, unskilled workers, has the lowest sex ratio.

Figure 5.27: the sex ratio by social class in Cardiff and the Vale of Glamorgan 1965-1986

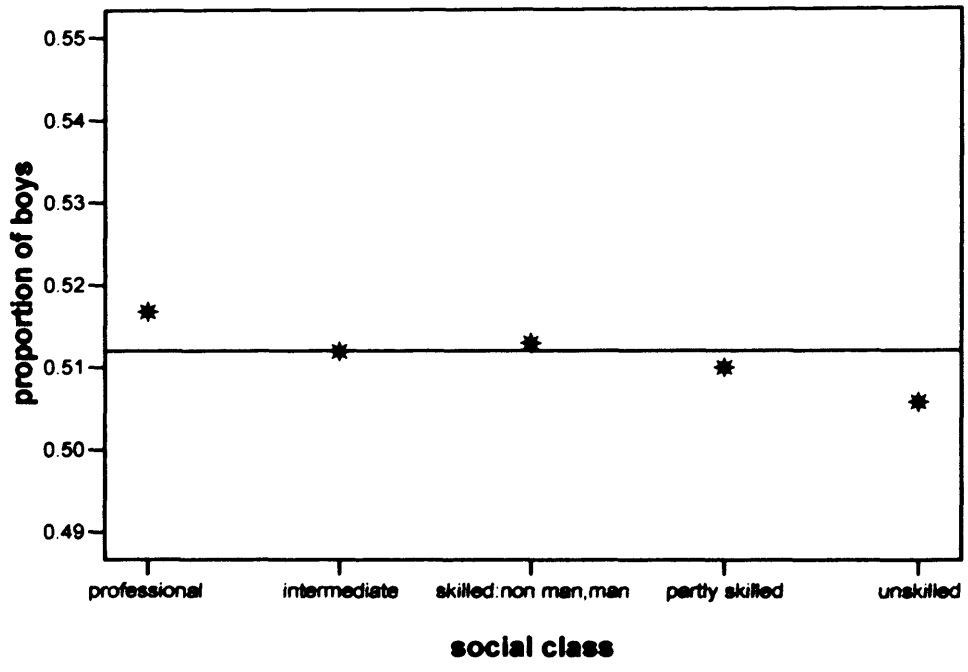


Table 5.6: the sex ratio with 95% confidence intervals by social class in Cardiff and the Vale of Glamorgan 1965-1986

Social Class	Proportion of Boys (95% CI)	Total Number of Births
Professional	0.517 (0.508, 0.526)	12224
Intermediate	0.512 (0.491, 0.534)	2087
Skilled	0.513 (0.509, 0.517)	64019
Partly Skilled	0.510 (0.504, 0.516)	30382
Unskilled	0.506 (0.498, 0.514)	14516

The disparity between the impact of social class and Townsend score on the sex ratio could be explained by the way in which these measures of socioeconomic deprivation are determined.

Townsend score is determined by factors such as car or home ownership or overcrowding which is influenced by the father's socioeconomic status as well as the mother's. However, the variable referred to as social class is determined only by the mother's employment status. Therefore, a mother who is unskilled could have a low Townsend score because the father is a professional. Furthermore, Townsend score is measured at the postcode level whereas social class is determined at the individual level therefore it is possible that a professional mother could have a high Townsend score because the majority of people in the area have a high Townsend score.

In addition to this social class variable, the employment status of the mother is recorded at the time that they first present themselves. The sex ratio for each of the two employment statuses are presented in Table 5.7. Table 5.7 reveals that the sex ratio is higher amongst employed mothers, which is expected based on the increase of the sex ratio with increased social class. However, it should be noted that some of the mothers are both classed as professionals and are down as unemployed possibly because they chose to take a career break to have children. Therefore, the employment status, though in many cases good, may on occasion be a poor indicator of the level of socioeconomic deprivation.

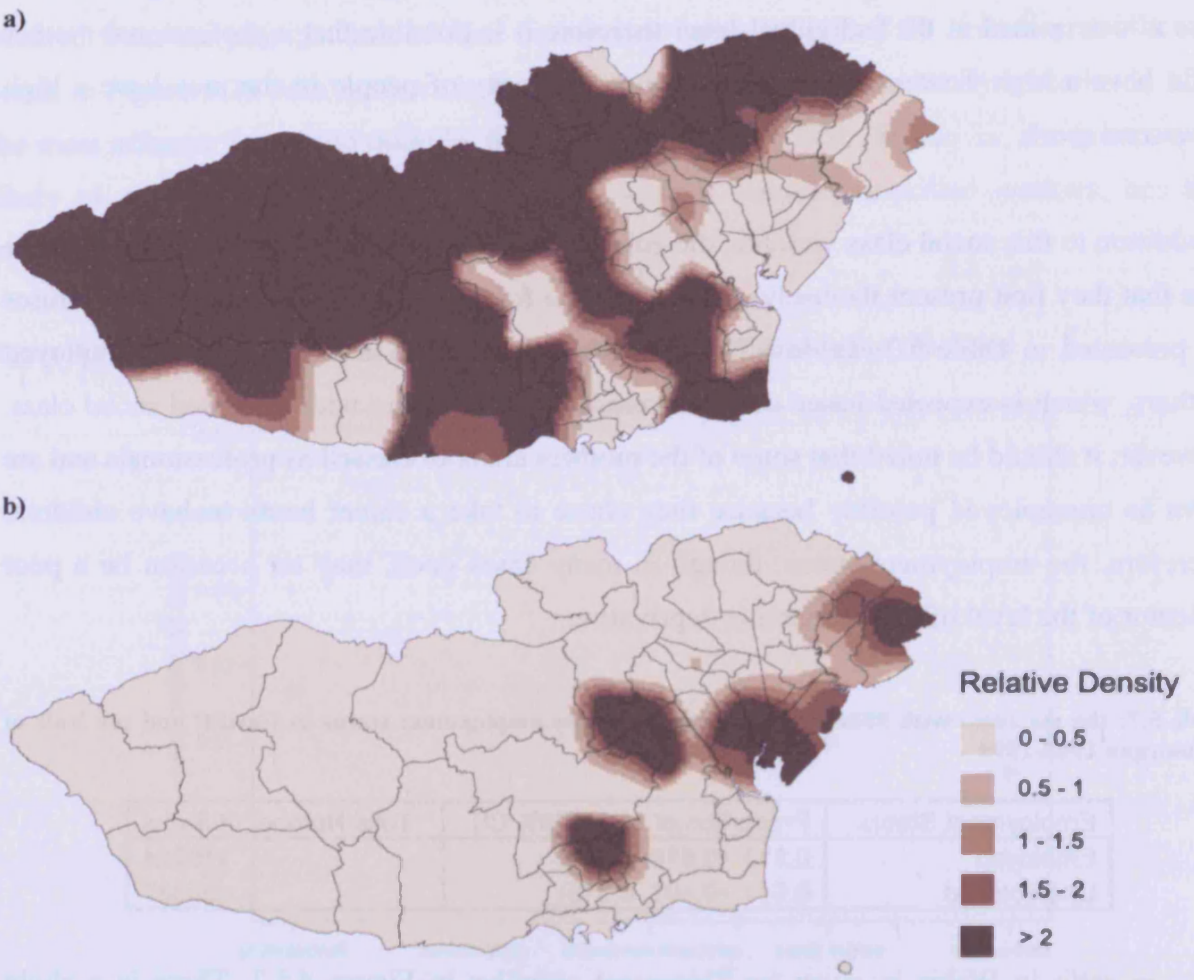
Table 5.7: the sex ratio with 95% confidence intervals by employment status in Cardiff and the Vale of Glamorgan 1965-1999

Employment Status	Proportion of Boys (95% CI)	Total Number of Births
Employed	0.513 (0.510, 0.516)	116314
Unemployed	0.511 (0.507, 0.515)	55887

The sex ratio in Wales is given by Townsend quintiles in Figure A5.7. There is a slight disparity in the pattern seen in sex ratio by Townsend quintile between the CHS and the Cardiff Birth Survey. However, despite the slight variation it is clear from the CHS data that those mothers in the most deprived Townsend quintile have the highest sex ratio whereas those with greater levels of affluence have lower sex ratios. This supports the result obtained from the Cardiff Birth Survey. However, it should be noted that this result conflicts with the Trivers-Willard hypothesis that mothers in a good condition are more likely to produce male offspring. The reason for this disparity may lie in the relationship between Townsend score and Body Mass Index (BMI) discussed later on. Figure A5.28 clearly shows that the BMI is lower amongst mothers with lower levels of deprivation and some studies have suggested that underweight mothers are less likely to have a male birth. Thus mothers in the most affluent

group could have a lower sex ratio because of the link between BMI and the sex ratio and not as a result of the relationship between the sex ratio and Townsend score.

Figure 5.28: relative density contours of births to a) the most affluent and b) the most deprived families in Cardiff and the Vale of Glamorgan in 2000-2004



The relative density contours for the location of families in the most deprived and most affluent Townsend quintiles are given in Figure 5.28 for 2000 to 2004. There is a higher density of families that fall in the most deprived Townsend quintile than expected in Ely, the North part of Barry, Grangetown, Butetown, Riverside, Cathays, Splott and the East side of Lisvane and St. Mellons. Families in the median Townsend quintile are most densely located in the South part of Barry, Canton, Riverside and Gabalfa. Finally, the regions of highest density of families that fall into the most affluent Townsend quintile are located in the Vale of Glamorgan, Radyr and St. Fagans, Whitchurch and Tongwynlais, Rhiwbina, Llanishen and the North part of Lisvane and St. Mellons. These patterns reflect that in general the most affluent families live in the more rural regions, whereas the most deprived families are more

likely to reside in the city centre and other built up areas. The regions with a high density of affluent families closely overlap those with a higher density of mothers with high maternal ages as expected. The density contours for all the Townsend quintiles and all three time periods are given in Figures A5.8 to A5.12.

5.4.3 Maternal Ethnicity

The ethnic origin of the mothers is recorded from 1989 onwards. The main ethnic groups reported include White, Black, Indian, Pakistani, Bangladeshi and Chinese. Over 3% of the mothers are recorded as others, many of which identified themselves as British with no specification of ethnicity, and over 6% do not have a recorded ethnic group. Mothers in the others or not known categories were omitted from the analysis. Figure 5.29 displays the sex ratio by maternal ethnicity and there does seem to be some variation in the sex ratio between ethnic groups. The sex ratio seems to be higher in Black, Indian, Pakistani and Chinese mothers and lowest in Bangladeshi mothers.

Figure 5.29: the sex ratio by maternal ethnicity in Cardiff and the Vale of Glamorgan 1989-2004

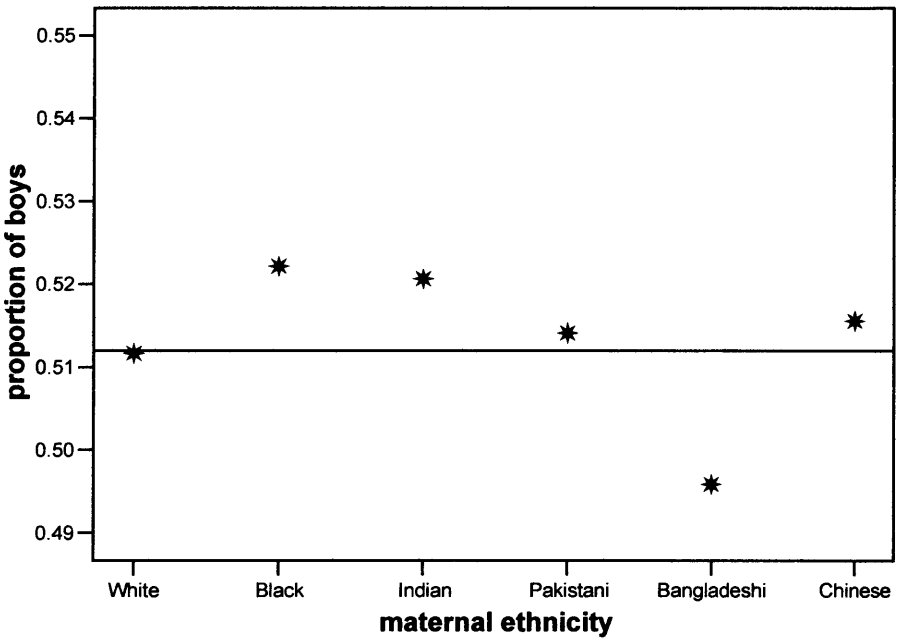


Table 5.8 indicates that there is overlap in the confidence intervals for the proportion of boys between ethnic groups which may be down to small sample sizes. In addition to small sample sizes, Khoury et al. claim that it is the paternal and not maternal ethnicity that influences the sex ratio, in which case any effect that is visible may only be a reflection of women generally

taking a partner of the same ethnicity [154]. In some cases women have partners of a different ethnic group which would dilute the ability to detect the influence of the paternal ethnicity. The association between ethnicity and gender is not statistically significant (χ^2 test: p-value = 0.816).

Table 5.8: the sex ratio with 95% confidence intervals by maternal ethnicity in Cardiff and the Vale of Glamorgan 1989-2004

Maternal Ethnicity	Proportion of Boys (95% CI)	Total Number of Births
White	0.512 (0.508, 0.516)	70429
Black	0.522 (0.497, 0.548)	1488
Indian	0.521 (0.489, 0.552)	968
Pakistani	0.514 (0.490, 0.539)	1595
Bangladeshi	0.496 (0.468, 0.524)	1206
Chinese	0.516 (0.464, 0.567)	353

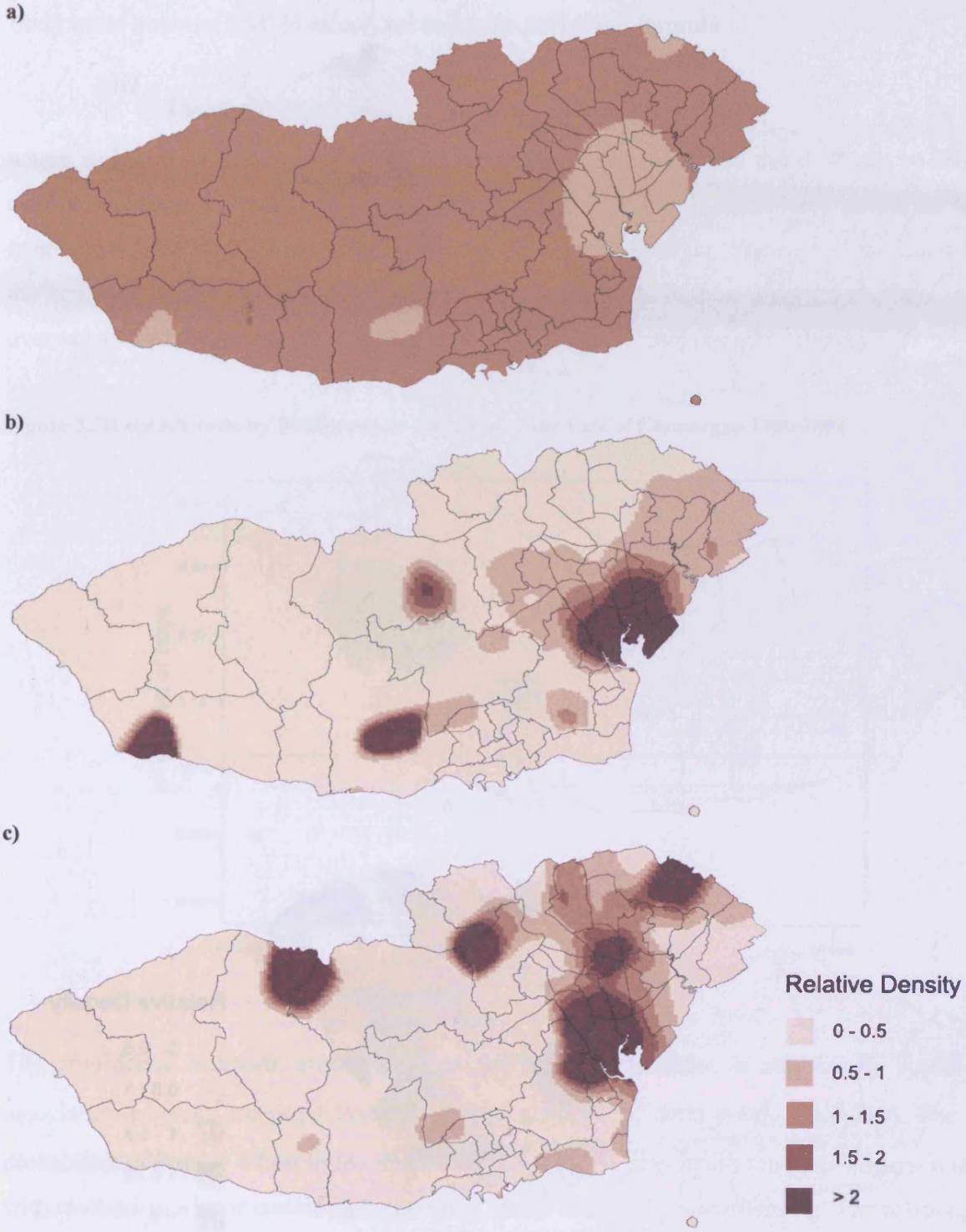
Figure 5.30 displays the relative density contours for the location of each of the named maternal ethnic groups for 2000 to 2004. White mothers are uniformly distributed across the whole of the Vale of Glamorgan and Cardiff. However, Black, Indian, Pakistani and Bangladeshi mothers are most densely located near the city centre in Grangetown, Butetown, Canton, Riverside, Cathays, Plasnewydd and Splott. On the other hand, Chinese mothers are not only densely located in the city centre but there are regions of excess density of Chinese mothers in Penarth, Lisvane and St. Mellons and Llanishen suggesting a greater level of integration. The equivalent relative density contours for all three time periods are given in Figures A5.13 to A5.18.

Figure 5.30: relative density contours of births to a) White, b) Black, c) Indian, d) Pakistani, e) Bangladeshi and f) Chinese mothers in Cardiff and the Vale of Glamorgan 2000-2004

a)

b)

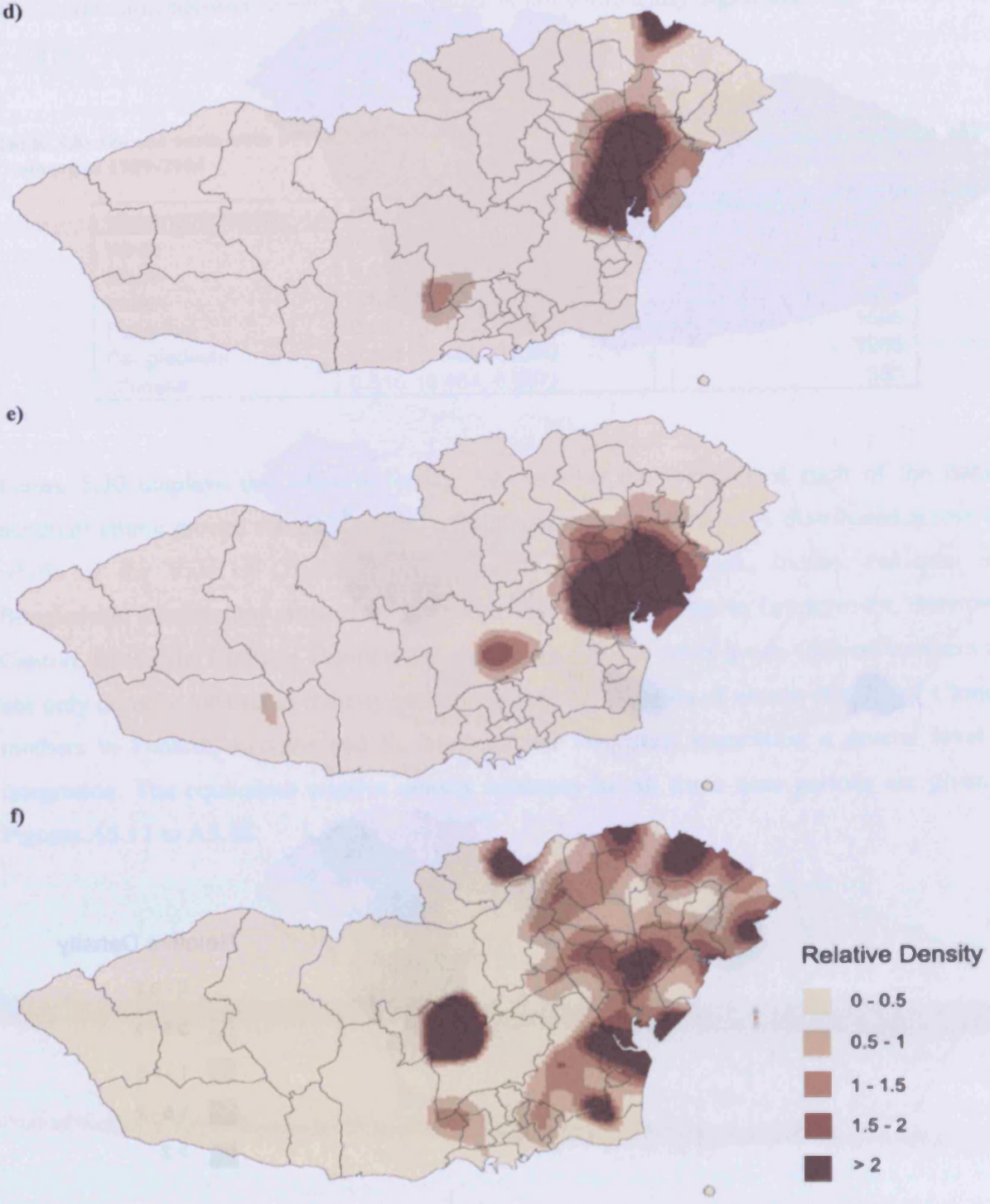
c)



d)

e)

f)



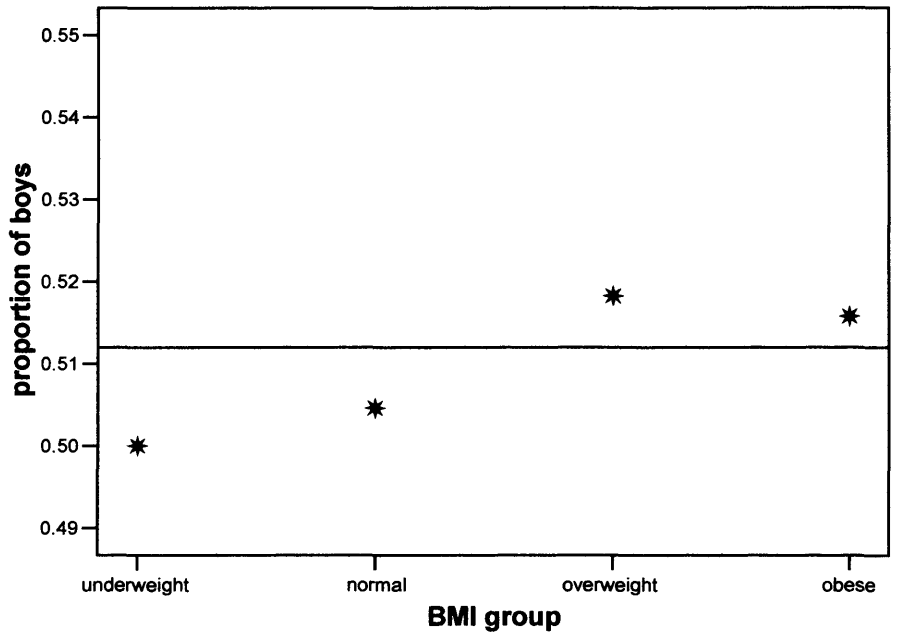
5.4.4 Maternal BMI

Body mass index or BMI is calculated using the following formula

$$BMI = \frac{weight}{(height)^2}$$

where weight is in kilograms and height in metres. Once calculated the BMI can be used to categorise people in terms of obesity. A BMI of 30 and over is considered obese, 25 to 30 overweight, 18.5 to 25 normal and less than 18.5 is underweight. Figure 5.31 indicates how the sex ratio varies between BMI groups. The sex ratio is highest amongst mothers in the overweight and obese categories whereas it is lowest in the underweight category.

Figure 5.31: the sex ratio by BMI group in Cardiff and the Vale of Glamorgan 1989-2004



The confidence intervals given in Table 5.9 overlap but there is statistically significant association between maternal BMI group and gender (χ^2 test: p-value = 0.014). The low probability of having a boy in the underweight condition may again relate to Trivers-Willard with mothers in a poor condition being more likely to have female offspring. The relationship between the sex ratio and BMI may explain the decreasing sex ratio with increased affluence, since affluent women tend to have a lower BMI than more economically deprived women.

Table 5.9: the sex ratio with 95% confidence intervals by maternal BMI in Cardiff and the Vale of Glamorgan 1989-2004

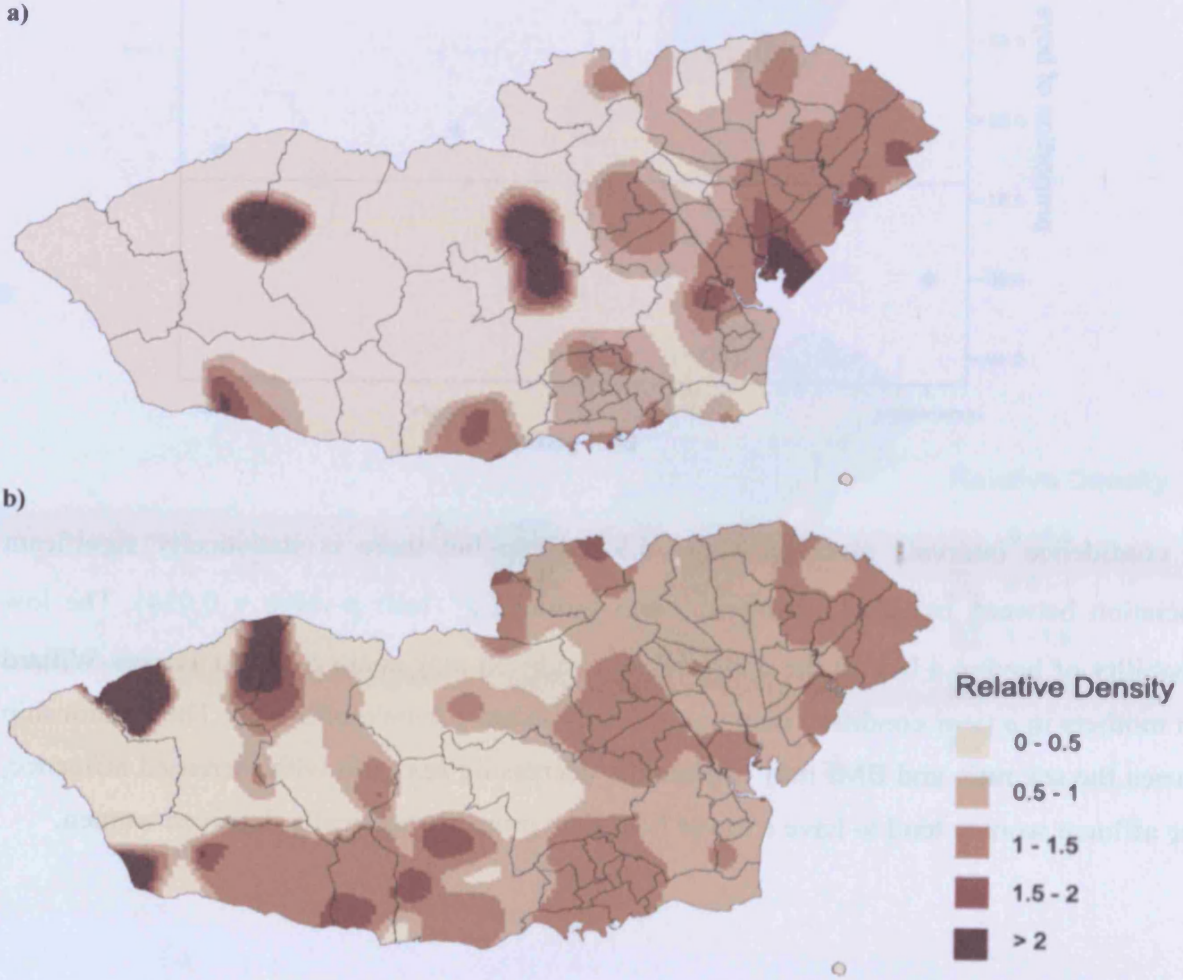
Maternal BMI Group	Proportion of Boys (95% CI)	Total Number of Births
Underweight	0.504 (0.481, 0.527)	1754
Normal	0.507 (0.502, 0.512)	39129
Overweight	0.517 (0.510, 0.524)	17458
Obese	0.518 (0.508, 0.528)	8920

The relative density contours for the location of underweight and, separately, obese mothers are given in Figure 5.32. In general, there are no noticeable regions with excess density of underweight or obese mothers where the population is of sufficient size. However, in 2000 to 2004 there is a relatively high density of underweight mothers in the Butetown area suggesting that the distribution of mothers in specific BMI groups is not always uniform. The relative density contours for all three time periods are given in Figures A5.19 and A5.20.

Figure 5.32: relative density contours of births to a) underweight and b) obese mothers in Cardiff and the Vale of Glamorgan 2000-2004

a)

b)



5.4.5 Parity

Parity is the number of previous live and still births that a mother has had before the current pregnancy. The relationship between the sex ratio and parity is presented in Figure 5.33. There does seem to be a variation in the sex ratio according to parity. The sex ratio for mothers with parity less than 5 seems to be generally lower than for mothers of parity 5 and over. The confidence intervals for the sex ratio by parity overlap substantially. However, there is statistically significant association between parity and gender (χ^2 test: p-value = 0.032).

Figure 5.33: the sex ratio by parity in Cardiff and the Vale of Glamorgan 1989-2004

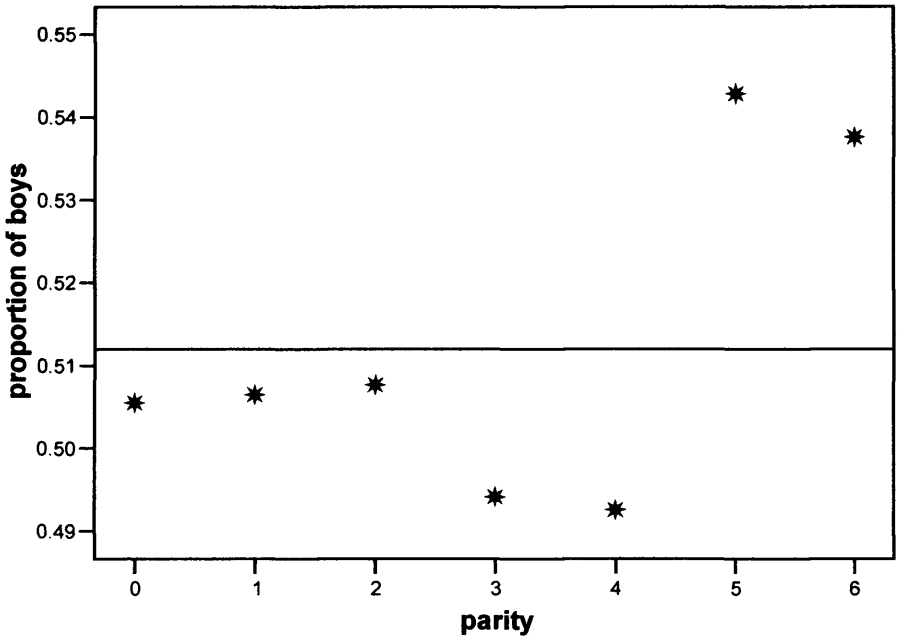


Table 5.10: the sex ratio with 95% confidence intervals by parity in Cardiff and the Vale of Glamorgan 1989-2004

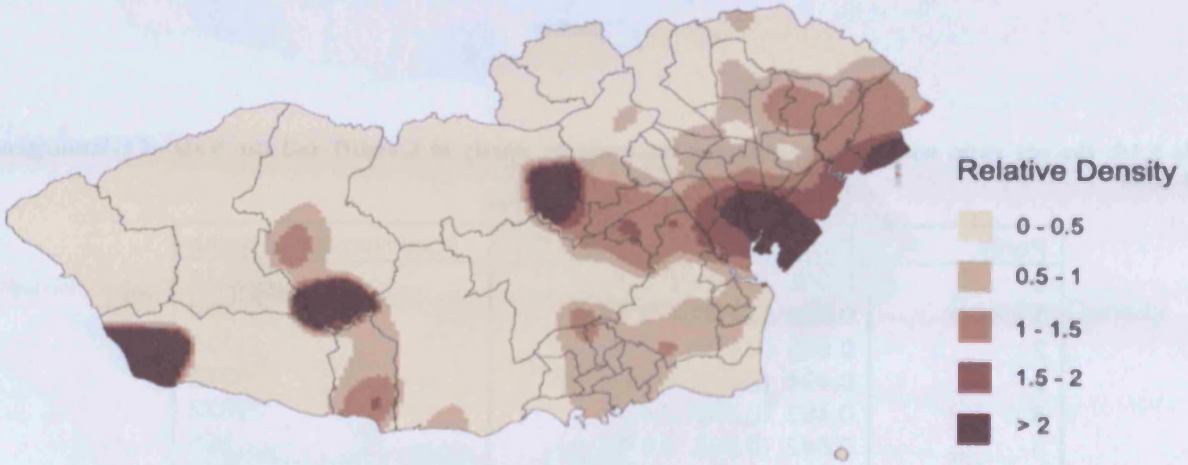
Parity	Proportion of Boys (95% CI)	Total Number of Births
0	0.506 (0.501, 0.511)	33376
1	0.506 (0.500, 0.512)	23870
2	0.508 (0.499, 0.517)	11135
3	0.494 (0.479, 0.509)	4276
4	0.493 (0.469, 0.518)	1622
5	0.542 (0.503, 0.579)	665
6	0.538 (0.478, 0.597)	266
7	0.473 (0.390, 0.558)	131
>= 8	0.578 (0.482, 0.670)	102

The sex ratio by parity is given for Wales in Figure A5.21 and for the US [283] in Figure A5.22. The pattern revealed for Wales is similar to that from the Cardiff Birth Survey with higher parities generally having a higher sex ratio. In the USA, the sex ratio generally decreases with increasing parity the reason for this apparent reverse effect is unclear but it may again relate to the different ethnic mix of the USA and the small number of births with high parity.

One explanation of the increase in the sex ratio with increasing parity relates to the characteristics of the Y-bearing sperm. Y-bearing sperm are smaller and more motile than X-bearing sperm. Debris is left in the maternal sex organs after past pregnancies. The more past pregnancies that a mother has had the greater the amount of debris there is. The build up of this debris may reduce the movement of sperm through the maternal reproductive system. The X-bearing sperm may be affected to a greater degree by this reduced movement because they are less motile. Consequently, fewer girls would be born to mothers with a higher parity. However, this possible mechanism is not supported by the data from the USA.

Figure 5.34 reveals that mothers with a parity of 5 or more are not uniformly distributed spatially. There is a higher density of mothers with a high parity in Grangetown, Butetown, Splott, Rumney and the East part of Lisvane and St. Mellons. The equivalent relative density contours for all three time periods are given in Figure A5.23.

Figure 5.34: relative density contours of births with a parity of 5 or more in Cardiff and the Vale of Glamorgan 2000-2004



5.4.6 Summary

Five possible risk factors were considered: maternal age, socioeconomic deprivation, maternal ethnicity, maternal BMI and parity. The sex ratio does vary by each of these risk factors but the differences in the sex ratio between groups are not always statistically significant. However, the effect sizes are very small and consequently large sample sizes would be required to achieve statistical significance. Therefore, a larger dataset would be required than that of the Cardiff Birth Survey and such a dataset that contains all the desired information was not available. Hence, it is not possible to rule out a relationship between the sex ratio and any of these factors.

The relative density contours for all five of the risk factors given above and in Appendix 5 reveal spatial patterns across Cardiff and the Vale, which means that if they did influence the sex ratio they may explain, to some extent, the spatial patterns in the sex ratio over the same region. Clearly, not all possible risk factors have been investigated and there still may be some that influence the spatial patterns in the sex ratio. The existence of any unknown risk factors can be assessed by adjusting the relative density contours for excess density of males or excess density of females by these risk factors. If the resultant contours display hotspots of excess density that are statistically significant then there are other risk factors that influence the spatial pattern in the sex ratio that have yet to be considered. If there were no statistically significant hotspots then it is likely that all possible risk factors have been adjusted for and consequently the spatial pattern in the sex ratio can be fully explained by the spatial pattern associated with the confounders.

5.5 Spatial Patterns in the Sex Ratio with Adjustment for Confounding

The next step was to adjust the contours for confounding. The reason for performing the adjustment is to assess whether or not investigation into other possible risk factors is warranted.

5.5.1 Mapping the Sex Ratio with Adjustment for Confounding

Maternal age, socioeconomic deprivation, maternal ethnicity, maternal BMI and parity are used to adjust the relative density contours of excess density of males and that of females. The

risk factors are adjusted for by fitting a logistic regression model where gender is the outcome and maternal age, Townsend score, maternal ethnicity, maternal BMI and parity are the independent variables. The resultant predicted probabilities are then used as weights to adjust the density for the location of all births to produce the expected density for male births and that of female births shown in Figure 5.35 and Figure 5.36, respectively.

Figure 5.35: probability contours for the expected density of male births in Cardiff and the Vale of Glamorgan a) 1990-1994; b) 1995-1999; c) 2000-2004

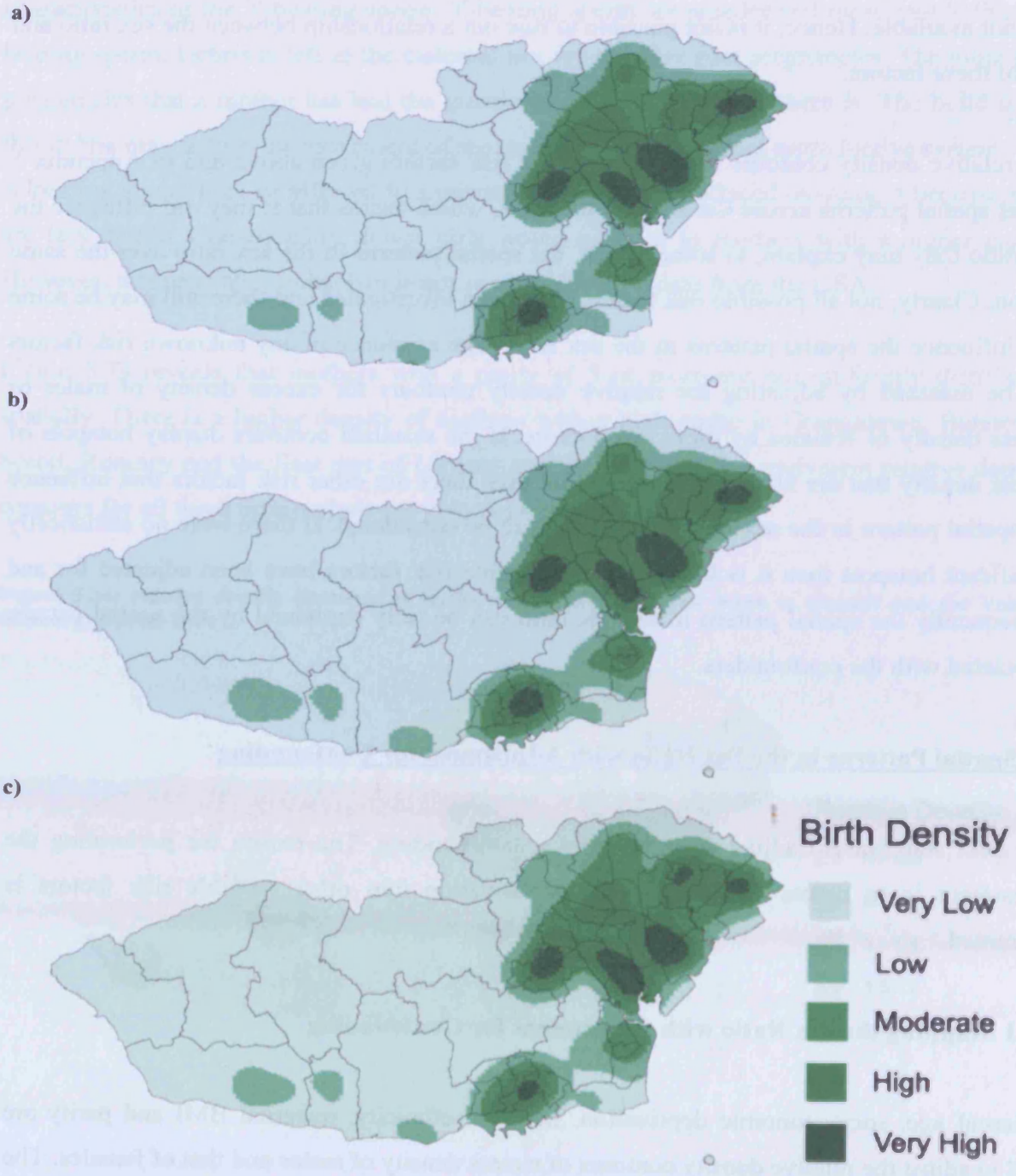
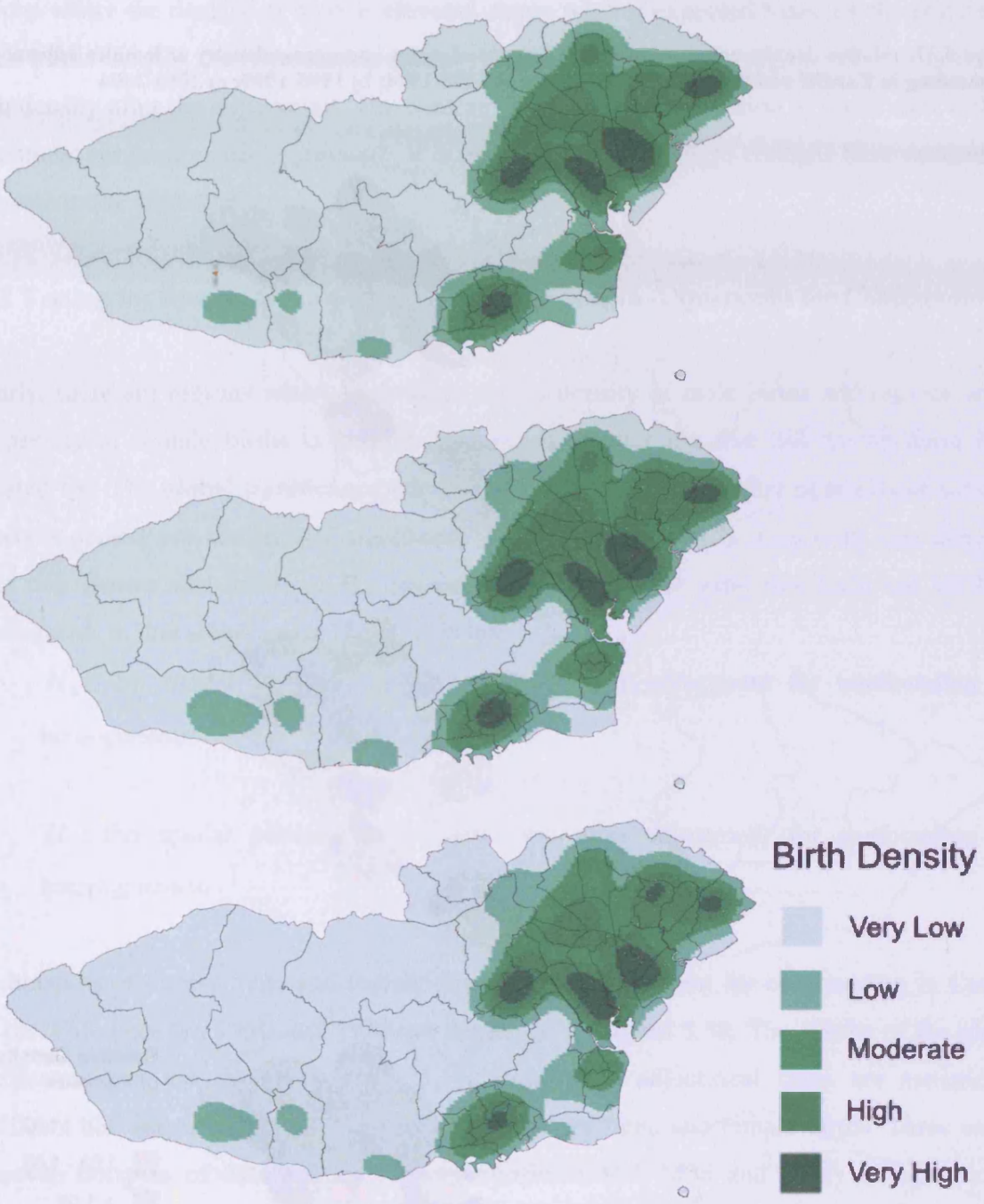


Figure 5.36: probability contours for the expected density of female births in Cardiff and the Vale of Glamorgan a) 1990-1994; b) 1995-1999; c) 2000-2004

a)

b)

c)



This adjusted density is then used as the denominator in producing relative density contours showing regions of excess males or of excess females after adjustment for the five risk factors. The variables have to be forced into the model because using any stepwise method of regression leads to a model with only a constant and no other terms since the dataset is too

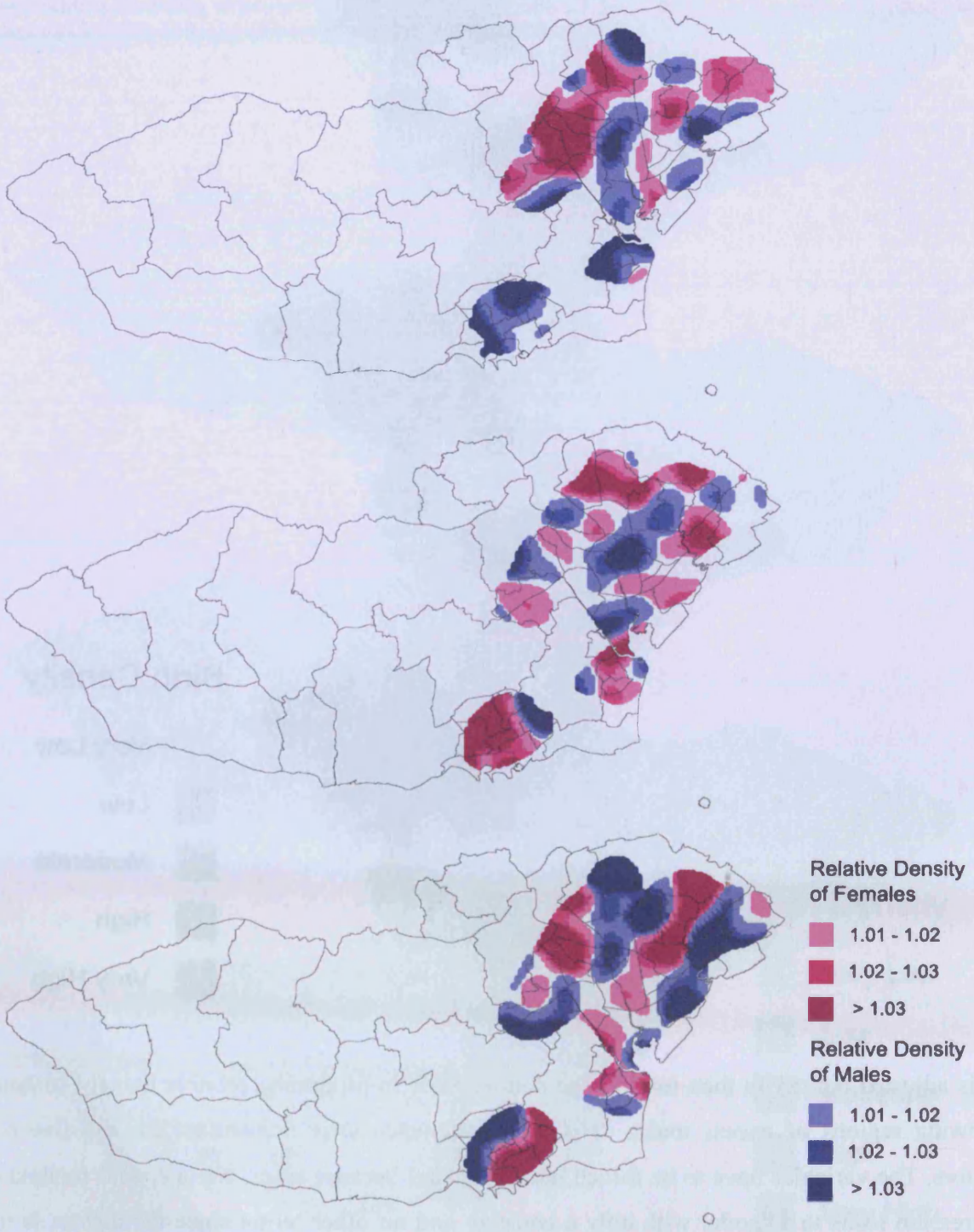
small to enable any of the five variables to be statistically significant factors. Another problem is that the same population for which the relative density contours were plotted for was used to fit the logistic regression model and so there will be overly optimistic fitting.

Figure 5.37: relative density contours of excess density of males and excess density of females adjusted for confounding in Cardiff and the Vale of Glamorgan a) 1990-1994; b) 1995-1999; c) 2000-2004

a)

b)

c)



The adjusted relative density contours for excess male and female birth density after adjustment for the five risk factors discussed are given in Figure 5.37. There are still some regions where the density of boys is elevated above what is expected based on the population of all births despite the adjustments made. Additionally, there are regions of excess female birth density after the adjustment. The contours are in a similar location to those seen without adjustment for confounding, although, it is evident that some slight changes have occurred in response to the adjustment.

5.5.2 Testing for Spatial Homogeneity of Sex Ratios with Adjustment for Confounding

Clearly, there are regions where there is an excess density of male births and regions where the density of female births is greater than expected after the five risk factors have been adjusted for. The global significance test can be used to assess whether or not these hotspots of excess density are statistically significant. If they are statistically significant then there are other risk factors that influence the spatial pattern in the sex ratio that have not yet been investigated. In this situation the hypotheses are

H_0: the spatial patterns in the sex ratio after adjustment for confounding are homogeneous

and

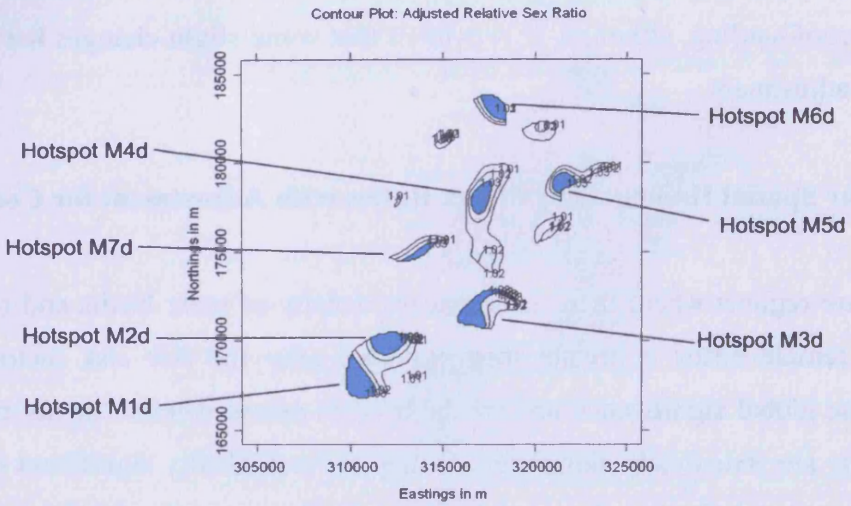
H_1: the spatial patterns in the sex ratio after adjustment for confounding are heterogeneous.

The hotspots of excess male and female density after adjustment for confounding in Cardiff and the Vale between 1990 and 1994 are displayed in Figure 5.38. The results of the global significance test are given in Table 5.11. Even after adjustment there are statistically significant elevations in the relative density surface for male and female births. Three out of the seven hotspots of excess male births identified (M1d, M3d and M7d) are statistically significant and two others (M2d and M6d) almost have statistically significant results. Furthermore, three of the five hotspots of excess female birth density (F1d, F2d and F4d) are also statistically significant. The existence of these statistically significant hotspots of excess density suggests that the alternative hypothesis is true thus providing evidence of a

heterogeneous spatial pattern in the sex ratio after adjustment for the five risk factors previously discussed.

Figure 5.38: hotspots of excess a) male and b) female birth density with adjustment for confounding in Cardiff and the Vale of Glamorgan 1990-1994

a)



b)

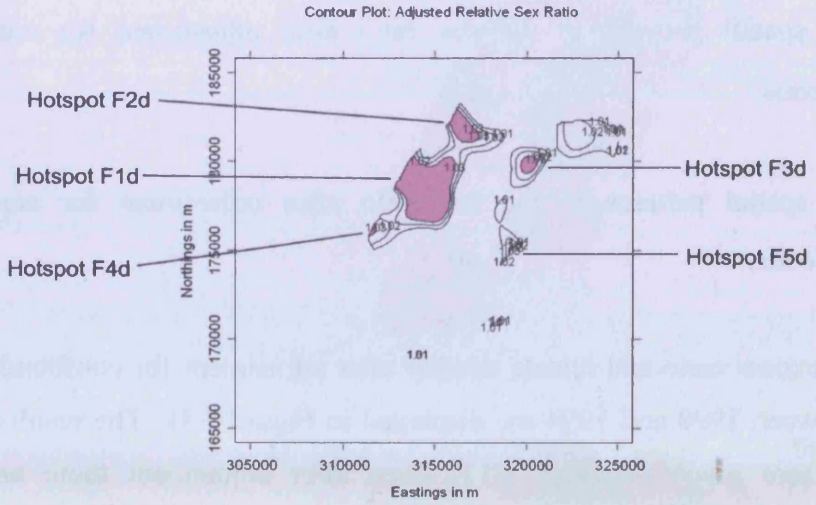


Table 5.11: results of the global significance test for hotspots of a) male and b) female birth density with adjustment for confounding in Cardiff and the Vale of Glamorgan 1990-1994

a)

Hotspot		M1d	M2d	M3d	M4d
Number of births		961	460	647	205
Number of male births		546	258	364	118
Observed mean relative density		1.071	1.040	1.062	1.033
p-value	test 1	<0.001	<0.001	<0.001	<0.001
	test 2	0.004	0.062	0.030	0.100
	test 3	<0.001	0.004	0.001	0.028

Hotspot		M5d	M6d	M7d
Number of births		202	364	81
Number of male births		117	209	59
Observed mean relative density		1.048	1.046	1.145
p-value	test 1	<0.001	<0.001	<0.001
	test 2	0.098	0.067	0.025
	test 3	0.023	0.001	<0.001

b)

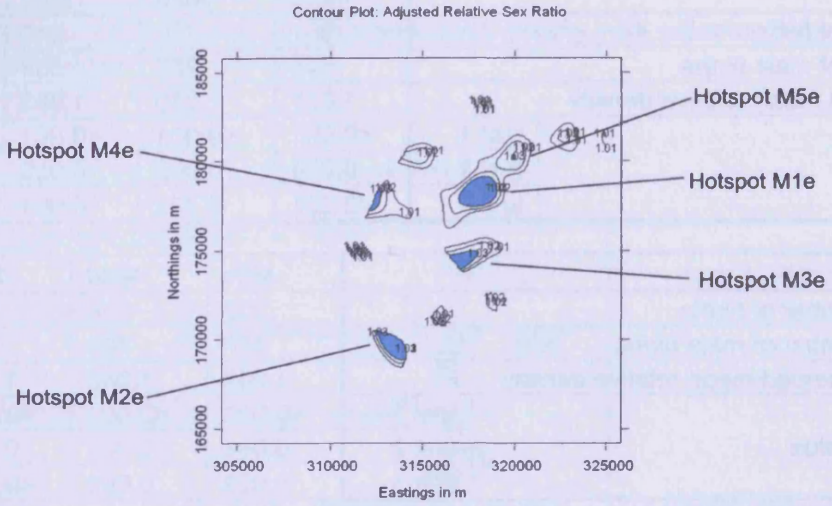
Hotspot		F1d	F2d	F3d	F4d
Number of births		1652	327	284	487
Number of female births		881	178	146	292
Observed mean relative density		1.059	1.090	1.033	1.136
p-value	test 1	<0.001	<0.001	<0.001	<0.001
	test 2	<0.001	0.002	0.112	<0.001
	test 3	0.001	0.049	0.285	<0.001

Hotspot		F5d
Number of births		103
Number of female births		56
Observed mean relative density		1.094
p-value	test 1	<0.001
	test 2	0.083
	test 3	0.234

Figure 5.39 displays the hotspots of excess male and female density in 1995 to 1999 and the results of the global significance test on these hotspots are given in Table 5.12. In 1995 to 1999 there are four statistically significant hotspots of excess male birth density (M1e to M4e) and two of excess female birth density (F1e and F2e), the existence of which suggest that the alternative hypothesis is true. Therefore, the global significance test results provide evidence that the spatial pattern in the sex ratio is heterogeneous after adjustment for the five risk factors.

Figure 5.39: hotspots of excess a) male and b) female birth density with adjustment for confounding in Cardiff and the Vale of Glamorgan 1995-1999

a)



b)

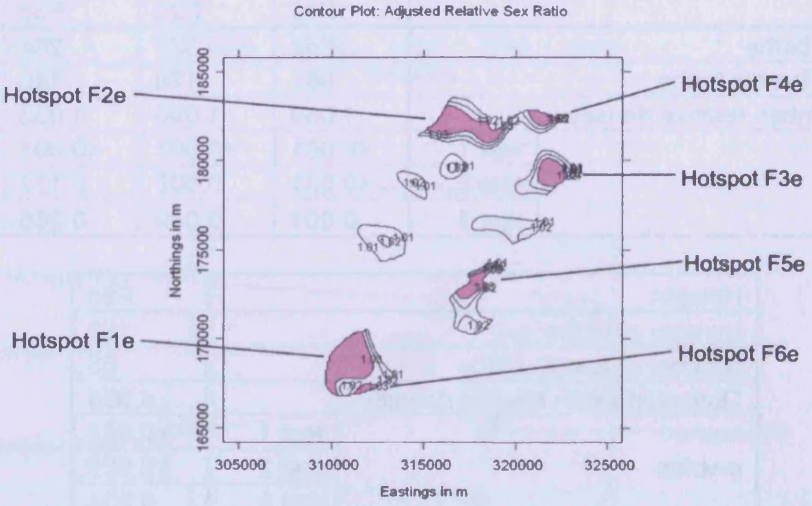


Table 5.12: results of the global significance test for hotspots of a) male and b) female birth density with adjustment for confounding in Cardiff and the Vale of Glamorgan 1995-1999

a)

Hotspot		M1e	M2e	M3e	M4e
Number of births		655	998	310	100
Number of male births		353	568	183	63
Observed mean relative density		1.045	1.082	1.113	1.158
p-value	test 1	<0.001	<0.001	<0.001	<0.001
	test 2	0.021	<0.001	0.003	0.015
	test 3	0.029	<0.001	<0.001	0.005

Hotspot		M5e
Number of births		195
Number of male births		113
Observed mean relative density		1.061
p-value	test 1	<0.001
	test 2	0.091
	test 3	0.017

b)

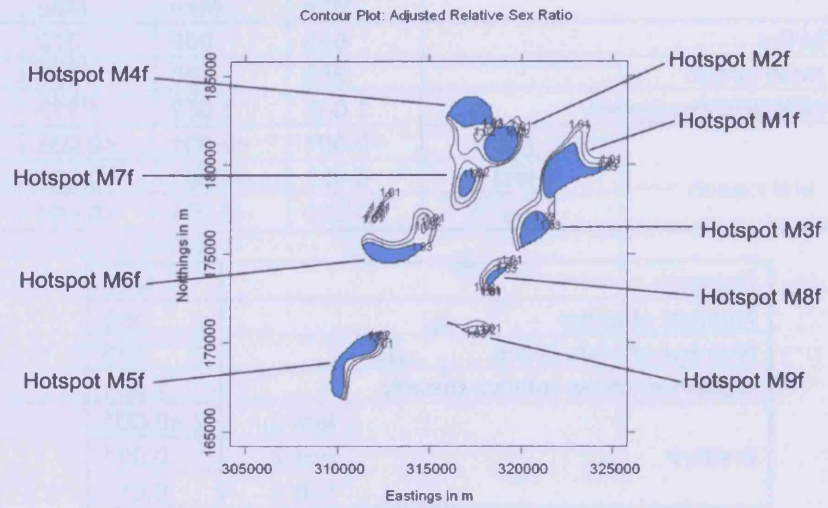
Hotspot		F1e	F2e	F3e	F4e
Number of births		1418	587	386	138
Number of female births		772	314	202	79
Observed mean relative density		1.058	1.046	1.038	1.044
p-value	test 1	<0.001	<0.001	<0.001	<0.001
	test 2	0.002	0.031	0.116	0.081
	test 3	<0.001	0.005	0.143	0.054

Hotspot		F5e	F6e
Number of births		204	83
Number of female births		120	50
Observed mean relative density		1.067	1.056
p-value	test 1	<0.001	<0.001
	test 2	0.089	0.091
	test 3	0.003	0.033

The hotspots of excess male and female density after adjustment for confounding in 2000 to 2004 are shown in Figure 5.40 and the results of the global significance test are in Table 5.13. Hotspots M5f to M7f, M9f, F1f to F3f and F6f have statistically significant results. The presence of four statistically significant hotspots of excess male density and four of excess female density suggest that the alternative hypothesis is true and that the relative density surface is heterogeneous even after adjustment for known confounders.

Figure 5.40: hotspots of excess a) male and b) female birth density with adjustment for confounding in Cardiff and the Vale of Glamorgan 2000-2004

a)



b)

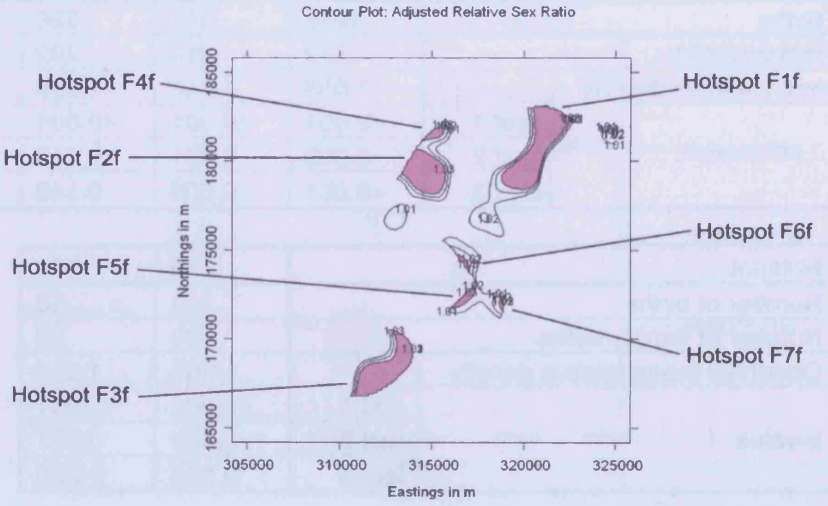


Table 5.13: results of the global significance test for hotspots of a) male and b) female birth density with adjustment for confounding in Cardiff and the Vale of Glamorgan 2000-2004

a)

Hotspot		M1f	M2f	M3f	M4f
Number of births		543	227	341	449
Number of male births		288	123	178	230
Observed mean relative density		1.060	1.053	1.039	1.048
p-value	test 1	<0.001	<0.001	<0.001	<0.001
	test 2	0.021	0.041	0.079	0.100
	test 3	0.109	0.113	0.278	0.270

Hotspot		M5f	M6f	M7f	M9f
Number of births		724	335	426	154
Number of male births		400	185	234	89
Observed mean relative density		1.080	1.094	1.078	1.122
p-value	test 1	<0.001	<0.001	<0.001	<0.001
	test 2	0.002	0.002	<0.001	0.029
	test 3	0.004	0.038	0.032	0.032

b)

Hotspot		F1f	F2f	F3f	F4f
Number of births		821	577	824	194
Number of female births		478	333	479	112
Observed mean relative density		1.054	1.080	1.065	1.094
p-value	test 1	<0.001	<0.001	<0.001	<0.001
	test 2	0.006	0.002	0.005	0.041
	test 3	<0.001	<0.001	<0.001	1.000

Hotspot		F5f	F6f	F7f
Number of births		61	106	53
Number of female births		35	64	31
Observed mean relative density		1.128	1.159	1.057
p-value	test 1	<0.001	<0.001	<0.001
	test 2	0.064	0.023	0.183
	test 3	0.180	0.028	0.143

The regions of excess males and of excess females suggest that there could be some risk factors further to the five that have been investigated that influence the sex ratio and lead to these heterogeneous spatial patterns. There are a number of possible risk factors for gender that have been discussed in the literature but data that can enable detailed investigation of these factors are not available particularly with relevant geographical information. For example, road traffic pollution particularly NO_2 and PM_{10} could be mapped using contours and checked against the patterns in the sex ratio. However, the required information is not available to a suitably high resolution and quality.

5.6 Conclusion

There is a statistically significant temporal trend in the sex ratio in England and Wales from 1965 to 2004 that is consistent with the declining trend in the sex ratio observed across the Westernised world. This decreasing trend in the sex ratio could be explained by temporal changes in levels of environmental contaminants. If exposure to environmental toxicants is a risk factor in the sex ratio then the heterogeneous spatial pattern of pollution would lead to a spatial pattern in the sex ratio.

Consequently, the sex ratio was mapped in Cardiff and the Vale of Glamorgan to assess whether or not there is a spatial pattern in this region. A heterogeneous spatial pattern was identified in the sex ratio with statistically significant regions of both excess male and female birth density. This spatially heterogeneous pattern could be a response to exposure to contaminants influencing the sex ratio. However, there are numerous other risk factors that have been suggested for sex ratios. Information on five risk factors was available including: maternal age, socioeconomic deprivation, maternal ethnicity, maternal BMI and parity and therefore these factors were investigated.

The sex ratio was high amongst mothers over 40 and mothers under 20 whereas it was lower amongst mothers between 20 and 39 years of age. Black and Indian mothers had high sex ratios, whereas Bangladeshi mothers had a low sex ratio. The sex ratio increases with increased maternal BMI. The sex ratio decreased with increasing affluence but the reason for the increased sex ratio with higher levels of deprivation may be down to higher obesity amongst more deprived mothers and the maternal BMI effect. There does seem to be an increased sex ratio amongst mothers with a high parity but an obvious trend is not present, which could be down to relatively small numbers of births for the higher parities.

Adjusting for these five factors, despite the lack of statistical significance, revealed that there are other risk factors for sex ratios that have not yet been considered. There are a number of possible risk factors that have been considered in the literature that have not been investigated here. These include timing of insemination, environmental conditions, stress and parental drug use. The most commonly discussed factor is that of exposure to environmental contaminants.

The level of exposure to environmental contaminants does follow spatial patterns which could explain, in part, the spatial patterns in the sex ratio. Unfortunately, adequate information on other possible risk factors is not available from the Cardiff Birth Survey so it has not been possible to investigate their relationship with the sex ratio. Therefore, a larger dataset is required and this dataset should contain information on a wider selection of possible risk factors. Additionally, a high resolution, good quality map of environmental exposure would assist further investigations into the sex ratio.

Chapter Six – Discussion

Spatial statistics has become an increasingly popular tool in investigations into the relationships between given health outcomes and putative environmental exposure from fixed sites. The reasons behind this popularity lie in the visual appeal of spatial statistics, increased availability of geographical information systems and a desire for a cleaner environment [1]. Current spatial statistical methods developed in the absence of individual level data generally focus on utilising area level data. However, improvements in geographical techniques [2], the development of a framework to assess the impact of environmental agents on health outcomes by the National Environmental Public Health Tracking Network (EPHTN) [3] and increased availability of point data for health outcomes based on address point motivates the need for new methodology that takes full advantage of the increased amount of information provided by working with individual level data [4]. The main aim of this thesis was to develop such methodology that facilitated this requirement. This was developed in three steps: mapping the risk of the specified health outcome, testing for spatial homogeneity of risk over the mapped region and testing for spatial association between the risk of the health outcome and putative risk factors.

Kernel density methodology is a powerful tool for mapping the risk of a health outcome using individual level data and thus provided an ideal framework from which new testing procedures could evolve. Consequently, starting from this foundation the global significance test for regions of elevated relative risk and the kernel density test for spatial association were developed. The relative risk contours that can be generated from kernel density methodology coupled with the two novel testing procedures cover all three steps in the spatial analysis of a health outcome using individual level data.

Once this methodology had been developed it was desirable to illustrate its uses. Birth outcomes are good examples to use for exhibiting spatial statistical techniques because the relatively short latency period reduces the bias induced by migration and temporal variations in exposure. Consequently, congenital malformations and sex ratios were chosen as the health outcomes of interest in this thesis. The additional benefit of using birth outcomes is the routine collection of data for both cases and controls including information on possible risk factors and geographical location. Both congenital malformations and sex ratios have

relatively unknown and complex aetiologies and hence any investigations into either of these health outcomes will be informative.

One of the most commonly discussed issues for congenital malformations is that of exposure to contaminants released from landfill sites. Although, this scenario was used in the evolution of the methodology, access to real exposure data was not forthcoming and thus the impact of exposure to toxicants in the vicinity of landfill sites on congenital malformations could not be used as an application for all of the new methodological procedures. However, the relative risk of congenital malformations around the Nant y Gwyddon landfill site in Wales was mapped and homogeneity was tested for, both with and without adjustment for known confounders.

In the absence of good quality exposure data the kernel density test for spatial association was demonstrated by comparing the spatial distributions for chromosomal and non-chromosomal anomalies. The results of the test suggested that chromosomal and non-chromosomal anomalies had different spatial distributions and from this it can be inferred that the two types of birth defects have different aetiologies. Additionally, the kernel density test was used to assess the procedure adopted for adjusting relative risk contours for known confounders and indicated that the adjustment was adequate.

The distribution of the sex ratio in Cardiff and the Vale of Glamorgan was mapped using both area level mapping procedures and relative risk contours which emphasises the benefits of using individual level data when mapping a health outcome. The motivation for considering the spatial distribution of the sex ratio lay in the hypothesis that observed temporal patterns in the sex ratio could be explained by exposure to environmental contaminants. If exposure to toxicants is a factor in explaining the sex ratio then a spatial pattern in the sex ratio is expected because the distribution of environmental pollution is spatially heterogeneous. In addition to generating relative risk contours, the global significance test for regions of elevated relative risk was performed, both with and without adjustment for known confounders. Regrettably, again lack of access to good quality, high resolution exposure data for any putative hazard prevented the use of the kernel density test for spatial association on the sex ratio example.

6.1 Methodology

6.1.1 Mapping Risk of the Health Outcome

The risk of a health outcome can be mapped using a summary measure determined for a set of regions defined by pre-specified boundaries and presenting this information in a choropleth map. The typical summary measure used is the standardised event ratio (SER) and smoothing techniques can be implemented to reduce sampling variation [157]. If the health outcome information is provided only at an area level for administrative regions then this is the logical approach to mapping risk. Kernel density estimation provides an alternative method of mapping risk where a relative risk surface can be determined by dividing the observed density surface by the expected density surface of the given health outcome [183, 184]. Contours are then presented based on the relative risk surface, which can locate regions where the relative risk is high. In contrast to mapping the SER, these relative risk contours require individual level data based on either address point or postcodes.

If the expected number of cases is low in a given region then the SER is subject to high sampling variation where high SER values can occur when there is, say, only one case. Similarly, elevations in the relative risk surface can occur because of one case in a region of low population density. The sampling variation in the SER can be reduced by using smoothing techniques but smoothing of this nature can make it difficult to differentiate between the effect of smoothing and true spatial dependence. The problem of high sampling variation in the relative risk contours can be avoided by filtering out regions of low population density after the relative risk surface has been fitted but before the contours are generated. Consequently, any elevations in the relative risk surface that could be the result of low population density as opposed to a high density of cases are not displayed on the map.

Finally, the main problem with mapping risk using SER's lies in the use of arbitrary pre-specified boundaries and ultimately means that kernel density mapping is superior. The use of regions in mapping a health outcome means that within region variation in the health outcome is overlooked. One region may contain one area where there is a high risk of the health outcome and one where there is a low risk which could lead to an overall SER of 1 suggesting no elevation in the risk. The relative risk surface on the other hand enables the identification of both the region of excess risk and that of low risk, providing appropriate values are selected

at which the contours are drawn. This ability to pinpoint the locations of increased or decreased risk of a given health outcome is advantageous from a public health perspective as it enables the implementation of locally relevant interventions. Consequently, if individual level data is available then the kernel density mapping technique is better than a summary statistic of the health outcome by region.

6.1.2 Testing for Spatial Homogeneity of the Health Outcome

There are numerous approaches that can be adopted for testing for the spatial homogeneity of risk of a given health outcome. A likelihood ratio based test for spatial homogeneity was developed here, where the study region is partitioned into sub-regions and the number of cases in each sub-region was assumed to have a binomial distribution. A likelihood ratio test statistic was then used to compare the observed situation to that of simulations under the null hypothesis of spatial homogeneity.

An alternative procedure for testing spatial homogeneity was developed from kernel density methodology. A threshold relative risk value was selected to define regions of excess relative risk, referred to as hotspots. The statistical significance of a hotspot was then determined using three component global significance tests. Test 1 is based on whether the relative risks for all individuals residing within the boundary of the hotspot are greater than the chosen threshold relative risk. Test 2 involves the mean relative risk value amongst individuals living within the boundary of the hotspot. Finally, test 3 is based on the number of cases of the health outcome of interest that lie within the boundaries of the hotspot.

There are a number of advantages to using the global significance test for regions of elevated relative risk over using the likelihood ratio based test for spatial homogeneity. Firstly, the latter is subject to the same problems as mapping with the SER. Therefore, the p-value for the likelihood ratio based test could indicate that there is spatial homogeneity even when there is within region variation in the risk of the health outcome. Secondly, the likelihood ratio based test involves assuming a binomial distribution, the parameters of which may be incorrectly estimated, whereas no such assumption is made in the global significance test.

The global significance test also has its advantages over other methods of testing for spatial homogeneity which have not been considered in detail in this thesis. Many methods are

designed for use with area level data and consequently can overlook within region variation, for example, Pearson's chi-squared test [187]. Some moving window tests, for example the geographical analysis machine [208], can be used with individual level data and thus overcome the problems related to using area level data. However, this group of testing procedures involve assuming the size and/or shape of the cluster *a priori* whereas relative risk contours do not. Other tests that work with individual level data include distance based tests like that proposed by Whittemore et al. [202]. Unfortunately, these distance based tests only test for spatial homogeneity and do not identify regions of excess risk whereas the global significance test involves identifying these regions from the outset.

One problem with the global significance test is that the relative risk value selected as the threshold for defining a region of excess density is arbitrary. However, the lower the threshold, the lower the mean relative risk value will be amongst individuals in the hotspot and thus it is less likely that statistical significance will be achieved in test 2. Moreover, the higher the threshold, the fewer cases of the health outcome there will be within the boundary of the hotspot and hence it is less likely that statistical significance will be achieved in test 3. Therefore, it is unlikely that statistically significant results in all the component global significance tests will be achieved when an inappropriate threshold value has been chosen.

A further problem with the global significance test lies in having three separate tests. If either statistical significance is achieved in all three tests or none of the tests achieve statistical significance then there is no problem. The difficulty lies in the situation where some of the component tests are statistically significant and others are not. In this scenario, it is not clear as to whether there is evidence to support or refute the null hypothesis. Therefore, the testing procedure could be improved by combining all three tests to generate one global significance test that leads to one unique p-value.

6.1.3 Testing for the Spatial Association between Exposure and the Health Outcome

There are many alternative ways of testing for the spatial association between the health outcome and exposure to a given risk factor. In this thesis the methods that were most commonly used in the literature were identified as the concentric circle approach (tradition approach) [224] and Stone's test [236] and these methods were considered in more detail.

There are numerous problems with the concentric circle approach. The choice of boundary that bisects the study region into exposed and unexposed is arbitrary and can influence the result of the test. The use of only two sub-regions means that within region variation can be easily overlooked. Finally, a large amount of information is lost when using only two categories of exposure and it is not possible to determine a dose-response effect.

In Stone's test, the study region is partitioned into more than two sub-regions which are ordered with decreasing exposure. The number of cases in each sub-region is assumed to have a binomial distribution. Stone uses the Poisson approximation to the binomial distribution under the assumption that the health outcome is rare and that the sample size is large. The proposed test statistic is based on a maximum likelihood ratio test where the maximum likelihood estimators for the parameters are given by a minimax formula that reflects the isotonic nature of the alternative hypothesis.

Stone's test is an improvement on the concentric circle approach as the use of more than two exposure groups provides more information on the dose-response relationship. Moreover, it is a powerful testing procedure and in an assessment of the efficiency of the test using simulations of congenital malformations around the Nant y Gwyddon landfill site the power at the 5% significance level was 84.9% (95% CI: 82.6%-87.0%) providing the full likelihood ratio test statistic is used.

However, the problem of overlooking within region variation is still a problem when using Stone's test which is not an issue when considering individual level data. Additionally, Stone's test becomes far less powerful when there is a smaller population size, a smaller effect size or a less variable exposure surface than the situation for which the power was quoted. The dramatic reduction in power could relate to the bias in the estimates in the relative risks for the sub-regions caused by the restriction imposed by using the minimax formula. This bias means that the relative risk always increases with increasing exposure, a characteristic that supports the alternative hypothesis, even under the null hypothesis.

Stone's paper does propose an individual level version of the testing procedure but it is not very powerful. Therefore, the likelihood ratio trend test for spatial association was developed based on individual level data. The likelihood ratio trend test is based on likelihood functions

derived from a Bernoulli distribution with the likelihood ratio defining the test statistic. The observed likelihood ratio is calculated and compared to likelihood ratios for simulations under the null hypothesis of no spatial association.

The likelihood ratio trend test for spatial association is more computationally difficult than Stone's test but it overcomes the problem of overlooking within region variation and gives more information on the dose-response relationship. The likelihood trend test is also a powerful test with a power of 88.7% (95% CI: 86.6%-90.5%) at the 5% significance level using simulations of congenital anomalies around Nant y Gwyddon. Furthermore, when the effect size is smaller, the population size is smaller or there is reduced variability in the mapped exposure in the study region then the likelihood ratio trend test is more powerful than Stone's test. An additional benefit to the likelihood ratio test is that this testing procedure is a two-tailed hypothesis test whereas Stone's test is one-tailed. This means that Stone's test is not equipped to detect protective risk factors. However, Stone's test was designed for use with putative point sources and it is implausible that exposure to one of these putative point sources of pollution would be protective.

An alternative method for testing for the spatial association between exposure to a risk factor and a given health outcome was developed from kernel density methodology referred to as the kernel density test for spatial association. The relative risk for each individual can be determined from the fitted relative risk surface and plotted against the corresponding exposure measure in a scatterplot. A threshold exposure model is chosen to dichotomise the population into the exposed and unexposed groups. The test statistics are the mean relative risk amongst the unexposed population, the proportion of relative risks above 1 in the exposed population and the slope of the linear regression model fitted to the scatterplot in the exposed region. These three test statistics are combined in to one test using discriminant analysis.

The kernel density test for spatial association does not work when the majority of exposure measures for individuals within the exposed region take only one value because the slope becomes unstable as seen when assessing the adjustment of relative risk contours for hospital of birth. In this situation the testing procedure can be used with only background mean and proportion but the power of the testing procedure will be reduced. Additionally the kernel density is similar to Stone's test in the respect that it is a one-tailed test and therefore will not detect a protective risk factor.

The kernel density test for spatial association is based on individual level data and thus overcomes the problems with overlooking within region variation and can provide more detailed information on the dose-response effect. The kernel density test is powerful with a power of 89.8% (95% CI: 87.8%-91.5%) at the 5% significance level based on simulations of congenital anomalies around the Nant y Gwyddon landfill site.

The kernel density testing procedure does lose efficiency when the effect size is reduced, the population size is decreased or the exposure measures are close to homogeneity in the study region. However, the efficiency of all testing procedures is influenced in this way. Clearly, the smaller the effect size the less important it is that any association is identified from a public health perspective. The problems of a small population size can be overcome by adopting a multiple site study which will increase the power. Although, care should be taken that good quality exposure data is available in a multiple site study. Finally, if the exposure measures are close to homogeneity in the study region then it is not a good study region in which to test for spatial association between exposure and a given health outcome.

Despite the ubiquitous reduced efficiency when there is a smaller effect size, a reduced population size or the exposure is more spatially uniform across the study region, the kernel density test far surpasses even the likelihood ratio trend test in terms of performance under all three of these conditions. Furthermore, the kernel density test for spatial association does not involve making assumptions about the distribution of cases unlike the likelihood ratio based testing procedures. Therefore, the kernel density test is recommended providing the exposure and health outcome data is of good quality and is provided at the individual level.

6.2 Congenital Malformations

The aetiology of congenital malformations is under constant debate in the literature yet despite this the causes of congenital malformations are largely unknown. One of the main discussions surrounding congenital malformations is concerned with increased risk in proximity to landfill sites. This can be considered using multiple site studies and single site studies. The problem with multiple site studies is that the quantity of waste, type of waste, local topography, local meteorology, leachate control and air dispersion control all influence the distribution of pollution and are site specific. Consequently, multiple site studies are

subject to problems that arise from these variations. Furthermore, the protection of the health of the public is a local issue and should be supported by studies that assess public health in a local setting. Consequently, the focus of this Chapter lies in a single landfill site, Nant y Gwyddon. However, a multiple site study could have been considered and would increase the power of the testing procedures adopted providing care was taken to combine landfill sites only if they have similar characteristics.

The crude relative risk of congenital malformations associated with living within 3km of the Nant y Gwyddon landfill site relative to living between 3km and 7km from the site is 1.303 (95% CI: 1.031-1.646) in 1983 to 1997. This is comparable to the results obtained from a study of 21 European landfill sites by Dolk et al. where the equivalent odds ratio for non-chromosomal anomalies was 1.33 (95% CI: 1.11-1.59) [64]. However, it is in contrast to the results obtained by Elliott et al. using 19,196 landfill sites in the UK where the relative risk of living within 2km of the landfill site is 1.01 (99% CI: 1.005-1.023) relative to remainder of the population [68]. The difference in the results for these studies could relate either to the choice of exposed region, the choice of the reference region or the inclusion by Elliott et al. of all landfill sites some of which will be benign thus reducing the landfill effect.

The concentric circle approach is overly simplistic with the use of only two exposure groups and these groups are defined using only distance from the site whereas other variables such as topographical and meteorological factors influence the pattern of exposure to contaminants released from landfill sites. Williams and Ogston identify a number of ways in which exposure to environmental contaminants can be modelled [284]. Relative risk contours are a better approach to assessing whether or not there are regions of excess congenital malformations in close proximity to the landfill site because no assumptions are made surrounding exposure.

The crude relative risk contours revealed that there were four regions over which the relative risk exceeded 2 between 1983 and 1997: two within 3km of the site, one approximately 7km from the site and one over 7km from the site. These four hotspots were statistically significant suggesting that there is some risk factor that is causing the observed spatial pattern in the relative risk of congenital malformations. Information on five putative risk factors was available on the dataset for 1983 to 1997 which could be fully or partially responsible for

these hotspots: year of birth, gender, maternal age, socioeconomic deprivation and hospital of birth.

The rate of congenital anomalies varies by year with two major changes in the reporting procedures for congenital malformations reflected in the temporal profile for the risk of birth defects: the introduction of the exclusion list in 1989 [277] and the change from only reporting congenital anomalies identified within 10 days of birth to being able to report birth defects whenever they were diagnosed [278]. Furthermore, deficiencies in reporting ascertainment by NCAS and changes in the diagnostic procedures led to changes in the rate of congenital malformations by year. In contrast, there is a fairly constant rate of congenital malformations by year of birth between 1998 and 2004 with the more fastidious data collection associated with the introduction of CARIS in 1998. Additionally, the rate of chromosomal anomalies in 1983 to 1997 is less variable by year because it is easier to diagnose and no chromosomal anomalies are included in the exclusion list.

The rate of congenital malformations is higher amongst male births compared to female births, which is largely the result of a higher rate of non-chromosomal anomalies. There is an increased risk of chromosomal anomalies associated with higher maternal age, which relates to the greater risk of Down's syndrome amongst mothers of an older age [23]. In contrast, the risk of non-chromosomal anomalies decreases with increased maternal age, which is supported by the knowledge that the risk of gastroschisis is raised amongst younger mothers [23]. The rate of congenital malformations increases slightly with increased deprivation. Although, there is a slightly higher rate in the most affluent quintile compared to the next affluent quintile, which reflects the higher proportion of older mothers in the most affluent quintile and the maternal age effect. The influence of the maternal age effect on the relationship between socioeconomic deprivation and congenital malformations is even more pronounced for chromosomal anomalies where the rate increases with increased affluence [37].

Some hospitals have higher reported rates of congenital malformations because mothers with a high risk of a birth defect can be referred to hospitals with specialist services thus increasing the rate at that hospital [23]. Additionally, the reported rate of congenital malformations varies between hospitals since it depends on diagnostic procedures and the diligence of the midwives

in notifying NCAS [278]. Moreover, some hospitals prioritise notification to NCAS more than others [268].

These five possible risk factors are adjusted for in the relative risk contours for congenital anomalies and this adjustment removes two of the regions over which the relative risk exceeds 2. However, there are still two hotspots remaining even after the adjustment: one to the northeast within 3km of Nant y Gwyddon and one to the northwest about 7km from the landfill site. Both these hotspots are statistically significant and therefore there is evidence to support the hypothesis that they are not the result of random variation. Hence, there could be other risk factors that have not yet been considered.

In order to explore possible risk factors that could explain the heterogeneous spatial pattern in the risk of congenital malformations, the kernel density test can be used to test for spatial association between mapped exposure for specific risk factors and the adjusted relative risk surface for congenital malformations. Unfortunately, good quality high resolution mapped exposure was not forthcoming and consequently the testing procedure could not be implemented in this way. However, it was possible to illustrate how the testing procedure could be used employing the example of testing for the spatial association between chromosomal and non-chromosomal anomalies.

The application of the methodology to comparing chromosomal and non-chromosomal anomalies indicated that after adjustment for the five confounding risk factors the relative risk surfaces for these two types of birth defects were not associated. The spatial non-concordance may be related to spatial pollution if this has a differential effect on the two types of congenital malformations. The possible disparities in the aetiologies of the two birth defects mean that the grouping of the chromosomal and non-chromosomal anomalies could dilute the true effects of certain risk factors. The problem extends further with many different types of chromosomal and non-chromosomal anomalies that may also differ in terms of aetiology. However, many investigators have no choice but to group together different types of congenital anomalies because each specific type of birth defect is a rare event so focusing on one type of congenital malformation leads to a substantial loss of power in any testing procedures implemented.

The results discussed above are based on data dating back to 1983 to 1997 and thus more recent data was considered to assess whether or not the elevations in the relative risk surface are still present around the Nant y Gwyddon landfill site. The relative risk contours for congenital malformations around the Nant y Gwyddon landfill site between 1998 and 2004 indicate that there are no longer any regions where the relative risk surface is elevated above 2. The reason for the disparity between the two time periods could lie in the relatively poor ascertainment rate for NCAS which was responsible for data collection in 1983 to 1997 [268]. Alternatively, it may be that the risk factor(s) responsible for the elevations in the relative risk contours in 1983 to 1997 may no longer be active in 1998 to 2004. For example, the Nant y Gwyddon landfill site closed in March 2002 which could explain why there are no longer any hotspots in its vicinity.

The kernel density test for spatial association was also used to test for association between the adjusted relative risk surface and the relative density surface for a given factor that was included in the adjustment. This tested the adequacy of the adjustment of the relative risk surface for confounding. The test could not be performed on hospital of birth because the nature of the relative density surface for East Glamorgan hospital led to instabilities in the slope as a test statistic. However, this is not a problem as the majority of births in the Nant y Gwyddon area were born at one hospital. An unfair comparison was made for year of birth where 1983 to 1989 was considered in the testing procedure whereas the logistic regression model adjusts for each year separately. The results do indicate that the adjustment for gender was adequate. Furthermore, despite statistically significant results for maternal age and socioeconomic deprivation the plotted points representing the observed test statistics suggest that the observed situation is more supportive of the null hypothesis. Thus, the adjustment for maternal age and socioeconomic deprivation is not perfect but is a move in the right direction. It is unlikely that any method used to adjust for confounding factors would be perfect. Thus, the adjustment used here was thought to be adequate.

6.3 Sex Ratios

The sex ratio is decreasing in the Westernised world and this decline is apparent in England and Wales between 1965 and 2004. There is a statistically significant temporal trend in the sex ratio in England and Wales during this time period. Hence, there is evidence that there are risk factors that influence the sex ratio and are responsible for the observed temporal patterns.

There are many putative risk factors discussed in the literature including: environmental contaminants, timing of insemination, environmental conditions, stress, nutrition, parental drug use, family status, parental dominance and parental ethnicity. The most prominent of these risk factors is that of exposure to environmental toxicants. If exposure to toxicants is a risk factor in the sex ratio then it is expected that the sex ratio will have a spatial, as well as temporal, pattern because environmental contaminants have a heterogeneous spatial pattern.

Consequently, the sex ratio was mapped in Cardiff and the Vale of Glamorgan to assess whether or not there is a spatial pattern in this region. A spatial pattern was identified in the study region with regions of excess male birth density and regions of excess female birth density clearly visible. Some of these regions of excess male and female birth density had statistical significance. Thus there is evidence that there is at least one risk factor for the sex ratio that is responsible for the heterogeneous spatial pattern observed. Information on five risk factors was available including: maternal age, socioeconomic deprivation, maternal ethnicity, maternal BMI and parity.

The sex ratio was high amongst mothers over 40 and mothers under 20 whereas it was lower amongst mothers between 20 and 39 years of age. Black and Indian mothers had high sex ratios whereas Bangladeshi mothers had a low sex ratio. Reports in the literature suggest that White fathers have higher sex ratios than Black fathers [97, 154]. The disparities between these reports and the result obtained in this thesis could be the result of considering the maternal not paternal ethnicity, small numbers in the groups for ethnic minorities or the poor quality of recording for the ethnic group category in the Cardiff Birth Survey.

The sex ratio increases with increased maternal BMI which supports the results obtained by Gibson et al. where the sex ratio was found to increase by the mid-upper arm muscle area [142]. Additionally, Cagnacci et al. found that women with a low pregnancy weight had a reduced sex ratio [139]. However, contrasting results have been obtained from other studies indicating that the complexity of nutrition means that it is not an easy risk factor to investigate [140].

The sex ratio increased with increasing deprivation which contradicts other studies on maternal dominance that indicate that women of a higher social standing are more likely have a son [97, 152]. The reason for the increased sex ratio with higher levels of deprivation may

be down to higher obesity amongst more deprived mothers and thus what is being detected here is in fact the maternal BMI effect.

There does seem to be an increased sex ratio amongst mothers with a high parity which corroborates findings reported in the literature [97]. However, an obvious trend of an increasing sex ratio with increased parity is not present which could be down to relatively small numbers of births for the higher parities.

These five risk factors discussed above were used to adjust the relative density surfaces for male and female births. Even after adjustment there were statistically significant regions of excess density of male and female births. This suggests that there is at least one risk factor other than the five that have already been considered which may be responsible for the heterogeneous spatial pattern in the sex ratio. One of these risk factors could be exposure to environmental contaminants however good quality, high resolution mapped exposure data was not available.

6.4 Future Work

The kernel density test for spatial association was compared to Stone's test in terms of performance. There are a number of other testing procedures that could also be considered, for example, score tests, linear risk score tests, logistic regression, Poisson regression or Poisson process techniques.

The global significance test produces three p-values and in some cases some of these p-values suggest statistical significance whilst others do not. Consequently, the logical next step is to combine all three tests in a way that allows the calculation of one global p-value. Additionally, when considering the procedures for testing for spatial association ROC curves were generated to assess the sensitivity of these methods. This process can be repeated for the testing procedures developed for assessing spatial homogeneity, thus enabling a direct comparison of the performance of each of the tests.

References

1 **Elliott, P. and Wartenburg, D.,** Spatial epidemiology: current approaches and future challenges. *Environmental Health Perspectives* 2004. **112**: 998-1006.

2 **Nuckols, J., Ward, M. and Jarup, L.,** Using geographic information systems for exposure assessment in environmental epidemiology studies. *Environmental Health Perspectives* 2004. **112**: 1007-1015.

3 **Mather, F., White, L., Langlois, E., Shorter, C., Swalm, C., Shaffer, J. and Hartley, W.,** Statistical methods for linking health, exposure, and hazards. *Environmental Health Perspectives* 2004. **112**: 1440-1445.

4 **Jarup, L.,** Health and environment information systems for exposure and disease mapping, and risk assessment. *Environmental Health Perspectives* 2004. **112**: 995-997.

5 **Vrijheid, M.,** Risk of congenital anomaly in relation to residence near hazardous waste landfill sites *London School of Hygiene and Tropical Medicine*. University of London, London 2000, pp 285.

6 **Kalter, H.,** Teratology in the 20th century: environmental causes of congenital malformations in humans and how they were established. *Neurotoxicology and Teratology* 2003. **25**: 131-282.

7 **Rushton, L.,** Health hazards and waste management. *British Medical Bulletin* 2003. **68**: 183-197.

8 Online encyclopedia, thesaurus, dictionary definitions and more 1999-2005 Answers Corporation. www.answers.com

9 **Fielder, H., Poon-King, C., Palmer, S., Moss, N. and Coleman, G.,** Assessment of impact on health of residents living near the Nant-y-Gwyddon landfill site: retrospective analysis. *British Medical Journal* 2000. **320**: 19-23.

10 **James, L., Matthews, I. and Nix, B.,** Spatial contouring of risk: a tool for environmental epidemiology. *Epidemiology* 2004. **15**: 287-292.

11 **Beischer, N., Mackay, E. and Colditz, P.,** *Obstetrics and the newborn: an illustrated textbook*, 3rd Edn: 1997.

12 **Campbell, S. and Lees, C.,** Conception, implantation and embryology *Obstetrics by ten teachers*, 17th Edn. Arnold, London 2000, pp 33-44.

13 **Rowlands, G.,** Reproduction and Genetics *A-level study guide: biology*. Pearson Education Limited 2000, pp 126-158.

14 **Chamberlain, G., Dewhurst, J. and Harvey, D.,** Foetal anatomy and physiology *Illustrated textbook of obstetrics*, 2nd Edn. Gower Medical Publishing, London 1991, pp 1-21.

15 Foetal development 2005 www.americanbaby.com/ab/category.jhtml?categoryid=/templatedata/ab/category 14/11/05

16 **Hanretty, K., Ramsden, I. and Callander, R.,** Physiology of reproduction *Obstetrics illustrated*, 6th Edn. Churchill Livingstone 2003, pp 1-14.

17 **James, D. and Pillai, M.,** *Obstetrics*, 2nd Edn. Churchill Livingstone: 1997.

18 Juedes, J. Foetal development timeline California Right to Life Education Fund. www.calright2life.org/fetal.htm 14/11/05

19 Fabes, R. and Martin, C. Prenatal and neonatal development Allyn and Bacon Companion Site. wps.ablongman.com/ab_fabes_exploring_2/0,4768,226579-,00.html 14/11/05

20 Timeline of the development of the human foetus The Fight Against Pop Culture. members.tripod.com/~peacepigeon/fetal.html 14/11/05

21 Foetal development 2001 Westside Pregnancy Resource Center. www.wprc.org/25.41.1.1.1.0.phtml 14/11/05

22 The unborn - a timeline: foetal development from conception to birth 2005 North Carolina Right to Life. www.ncrtl.org/Lifeline.htm 14/11/05

23 **Dolk, H.,** Epidemiologic approaches to identifying environmental causes of birth defects. *American Journal of Medical Genetics Part C: Seminars in Medical Genetics* 2004. **125C**: 4-11.

24 **Hanretty, K., Ramsden, I. and Callander, R.,** The newborn baby *Obstetrics illustrated*, 6th Edn. Churchill Livingstone 2003, pp 364-394.

25 **Campbell, S. and Lees, C.,** Prenatal diagnosis *Obstetrics by ten teachers*, 17th Edn. Arnold, London 2000, pp 175-185.

26 **Chamberlain, G., Dewhurst, J. and Harvey, D.,** The abnormal newborn baby *Illustrated textbook of obstetrics*, 2nd Edn. Gower Medical Publishing, London 1991, pp 207-219.

27 **Sanders, R., Blackmon, L., Hogge, W., Spevak, P. and Wulfsberg, E.,** *Structural foetal abnormalities: the total picture*, 2nd Edn. Mosby: 2002.

28 **Turner, T.,** Congenital abnormalities. In **Fraser, D. and Cooper, M.** (Eds.) *Myles textbook for midwives*, 14th Edn. Churchill Livingstone 2003, pp 841-864.

29 Wales Antenatal Screening Project. Welsh Assembly Government 2003.

30 **Campbell, S. and Lees, C.,** Antenatal imaging and assessment of foetal wellbeing *Obstetrics by ten teachers*, 17th Edn. Arnold, London 2000, pp 67-85.

31 **Twining, P.,** First trimester detection of foetal anomalies. In **Twining, P., McHugo, J. and Pilling, D.** (Eds.) *Textbook of foetal abnormalities*. Churchill Livingstone 2000, pp 41-69.

32 **Farndon, P.,** Foetal anomalies - the geneticist's approach. In **Twining, P., McHugo, J. and Pilling, D.** (Eds.) *Textbook of foetal abnormalities*. Churchill Livingstone 2000, pp 457-477.

33 CARIS review including 1998-2005 data. CARIS 2006.

34 **Sullivan, A. and Kirk, B.,** Specialised foetal investigations. In **Fraser, D. and Cooper, M.** (Eds.) *Myles textbook for midwives*, 14th Edn. Churchill Livingstone 2003, pp 411-432.

35 Karyotyping 2005 Medline Plus: A service for the US National Library and the National Institute. www.nlm.nih.gov/medlineplus/ency/article/003935.htm 10/01/06

36 **Dolk, H., Loane, M., Garne, E., de Walle, H., Queisser-Luft, A., De Vigan, C., Addor, M., Gener, B., Haeusler, M., Jordan, H., Tucker, D., Stoll, C., Feijoo, M., Lillis, D. and Bianchi, F.,** Trends and geographic inequalities in the prevalence of Down syndrome in Europe, 1980-1999. *Revue d'Epidemiologie et de Sante Publique* 2005. **53**: S87-S95.

37 **Dolk, H. and Vrijheid, M.,** The impact of environmental pollution on congenital anomalies. *British Medical Bulletin* 2003. **68**: 25-45.

38 **Archer, N., Langlois, P., Suarez, L., Brender, J. and Shanmugam, R.,** Association of paternal age with prevalence of selected birth defects. *Birth Defects Research (Part A)* 2007. **79**: 27-34.

39 Dolk, H., Doyle, P., Garne, E., Little, J., Loane, M., Robert, E., Seller, M. and Vrijheid, M. The environmental causes of congenital anomalies: a review of the literature 2004 www.eurocat.ulster.ac.uk/pubdata 10/01/06

40 Thalidomide as a teratogen Developmental Biology Online. www.devbio.com/printer.php?ch=21&id=200 10/01/06

41 Hirata, G. Congenital anomalies and teratogenesis 2002 Department of Pediatrics, University of Hawaii. www.hawaii.edu/medicine/pediatrics/pedtext/s04c02.html 10/01/06

42 **Bille, C., Murray, J. and Olsen, S.,** Folic acid and birth malformations. *British Medical Journal* 2007. **334**: 433-434.

43 **Potera, C.,** New evidence on vitamins, minerals, and pregnancy. *American Journal of Nursing* 2007. **107**: 19.

44 Foetal drugs effects Intoxikon International. www.intoxikon.com 10/01/06

45 Tuormaa, T. The adverse effects of the tobacco smoking on reproduction FORESIGHT. www.foresight-preconception.org.uk/summaries/frames/smoking-nf.html 10/01/06

46 **Webster, W.,** Teratogen update: congenital rubella. *Teratology* 1998. **58**: 13-23.

47 General Practice Notebook - Congenital malformations in maternal diabetes www.gpnotebook.co.uk/simplepage.cfm?*ID=1523580971* 10/01/06

48 **Bajaj, J., Misra, A., Rajalakshmi, M. and Madan, R.,** Environmental release of chemicals and reproductive ecology. *Environmental Health Perspectives* 1993. **101**: 125-130.

49 **Brender, J., Suarez, L., Felkner, M., Gilani, Z., Stinchcomb, D., Moody, K., Henry, J. and Hendricks, K.,** Maternal exposure to arsenic, cadmium, lead and mercury and neural tube defects in offspring. *Environmental Research* 2006. **101**: 132-139.

50 **Lawson, C., Schnorr, T., Whelan, E., Deddens, J., Dankovic, D., Piacitelli, L., Sweeney, M. and Connally, L.,** Paternal occupational exposure to 2,3,7,8-tetrachlorodibenzo-p-dioxin and birth outcomes of offspring: birth weight, preterm delivery, and birth defects. *Environmental Health Perspectives* 2004. **112**: 1403-1408.

51 **Shaw, G., Ranatunga, D., Quach, T., Neri, E., Correa, A. and Neutra, R.,** Trihalomethane exposures from municipal water supplies and selected congenital malformations. *Epidemiology* 2003. **14**: 191-199.

52 **Cordier, S., Chevrier, C., Robert-Gnansia, E., Lorente, C., Brula, P. and Hours, M.,** Risk of congenital anomalies in the vicinity of municipal solid waste incinerators. *Occupational and Environmental Medicine* 2004. **61**: 8-15.

53 **Dummer, T., Dickinson, H. and Parker, L.,** Adverse pregnancy outcomes around incinerators and crematoriums in Cumbria, north west England, 1956-93. *Journal of Epidemiology and Community Health* 2003. **57**: 456-461.

54 **Cresswell, P., Scott, J., Pattenden, S. and Vrijheid, M.,** Risk of congenital anomalies near the Byker waste combustion plant. *Journal of Public Health Medicine* 2003. **25**: 237-242.

55 **Gilboa, S., Mendola, P., Olshan, A., Longlois, P., Savitz, D., Loomis, D., Herring, A. and Fixler, D.,** Relation between ambient air quality and selected birth defects, seven county study, Texas, 1997-2000. *American Journal of Epidemiology* 2005. **162**: 238-252.

56 **Vrijheid, M.,** Health effects of residence near hazardous waste landfill sites: a review of epidemiologic literature. *Environmental Health Perspectives* 2000. **108**: 101-112.

57 **Brender, J., Zhan, F., Suarez, L., Langlois, P. and Moody, K.,** Maternal residential proximity to waste sites and industrial facilities and oral clefts in offspring. *Journal of Occupational and Environmental Medicine* 2006. **48**: 565-572.

58 **Najem, G. and Voyce, L.,** Health effects of a thorium waste disposal site. *American Journal of Public Health* 1990. **80**: 478-480.

59 **Gilbreath, S. and Kass, P.,** Foetal and neonatal deaths and congenital anomalies associated with open dumpsites in Alaska Native villages. *International Journal of Circumpolar Health* 2006. **65**: 133-147.

60 **Bentov, Y., Kordysh, E., Hershkovitz, R., Belmaker, I., Polyakov, M., Bilenko, N. and Sarov, B.,** Major congenital malformations and residential proximity to a regional industrial park including a national toxic waste site: an ecological study. *Environmental Health: a Global Access Science Source* 2006. **5**.

61 **Kloppenborg, S., Brandt, U., Gulis, G. and Ejstrud, B.,** Risk of congenital anomalies in the vicinity of waste landfills in Denmark; an epidemiological study using GIS. *Central European Journal of Public Health* 2005. **13**: 137-143.

62 **Boyle, E., Johnson, H., Kelly, A. and McDonnell, R.,** Congenital anomalies and proximity to landfill sites. *Irish Medical Journal* 2004. **97**: 16-18.

63 **Dolk, H.,** EUROCAT: 25 years of European surveillance of congenital anomalies. *Archives of Disease in Childhood - Fetal and Neonatal Edition* 2006. **90**: 355-358.

64 **Dolk, H., Vrijheid, M., Armstrong, B., Abramsky, L., Bianchi, F., Garne, E., Nelen, V., Robert, E., Scott, J., Stone, D. and Tenconi, R.,** Risk of congenital anomalies near hazardous-waste landfill sites in Europe: the EUROHAZCON study. *The Lancet* 1998. **352**: 423-427.

65 **Vrijheid, M., Dolk, H., Armstrong, B., Abramsky, L., Bianchi, F., Fazarinc, I., Garne, E., Ide, R., Nelen, V., Robert, E., Scott, J., Stone, D. and Tenconi, R.,** Chromosomal congenital anomalies and residence near hazardous waste landfill sites. *The Lancet* 2002. **359**: 320-322.

66 **Vrijheid, M., Dolk, H., Armstrong, B., Boschi, G., Busby, A., Jorgensen, T. and Pointer, P.,** Hazard potential ranking of hazardous waste landfill sites and risk of congenital anomalies. *Occupational and Environmental Medicine* 2002. **59**: 768-776.

67 **Harrison, R.,** Hazardous waste landfill sites and congenital anomalies. *Occupational and Environmental Medicine* 2003. **60**: 79-80.

68 **Elliott, P., Briggs, D., Morris, S., de Hoogh, C., Hurt, C., Jenson, T. K., Maitland, I., Richardson, S., Wakefield, J. and Jarup, L.,** Risk of adverse birth outcomes in populations living near landfill sites. *British Medical Journal* 2001. **323**: 363-368.

69 **Irvine, H.,** Landfill sites: is the risk of adverse birth outcomes really different in Scotland? *Scottish Medical Journal* 2003. **48**: 102-104.

70 **Morris, S., Thomson, A., Jarup, L., de Hoogh, C., Briggs, D. and Elliott, P.,** No excess risk of adverse birth outcomes in populations living near special waste landfill sites in Scotland. *Scottish Medical Journal* 2003. **48**: 105-107.

71 **Fielder, H., Palmer, S., Poon-King, C., Moss, N. and Coleman, G.,** Addressing environmental health concerns near Trecatti landfill site, United Kingdom. *Archives of Environmental Health* 2001. **56**: 529-535.

72 **Dolk, H.,** Commentary: impact on health needs assessing: from different angles. *British Medical Journal* 2000. **320**: 23.

73 **Palmer, S., Dunstan, F., Fielder, H., Fone, D., Higgs, G. and Senior, M.,** Risk of congenital anomalies after the opening of landfill sites. *Environmental Health Perspectives* 2005. **113**: 1362-1365.

74 **Chahnazarian, A.,** Determinants of the sex ratio at birth: review of the recent literature. *Social Biology* 1988. **35**: 214-235.

75 **Hesketh, T. and Xing, Z.,** Abnormal sex ratios in human populations: causes and consequences. *Proceedings of the National Academy of Science* 2006. **103**: 13271-13275.

76 **Davis, D., Gottlieb, M. and Stampnitzky, J.,** Reduced ratio of male to female births in several industrial countries: a sentinal health indicator? *Journal of the American Medical Association* 1998. **279**: 1018-1023.

77 **Marcus, M., Kiely, J., Xu, F., McGeehin, M., Jackson, R. and Sinks, T.,** Changing sex ratio in the United States, 1969-1995. *Fertility and Sterility* 1998. **70**: 270-273.

78 **Gutierrez-Adan, A., Pintado, B. and de la Fuente, J.,** Demographic and behavioural determinants of the reduction of male-female birth ratio in Spain from 1981 to 1997. *Human Biology* 2000. **72**: 891-898.

79 **Edlund, L.,** Son preference, sex ratios, and marriage patterns. *Journal of Political Economy* 1999. **107**: 1275-1304.

80 **Majerus, M.,** *Sex wars: genes, bacteria and biased sex ratios.* Princeton University Press: 2003.

81 **Trivers, R. and Willard, D.,** Natural selection of parental ability to vary the sex ratio of offspring. *Science* 1973. **179**: 90-92.

82 **Charnov, E.,** *The theory of sex allocation.* Princeton University Press: 1982.

83 **Sieff, D.,** Explaining biased sex ratios in human populations: a critique of recent studies. *Current Anthropology* 1990. **31**: 25-48.

84 **Gaulin, S. and Robbins, C.,** Trivers-Willard effect in contemporary North American society. *American Journal of Physical Anthropology* 1991. **85**: 61-69.

85 **Clutton-Brock, T.,** Sons and daughters. *Nature* 1982. **298**: 11-13.

86 **Hardy, I.,** *Sex ratios: concepts and research methods.* Cambridge University Press: 2002.

87 **Norberg, K.,** Partnership status and the human sex ratio at birth. *Proceedings. Biological Sciences/The Royal Society* 2004. **271**: 2403-2410.

88 **Marleau, J. and Saucier, J.,** Pregnant women's social status, stress, self-esteem, and their infants' sex ratio at birth. *Perceptual and Motor Skills* 2000. **91**: 697-702.

89 **James, W.,** Possible constraints on adaptive variation in sex ratio at birth in humans and other primates. *Journal of Theoretical Biology* 2006. **238**: 383-394.

90 **Koziel, S. and Ulijaszek, S.,** Waiting for Trivers and Willard: do the rich really favour sons? *American Journal of Physical Anthropology* 2001. **115**: 71-79.

91 **Burley, N.,** Sex ratio manipulation and selection for attractiveness. *Science* 1981. **211**: 721-722.

92 **Pergament, E., Todydemir, P. and Fiddler, M.,** Sex ratio: a biological perspective of 'Sex and the City'. *Reproductive Biomedical Online* 2002. **5**: 43-46.

93 **Assche, S.,** A different perspective on the imbalance of reported sex ratios at birth in rural China. *Stanford Journal of East Asian Affairs* 2004. **4**: 50-67.

94 **Jayaraj, D. and Subramanian, S.,** Women's wellbeing and the sex ratio at birth: some suggestive evidence from India. *Journal of Developmental Studies* 2004. **40**: 91.

95 **James, W.,** Further evidence that mammalian sex ratios at birth are partially controlled by parental hormone levels around the time of conception. *Human Reproduction* 2004. **19**: 1250-1256.

96 **James, W.,** Evidence that mammalian sex ratios at birth are partially controlled by parental hormone levels at the time of conception. *Journal of Theoretical Biology* 1996. **180**: 271-286.

97 **James, W.,** The human sex ratio. Part 1: a review of the literature. *Human Biology* 1987. **59**: 721-752.

98 **James, W.,** Offspring sex ratios at birth as markers of paternal endocrine disruption. *Environmental Research* 2006. **100**: 77-85.

99 **Astolfi, P., Cuccia, M. and Martinetti, M.,** Paternal HLA genotype and offspring sex ratio. *Human Biology* 2000. **71**: 315-319.

100 **Krackow, S.,** Potential mechanisms for sex ratio adjustment in mammals and birds. *Biological Reviews. Cambridge* 1995. **70**: 225-241.

101 **Boklage, C.,** The epigenetic environment: secondary sex ratio depends on differential survival in embryogenesis. *Human Reproduction* 2005. **20**: 583-587.

102 **Mackenzie, C., Lockridge, A. and Keith, M.,** Declining sex ratio in a First Nation community. *Environmental Health Perspectives* 2005. **113**: 1295-1298.

103 **Jongbloet, P., Zielhaus, G., Groenewould, H. and Pasker-de Jong, P.,** The secular trends in male:female ratio at birth in postwar industrialized countries. *Environmental Health Perspectives* 2001. **109**: 749-752.

104 **Figa-Talamanca, I., Carbone, P., Lauria, L., Spinelli, A. and Ulizzi, L.,** Environmental factors and the proportion of males at birth in Italy. *Archives of Environmental Health* 2003. **58**: 119-124.

105 **Milham, S.,** Unusual sex ratio of births to carbon setter fathers. *American Journal of Industrial Medicine* 1993. **23**: 829-831.

106 **Lichtenfels, A., Gomes, J., Pieri, P., El Khouri Miraglia, S., Hallak, J. and Saldiva, P.,** Increased levels of air pollution and a decrease in the human and mouse male-to-female ratio in Sao Paulo, Brazil. *Fertility and Sterility* 2007. **87**: 230-232.

107 **Sakamoto, M., Nakano, A. and Akagi, H.,** Declining Minamata male birth ratio associated with increased male foetal death due to heavy methylmercury pollution. *Environmental Research* 2001. **87**: 92-98.

108 **Webster, P.,** Canadian petrochemical plants blamed for gender imbalance. *The Lancet* 2006. **367**: 462.

109 **Simonsen, C., Roge, R., Christiansen, U., Larsen, T. and Bonde, J.,** Effects of paternal blood lead levels on offspring sex ratio. *Reproductive Toxicology* 2006. **22**: 3-4.

110 **Jarrell, J., Weisskopf, M., Weuve, J., Tellez-Rojo, M., Hu, H. and Hernandez-Avila, M.,** Maternal lead exposure and the secondary sex ratio. *Human Reproduction* 2006. **21**: 1901-1906.

111 **Saadat, M.,** Change in sex ratio at birth in Sardasht (north west of Iran) after chemical bombardment. *Journal of Epidemiology and Community Health* 2006. **60**: 183.

112 **Washam, C.,** PCB's legacy: fewer boys. *Environmental Health Perspectives* 2003. **111**: A146.

113 **del Rio Gomez, I., Marshall, T., Tsai, P., Shao, Y. and Guo, Y.,** Number of boys born to men exposed to polychlorinated byphenyls. *Lancet* 2002. **360**: 143-144.

114 **Weisskopf, M., Anderson, H. and Hanrahan, L.,** Decreased sex ratio following maternal exposure to polychlorinated biphenyls from contaminated Great Lakes sport-caught fish: a retrospective cohort study. *Environmental Health: A Global Access Science Source* 2003. **2**: 2.

115 **Karmaus, W., Huang, S. and Cameron, L.,** Parental concentration of dichlorodiphenyl dichloroethene and polychlorinated biphenyls in Michigan fish eaters and sex ratio in offspring. *Journal of Occupational and Environmental Medicine* 2002. **44**: 8-13.

116 **Ryan, J., Amirova, Z. and Carrier, G.,** Sex ratios of children of Russian pesticide producers exposed to dioxin. *Environmental Health Perspectives* 2002. **110**: A699-A701.

117 **Jongbloet, P., Roeleveld, N. and Groenewould, H.,** Where the boys aren't: dioxin and the sex ratio. *Environmental Health Perspectives* 2002. **110**: 1-3.

118 **Mocarelli, P., Gerthoux, P., Ferrari, E., Patterson, D., Kieszak, S., Brambilla, P., Vincoli, N., Signorini, S., Tramacere, P., Carreri, V., Sampson, E., Turner, W.**

and Needham, L., Paternal concentrations of dioxin and sex ratio of offspring. *Lancet* 2000. **355**: 1858-1863.

119 **Taylor, K., Jackson, L., Lynch, C., Kostyniak, P. and Louis, G.,** Preconception maternal polychlorinated biphenyl concentrations and the secondary sex ratio. *Environmental Research* 2007. **103**: 99-105.

120 **Anasari-Lari, M., Saadat, M. and Hadi, N.,** Influence of GSTT1 null genotype on the offspring sex ratio of gasoline filling station workers. *Journal of Epidemiology and Community Health* 2004. **58**: 393-394.

121 **Guerrero, R.,** Association of the type and time of insemination within the menstrual cycle with the human sex ratio at birth. *The New England Journal of Medicine* 1974. **291**: 1056-1059.

122 **Harlap, S.,** Gender of infants conceived on different days of the mentrual cycle. *The New England Journal of Medicine* 1979. **300**: 1445-1448.

123 **Gutierrez-Adañ, A., Perez-Garnelo, Granados, J., Garde, J., Perez-Guzman, M., Pintado, B. and de la Fuente, J.,** Relationship between sex ratio and time of insemination according to both time of ovulation and maturational state of oocyte. *Zygote* 1999. **7**: 37-43.

124 **Martin, J.,** Length of follicular phase, time of insemination, coital rate and the sex of offspring. *Human Reproduction* 1997. **12**: 611-616.

125 **Smits, L., de Bie, R., Essed, G. and van den Brandt, P.,** Time to pregnancy and sex of offspring: cohort study. *British Medical Journal* 2005. **331**: 1437-1438.

126 **Cagnacci, A., Renzi, A., Arangino, S., Alessandrini, C. and Volpe, A.,** Interplay between maternal weight and seasons in determining the secondary sex ratio of human offspring. *Fertility and Sterility* 2005. **84**: 246-248.

127 **Nonaka, K., Desjardins, B., Charbonneau, H., Legare, J. and Miura, T.,** Human sex ratio at birth and mother's birth season: multivariate analysis. *Human Biology* 1999. **71**: 875-884.

128 **Lerchl, A.,** Sex ratios at birth and environmental temperatures. *Naturwissenschaften* 1999. **86**: 340-342.

129 **Lyster, W.,** Altered sex ratio after the London smog of 1952 and the Brisbane flood of 1965. *The Journal of Obstetrics and Gynaecology of the British Commonwealth* 1974. **81**: 626-631.

130 **Fukuda, M., Fukuda, K., Shimizu, T. and Moller, H.,** Decline in sex ratio at birth after Kobe earthquake. *Human Reproduction* 1998. **13**: 2321-2322.

131 **Hansen, D., Moller, H. and Olsen, J.,** Severe periconceptional life events and the sex ratio in offspring: follow up study based on five national registers. *British Medical Journal* 1999. **319**: 548-549.

132 **Catalano, R., Bruckner, T., Gould, J., Eskenazi, B. and Anderson, E.,** Sex ratios in California following the terrorist attacks of September 11, 2001. *Human Reproduction* 2005. **20**: 1221-1227.

133 **Dobson, R.,** Fewer boys born in New York after 9/11 attacks. *British Medical Journal* 2006. **333**: 516.

134 **Catalano, R.,** Sex ratios in the two Germanies: a test of the economic stress hypothesis. *Human Reproduction* 2003. **18**: 1972-1975.

135 **Catalano, R. and Bruckner, T.,** Economic antecedents of the Swedish sex ratio. *Social Science and Medicine* 2005. **60**: 537-543.

136 **Graffelman, J. and Hoekstra, R.,** A statistical analysis of the effect of warfare on the human secondary sex ratio. *Human Biology* 2000. **72**: 433-445.

137 **Polasek, O., Kolcic, I., Kolaric, B. and Rudan, I.,** Short communication: sex ratio at birth and war in Croatia (1991-1995). *Human Reproduction* 2005. **20**: 2489-2491.

138 **Zorn, B., Sucur, V., Stare, J. and Meden-Vrtovec, H.,** Decline in sex ratio at birth after 10-day war in Slovenia. *Human Reproduction* 2002. **17**: 3173-3177.

139 **Cagnacci, A., Renzi, A., Arangino, S., Alessandrini, C. and Volpe, A.,** Influences of maternal weight on the secondary sex ratio of human offspring. *Human Reproduction* 2004. **19**: 442-444.

140 **Stein, A., Barnett, P. and Sellen, D.,** Maternal undernutrition and the sex ratio at birth in Ethiopia: evidence from a national sample. *Proceedings. Biological Sciences/The Royal Society* 2004. **271**: S37-S39.

141 **Williams, R. and Gloster, S.,** Human sex ratio as it relates to caloric availability. *Social Biology* 1992. **39**: 285-291.

142 **Gibson, M. and Mace, R.,** Strong mothers bear more sons in rural Ethiopia. *Proceedings. Biological Sciences/The Royal Society* 2003. **270**: S108-S109.

143 **Stein, A., Zybert, P. and Lumey, L.,** Acute undernutrition is not associated with excess of females at birth in humans: the Dutch Hunger Winter. *Proceedings. Biological Sciences/The Royal Society* 2004. **271**: S138-S141.

144 **Fukuda, M., Fukuda, K., Shimizu, T., Anderson, C. and Byskov, A.,** Parental periconceptional smoking and male:female ratio of newborn infants. *Lancet* 2002. **359**: 1407-1408.

145 **Mills, J., England, L., Granath, F. and Cnattingius, S.,** Cigarette smoking and the male-female sex ratio. *Fertility and Sterility* 2003. **79**: 1243-1245.

146 **Heron, J. and Ness, A.,** Lack of association between smoking behaviour and the sex ratio of offspring in the Avon Longitudinal Study of Parents and Children. *Fertility and Sterility* 2004. **81**: 700-702.

147 **Viloria, T., Rubio, M., Rodrigo, L., Calderon, G., Mercader, A., Mateu, E., Meseguer, M., Remohi, J. and Pellicer, A.,** Smoking habits of parents and male:female ratio in spermatozoa and preimplantation embryos. *Human Reproduction* 2005. **20**: 2517-2522.

148 **Shiono, P., Harlap, S. and Ramcharan, S.,** Sex of offspring of women using oral contraceptives, rhythm, and other methods of birth control around the time of conception. *Fertility and Sterility* 1982. **37**: 367-372.

149 **Luna, M., Duke, M., Copperman, A., Grunfield, L., Sandler, B. and Barritt, J.,** Blastocyst embryo transfer is associated with a sex-ratio imbalance in favour of male offspring. *Fertility and Sterility* 2007. **87**: 519-523.

150 **Whiting, J.,** The effect of polygyny on sex ratio at birth. *American Anthropologist* 1993. **95**: 435-442.

151 **Biggar, R., Wohlfahrt, J., Westergaard, T. and Melbye, M.,** Sex ratios, family size, and birth order. *American Journal of Epidemiology* 1999. **150**: 957-962.

152 **Grant, V. and Yang, S.,** Achieving women and declining sex ratio. *Human Biology* 2003. **75**: 917-927.

153 **Astolfi, P. and Zonta, L.,** Brief communication sex ratio and parental age gap. *Human Biology* 1999. **71**: 135.

154 **Khoury, M., Erickson, J. and James, L.,** Paternal effects on the human sex ratio at birth: evidence from interracial crosses. *American Journal of Human Genetics* 1984. **36**: 1103-1111.

155 **Snow, J.,** *On the mode of cholera*, Second Edn. John Churchill, London: 1855.

156 **Elliott, P., Martuzzi, M. and Shaddick, G.,** Spatial statistical methods in environmental epidemiology: a critique. *Statistical Methods in Medical Research* 1995. **4**: 137-159.

157 **Elliott, P., Wakefield, J., Best, N. and Briggs, D.,** *Spatial epidemiology: methods and applications*. Oxford University Press, Oxford: 2000.

158 **Lawson, A. B., Brown, W. J. and Vidal Rodeiro, C. L.,** *Disease mapping with WinBUGS and MLwiN.* John Wiley & Sons: 2003.
159 **Haining, R.,** *Spatial data analysis in the social and environmental sciences.* Cambridge University Press: 1990.
160 **Cressie, N.,** *Statistics for spatial data.* Wiley, New York: 1993.
161 **Lawson, A., Biggeri, A., Bohning, D., Lesaffre, E., Viel, J. and Bertollinin, R.,** *Disease mapping and risk assessment for public health.* Wiley, Chichester: 1999.
162 **Lawson, A.,** *Statistical methods in spatial epidemiology.* Wiley, Chichester: 2001.
163 **Ebdon, D.,** Spatial statistics *Statistics in Geography*, 2nd Edn. Billing and Sons, Worcester 1985, pp 128-175.
164 **Ripley, B.,** *Spatial statistics.* Wiley, New York: 1952.
165 **Cliff, A. and Ord, J.,** *Spatial processes: models and applications.* Pion, London: 1981.
166 **Langley, P. and Batty, M.,** *Spatial analysis: modelling in a GIS environment.* Wiley, New York: 1996.
167 **Lawson, A. and Williams, F.,** *An introductory guide to disease mapping.* Wiley, Chichester: 2001.
168 **Davies, J. and McCullagh, M.,** *Display and analysis of spatial data.* Wiley, New York: 1975.
169 **Robinson, T.,** Spatial statistics and geographical information systems in epidemiology and public health. *Advances in Parasitology* 2000. **47**: 81-128.
170 **Hills, M. and Alexander, F.,** Statistical methods used in assessing the risk of disease near a source of possible environmental pollution: a review. *Journal of the Royal Statistical Society Part A* 1989. **152**: 353-363.
171 **Gesler, W.,** The uses of spatial analysis in medical geography: a review. *Society for Science and Medicine* 1986. **23**: 963-973.
172 **Pascutto, C., Wakefield, J., Best, N., Richardson, S., Bernardinelli, L., Staines, A. and Elliott, P.,** Statistical issues in the analysis of disease mapping data. *Statistics in Medicine* 2000. **19**: 2493-2519.
173 **Rushton, L. and Elliott, P.,** Evaluating evidence on environmental health risks. *British Medical Bulletin* 2003. **68**: 113-128.
174 **Clayton, D. and Kaldor, J.,** Empirical Bayes estimates of age-standardized relative risks for use in disease mapping. *Biometrics* 1987. **43**: 671-681.
175 **Best, N., Richardson, S. and Thomson, A.,** A comparison of Bayesian spatial models for disease mapping. *Statistical Methods in Medical Research* 2005. **14**: 35-59.
176 **Butinx, F., Geys, H., Lousbergh, D., Broeders, G., Cloes, E., Dhollander, D., Op De Beeck, L., Vanden Brande, J., Van Waes, A. and Molenberghs, G.,** Geographical differences in cancer incidence in the Belgian province of Limburg. *European Journal of Cancer* 2003. **39**: 2058-2072.
177 **Jarup, L., Best, N., Toledano, M., Wakefield, J. and Elliott, P.,** Geographical epidemiology of prostate cancer in Great Britain. *International Journal of Cancer* 2002. **97**: 695-699.
178 **Toledano, M., Jarup, L., Best, N., Wakefield, J. and Elliott, P.,** Spatial variation and temporal trends of testicular cancer in Great Britain. *British Journal of Cancer* 2001. **84**: 1482-1487.
179 **Cislaghi, C., Biggeri, A., Braga, M., Lagazio, C. and Marchi, M.,** Exploratory tools for disease mapping in geographical epidemiology. *Statistics in Medicine* 1995. **14**: 2363-2381.
180 **Aylin, P., Maheswaren, R., Wakefield, J., Cockings, S., Jarup, L., Arnold, R., Wheeler, G. and Elliott, P.,** A national facility for small area disease mapping and

rapid initial assessment of apparent disease clusters around a point source: the UK Small Area Health Statistics Unit. *Journal of Public Health Medicine* 1999. **21**: 289-298.

181 **Martuzzi, M., Grundy, C. and Elliott, P.,** Perinatal mortality in an English health region: geographical distribution and association with socio-economic factors. *Paediatric and Perinatal Epidemiology* 1998. **12**: 263-276.

182 **Eaton, N., Shaddick, G., Dolk, H. and Elliott, P.,** Small-area study of the incidence of neoplasms of the brain and central nervous system among adults in the West Midlands Region, 1974-86. *British Journal of Cancer* 1997. **75**: 1080-1083.

183 **Bowman, A. W. and Azzalini, A.,** *Applied smoothing techniques for data analysis: the kernel density approach with S-Plus illustrations*: 1997.

184 **Silverman, B.,** *Density estimation for statistics and data analysis (monographs on statistics and applied probability.* Chapman and Hall: 1986.

185 **Lawson, A. and Williams, F.,** Armadale: a case-study in environmental epidemiology. *Journal of the Royal Statistical Society Part A* 1994. **157**: 285-298.

186 **Williams, F., Lawson, A. and Lloyd, O.,** Low sex ratios of births in areas at risk from air pollution from incinerators, as shown by geographical analysis and 3-dimensional mapping. *International Journal of Epidemiology* 1992. **21**: 311-319.

187 **Ohno, Y., Aoki, K. and Aoki, N.,** A test of significance for geographic clusters of disease. *International Journal of Epidemiology* 1979. **8**: 273-281.

188 **Potthoff, R. and Whittinghill, M.,** Testing for homogeneity: I. the binomial and multinomial distributions. *Biometrika* 1966. **53**: 167-182.

189 **Potthoff, R. and Whittinghill, M.,** Testing for homogeneity: II. the Poisson distribution. *Biometrika* 1966. **53**: 183-190.

190 **Dolk, H., Shaddick, G., Walls, P., Grundy, C., Thakrar, B., Kleinschmidt, I. and Elliott, P.,** Cancer incidence near radio and television transmitters in Great Britain: I. Sutton Coldfield Transmitter. *American Journal of Epidemiology* 1997. **145**: 1-9.

191 **Michelozzi, P., Capon, A., Kirchmayer, U., Forastiere, F., Biggeri, A., Barca, A. and Perucci, C.,** Adult and childhood leukaemia near a high-power radio station in Rome, Italy. *American Journal of Epidemiology* 2002. **155**: 1096-1103.

192 **Moran, P.,** The interpretation of statistical maps. *Journal of the Royal Statistical Society Part B* 1948. **10**: 243-251.

193 **Geary, R.,** The contiguity ratio and statistical mapping. *The Incorporated Statistician* 1954. **5**: 115-145.

194 **Getis, A.,** Interaction modelling using second-order analysis. *Environment and Planning A* 1984. **16**: 173-183.

195 **Getis, A. and Franklin, J.,** Second-order neighborhood analysis of mapped point patterns. *Ecology* 1987. **68**: 473-477.

196 **Getis, A.,** Screening for spatial dependence in regression analysis. *Papers of the Regional Science Association* 1990. **69**: 69-81.

197 **Hungerford, L.,** Use of spatial statistics to identify and test significance in geographic disease patterns. *Preventive Veterinary Medicine* 1991. **11**: 237-242.

198 **Getis, A. and Ord, J.,** The analysis of spatial association by use of distance statistics. *Geographical Analysis* 1992. **24**: 189-206.

199 **Anselin, L.,** Local indicators of spatial association - LISA. *Geographical Analysis* 1995. **27**: 93-115.

200 **Kitron, U., Michael, J., Swanson, J. and Haramis, L.,** Spatial analysis of the distribution of LaCrosse encephalitis in Illinois, using a geographic information system and local and global spatial statistics. *American Journal of Tropical Medicine and Hygiene* 1997. **57**: 469-475.

201 **Glick, B.,** The spatial autocorrelation of cancer mortality. *Society for Science and Medicine* 1979. **13D**: 123-130.

202 **Whittemore, A., Friend, N., Brown, B. and Holly, E.,** A test to detect clusters of disease. *Biometrika* 1987. **74**: 631-635.

203 **Tango, T.,** A class of tests for detecting 'general' and 'focused' clustering of rare diseases. *Statistics in Medicine* 1995. **14**: 2323-2334.

204 **Song, C. and Kulldorff, M.,** Likelihood based tests for spatial randomness. *Statistics in Medicine* 2006. **25**: 825-839.

205 **Olsen, S. F., Martuzzi, M. and Elliott, P.,** Cluster analysis and disease mapping - why, when, and how? A step by step guide. *British Medical Journal* 1996. **313**: 863-866.

206 **Wartenburg, D.,** Investigating disease clusters: why, when and how? *Journal of the Royal Statistical Society, A* 2001. **164**: 13-22.

207 **Elliott, P. and Wakefield, J.,** Disease clusters: should they be investigated, and, if so, when and how? *Journal of the Royal Statistical Society, A* 2001. **164**: 3-12.

208 **Openshaw, S., Charlton, M., Wymer, C. and Craft, A.,** A mark I geographical analysis machine for the automated analysis of point data sets. *International Journal of Geographical Information Systems* 1987. **1**: 335-358.

209 **Openshaw, S.,** Using a geographic analysis machine to detect the presence of spatial clustering and the location of clusters in synthetic data. In **Alexander, F. and Boyle, P.** (Eds.) *Methods for investigating localized clustering of disease*. IARC Scientific Publications 135, Lyon 1996.

210 **Besag, J. and Newell, J.,** The detection of clusters in rare diseases. *Journal of the Royal Statistical Society, A* 1991. **154**: 143-155.

211 **Cuzick, J. and Edwards, R.,** Spatial clustering for inhomogeneous populations. *Journal of the Royal Statistical Society Part B* 1990. **52**: 73-104.

212 **Rogerson, P.,** Statistical methods for the detection of spatial clustering in case-control data. *Statistics in Medicine* 2006. **25**: 811-823.

213 **Turnbull, B., Iwano, E., Burnett, W., Howe, H. and Clark, L.,** Monitoring for clusters of disease: application to leukemia incidence in upstate New York. *American Journal of Epidemiology* 1990. **132**: S136-S143.

214 **Kulldorff, M. and Nagarwalla, N.,** Spatial disease clusters: detection and inference. *Statistics in Medicine* 1995. **14**: 799-810.

215 **Takahashi, K. and Tango, T.,** An extended power of cluster detection tests. *Statistics in Medicine* 2006. **25**: 841-852.

216 **Kulldorff, M., Huang, L., Pickle, L. and Duczmal, L.,** An elliptic spatial scan statistics. *Statistics in Medicine* 2006. **25**: 3929-3943.

217 **Hjalmars, U., Kulldorff, M., Gustafsson, G. and Nagarwalla, N.,** Childhood leukaemia in Sweden: using GIS and a spatial scan statistic for cluster detection. *Statistics in Medicine* 1996. **15**: 707-715.

218 **Kulldorff, M., Feuer, E., Miller, B. and Freedman, L.,** Breast cancer clusters in the northeast United States: a geographic analysis. *American Journal of Epidemiology* 1997. **146**: 161-170.

219 **Assuncao, R., Costa, M., Tavares, A. and Ferreira, S.,** Fast detection of arbitrarily shaped disease clusters. *Statistics in Medicine* 2006. **25**: 723-742.

220 **Rose, D., Boyar, A. and Wynder, E.,** International comparisons of mortality rates for cancer of the breast, ovary, prostate, and colon, and per capita food consumption. *Cancer* 1986. **58**: 2363-2371.

221 **Armstrong, B. and Doll, R.,** Environmental factors and cancer incidence and mortality in different countries, with special reference to dietary practices. *International Journal of Cancer* 1975. **15**: 617-631.

222 **Alexander, F., Cartwright, R., McKinney, P. and Rickets, T.,** Leukaemia incidence, social class and estuaries: an ecological analysis. *Journal of Public Health Medicine* 1990. **12**: 109-117.

223 **Chen, J., Geissler, C., Parpia, B., Li, J. and Campbell, T.,** Antioxidant status and cancer mortality in China. *International Journal of Epidemiology* 1992. **21**: 625-635.

224 **Elliott, P., Westlake, A., Hills, M., Kleinschmidt, I., Rodrigues, L., McGale, P., Marshall, K. and Rose, G.,** The Small Area Health Statistics Unit: a national facility for investigating health around point sources of environmental pollution in the United Kingdom. *Journal of Epidemiology and Community Health* 1992. **46**: 345-349.

225 **Draper, G., Vincent, T., Kroll, M. and Swanson, J.,** Childhood cancer in relation to distance from high voltage power lines in England and Wales: a case-control study. *British Medical Journal* 2005. **330**: 1290-1293.

226 **Jarup, L., Briggs, D., de Hoogh, C., Morris, S., Hurt, C., Lewin, A., Maitland, I., Richardson, S., Wakefield, J. and Elliott, P.,** Cancer risks in populations living near landfill sites in Great Britain. *British Journal of Cancer* 2002. **86**: 1732-1736.

227 **Cook-Mozaffari, P., Darby, S. and Doll, R.,** Cancer near potential sites of nuclear installations. *The Lancet* 1989. **334**: 1145-1147.

228 **Hellstrom, L., Elinder, C., Dahlberg, B., Lundberg, M., Jarup, L., Persson, B. and Axelson, O.,** Cadmium exposure and end-stage renal disease. *American Journal of Kidney Disease* 2001. **38**: 1001-1008.

229 **Dolk, H., Thakrar, B., Walls, P., Landon, M., Grundy, C., Saez Lloret, I., Wilkinson, P. and Elliott, P.,** Mortality among residents near cokeworks in Great Britain. *Occupational and Environmental Medicine* 1999. **56**: 34-40.

230 **Maheswaren, R. and Elliott, P.,** Stroke mortality associated with living near main roads in England and Wales: a geographical study. *Stroke* 2003. **34**: 2776-2780.

231 **Morris, S. E., Sale, R. C., Wakefield, J. C., Falconer, S., Elliott, P. and Boucher, B. J.,** Hospital admissions for asthma and chronic obstructive airways disease in East London hospitals and proximity of residence to main roads. *Journal of Epidemiology and Community Health* 2000. **54**: 75-76.

232 **Venn, A., Yemaneberhan, H., Lewis, S., Parry, E. and Britton, J.,** Proximity of the home to roads and the risk of wheeze in an Ethiopian population. *Occupational and Environmental Medicine* 2005. **62**: 376-380.

233 **Livingstone, A., Shaddick, G., Grundy, C. and Elliott, P.,** Do people living near inner city main roads have more asthma needing treatment? Case-control study. *British Medical Journal* 1996. **312**: 676-677.

234 **Hatch, M. and Susser, M.,** Background gamma radiation and childhood cancers within ten miles of a US nuclar plant. *International Journal of Epidemiology* 1990. **19**: 546-552.

235 **Gardner, M., Snee, M., Hall, A., Powell, C., Downes, S. and Terrell, J.,** Results of case-control study of leukaemia and lymphoma among young people near Sellafield nuclear plant in West Cumbria. *British Medical Journal* 1990. **300**: 423-429.

236 **Stone, R. A.,** Investigation of excess environmental risks around putative sources: statistical problems and a proposed test. *Statistics in Medicine* 1988. **7**: 649-660.

237 **Shaddick, G. and Elliott, P.,** Use of Stone's method in studies of disease risk around a point sources of environmental pollution. *Statistics in Medicine* 1996. **15**: 1927-1934.

238 **Morton-Jones, T., Diggle, P. and Elliott, P.,** Investigation of excess environmental risk around putative sources: Stone's test with covariate adjustment. *Statistics in Medicine* 1999. **18**: 189-197.

239 **Aylin, P., Bottle, A., Wakefield, J., Jarup, L. and Elliott, P.,** Proximity to coke works and hospital admissions for respiratory and cardiovascular disease in England and Wales. *Thorax* 2001. **56**: 228-233.

240 **Dolk, H., Pattenden, S., Vrijheid, M., Thakrar, B. and Armstrong, B.,** Perinatal and infant mortality and low birth weight among residents near cokeworks in Great Britain. *Archives of Environmental Health* 2000. **55**: 26-33.

241 **Parodi, S., Vercelli, M., Stella, A., Stagnaro, E. and Valerio, F.,** Lymphohaematopoietic system cancer incidence in an urban area near a coke oven plant: an ecological investigation. *Occupational and Environmental Medicine* 2003. **60**: 187-193.

242 **Chellini, E., Cherubini, M., Chetoni, L., Costantini, A., Biggeri, A. and Vannucchi, G.,** Risk of respiratory cancer around a sewage plant in Prato, Italy. *Archives of Environmental Health* 2002. **57**: 548-553.

243 **Wilkinson, P., Thakrar, B., Shaddick, G., Stevenson, S., Pattenden, S., Landon, M., Grundy, C. and Elliott, P.,** Cancer incidence and mortality around the Pan Britannica Industries pesticide factory, Waltham Abbey. *Occupational and Environmental Medicine* 1997. **54**: 101-107.

244 **Wilkinson, P., Thakrar, B., Walls, P., Landon, M., Falconer, S., Grundy, C. and Elliott, P.,** Lymphohaematopoietic malignancy around all industrial complexes that include major oil refineries in Great Britain. *Occupational and Environmental Medicine* 1999. **56**: 577-580.

245 **Sans, S., Elliott, P., Kleinschmidt, I., Shaddick, G., Pattenden, S., Walls, P., Grundy, C. and Dolk, H.,** Cancer incidence and mortality near the Baglan Bay petrochemical works, South Wales. *Occupational and Environmental Medicine* 1995. **52**: 217-224.

246 **Elliott, P., Hills, M., Beresford, J., Kleinschmidt, I., Jolley, D., Pattenden, S., Rodrigues, L., Westlake, A. and Rose, G.,** Incidence of cancers of the larynx and lung near incinerators of waste solvents and oils in Great Britain. *The Lancet* 1992. **339**: 854-858.

247 **Elliott, P., Shaddick, G., Kleinschmidt, I., Jolley, D., Walls, P., Beresford, J. and Grundy, C.,** Cancer incidence near municipal solid waste incinerators in Great Britain. *British Journal of Cancer* 1996. **73**: 702-710.

248 **Dolk, H., Elliott, P., Shaddick, G., Walls, P. and Thakrar, B.,** Cancer incidence near radio and television transmitters in Great Britain: II. all high power transmitters. *American Journal of Epidemiology* 1997. **145**: 10-17.

249 **Bithell, J., Dutton, S., Draper, G. and Neary, N.,** Distribution of childhood leukaemias and non-Hodgkin's lymphomas near nuclear installations in England and Wales. *British Medical Journal* 1994. **309**: 501-505.

250 **Viel, J., Pobel, D. and Carre, J.,** Incidence of leukaemia in young people around the La Hague nuclear waste reprocessing plant: a sensitivity analysis. *Statistics in Medicine* 1995. **14**: 2459-2472.

251 **Tarone, R.,** The use of historical control information in testing for a trend in Poisson means. *Biometrics* 1982. **38**: 457-462.

252 **Waller, L., Turnbull, B., Clark, L. and Nasca, P.,** Spatial pattern analysis to detect rare disease clusters. In **Lange, N. and Ryan, L.** (Eds.) *Case studies in biology.* Wiley, New York 1994, pp 3-23.

253 **Lawson, A.,** On the analysis of mortality events associated with a prespecified fixed point. *Journal of the Royal Statistical Society Part A* 1993. **156**: 363-377.

254 **Bithell, J,** The choice of test for detecting raised disease risk near a point source. *Statistics in Medicine* 1995. **14**: 2309-2322.

255 **Upton, G. and Cook, I.,** *Oxford dictionary of statistics*. Oxford University Press: 2002.

256 **Diggle, P.,** A point process modelling approach to raised incidence of a rare phenomenon in the vicinity of a prespecified point. *Journal of the Royal Statistical Society Part A* 1990. **153**: 349-362.

257 **Maheswaren, R., Morris, S., Falconer, S., Grossinho, A., Perry, I., Wakefield, J. and Elliott, P.,** Magnesium in drinking water supplies and mortality from acute myocardial infarction in north west England. *Heart* 1999. **82**: 455-460.

258 **Richardson, S., Monfort, C., Green, M., Draper, G. and Muirhead, C.,** Spatial variation of natural radiation and childhood leukaemia incidence in Great Britain. *Statistics in Medicine* 1995. **14**: 2487-2501.

259 **Marsh, G., Stone, R., Esmen, N., Gula, M., Gause, C., Peterson, N., Meaney, F., Rodney, S. and Prybylski, D.,** A case-control study of lung cancer mortality in six Gila Basin, Arizona smelter towns. *Environmental Research* 1997. **75**: 56-72.

260 **Marsh, G., Stone, R., Esmen, N., Gula, M., Gause, C., Peterson, N., Meaney, F., Rodney, S. and Prybylski, D.,** A case-control study of lung cancer mortality in four rural Arizona smelter towns. *Archives of Environmental Health* 1998. **53**: 15-31.

261 **Nyberg, F., Gustavsson, P., Jarup, L., Bellander, T., Berglind, N., Jakobsson, R. and Pershagen, G.,** Urban air pollution and lung cancer in Stockholm. *Epidemiology* 2000. **11**: 487-495.

262 **Biggeri, A., Barbone, F., Lagazio, C., Bovenzi, M. and Stanta, G.,** Air pollution and lung cancer in Trieste, Italy: spatial analysis of risk as a function of distance from sources. *Environmental Health Perspectives* 1996. **104**: 750-754.

263 **Klonoff-Cohen, H., Lam, P. and Lewis, A.,** Outdoor carbon monoxide, nitrogen dioxide, and sudden infant death syndrome. *Archives of Disease in Children* 2005. **90**: 750-753.

264 **Pikhart, H., Bobak, M., Kriz, B., Danova, J., Celko, M. A., Prikazsky, V., Pryl, K., Briggs, D. and Elliott, P.,** Outdoor air concentrations of nitrogen dioxide and sulphur dioxide and prevalence of wheezing in school children. *Epidemiology* 2000. **11**: 153-160.

265 **Pikhart, H., Bobak, M., Gorynski, P., Wojtyniak, B., Danova, J., Celko, M., Kriz, B., Briggs, D. and Elliott, P.,** Outdoor sulphur dioxide and respiratory symptoms in Czech and Polish school children: a small-area study (SAVIAH). *International Archives of Occupational and Environmental Health* 2001. **74**: 574-578.

266 **Aylin, P., Morris, S., Wakefield, J., Grossinho, A., Jarup, L. and Elliott, P.,** Temperature, housing, deprivation and their relationship to excess winter mortality in Great Britain, 1986-1996. *International Journal of Epidemiology* 2001. **30**: 1100-1108.

267 **Toledano, M., Nieuwenhuijsen, M., Best, N., Whitaker, H., Hamley, P., De Hoogh, C., Fawell, J., Jarup, L. and Elliott, P.,** Relation of trihalomethane concentrations in public water supplies to stillbirth weight in three regions in England. *Environmental Health Perspectives* 2005. **113**: 225-232.

268 **Boyd, P., Armstrong, B., Dolk, H., Botting, B., Pattenden, S., Abrahmsky, L., Rankin, J., Vrijheid, M. and Wellesley, D.,** Congenital anomaly surveillance in England - ascertainment of deficiencies in the national system. *British Medical Journal* 2005. **330**: 27.

269 **Legendre, P.,** Spatial autocorrelation: trouble or new paradigm. *Ecology* 1993. **74**: 1659-1673.

270 **Duczmal, L. and Buckeridge, D.,** A workflow spatial scan statistic. *Statistics in Medicine* 2006. **25**: 743-754.

271 **Breslow, N., Lubin, J., Marek, P. and Langholz, B.,** Multiplicative models and cohort analysis. *Journal of the American Statistical Assocation* 1983. **78**: 1-12.

272 **James, L. M.,** Use of non-parametric kernel density techniques to determine factors that influence congenital anomaly rates in Wales *Epidemiology, Statistics and Public Health.* University of Wales College of Medicine, Cardiff 2003.

273 **Dunstan, F. D. J., Nix, A. B. J., Reynolds, J. F. and Rowlands, R. J.,** Regression and correlation *Introductory statistics: statistical inference.* R.N.D Publications 1983, pp 364-420.

274 **Lachenbruch, P. A.,** *Discriminant analysis.* Hafner Press: 1974.

275 **Johnson, R. and Wichern, D.,** *Applied multivariate statistical analysis,* 3rd Edn. Prentice-Hall International Editions: 1992.

276 **Garson, D.,**Discriminant function analysis (three groups): SPSS output

277 The National Congenital Anomaly System www.statistics.gov.uk/downloads/theme_health/National_Congenital_Anomaly_System/CA_text_v2.pdf 30/10/06

278 **Yates, R. and Cook, L.,**Geographical variation in congenital anomaly notifications in England and Wales 30/10/06

279 **Rankin, J., Pattenden, S., Abrahmsky, L., Boyd, P., Jordan, H., Stone, D., Vrijheid, M., Wellesley, D. and Dolk, H.,** Prevalence of congenital anomalies in five British regions, 1991-99. *Archives of Disease in Childhood - Fetal and Neonatal Edition* 2005. **90**: 374-379.

280 CARIS reporting system 2006 www.wales.nhs.uk/sites3/page.cfm?orgid=416&pid=3480 30/10/06

281 Annual update: congenital anomaly statistics: notifications, 2005, England and Wales. *Health Statistics Quarterly* 2007. **33**: 93-96.

282 Congenital anomaly statistics notifications: a statistical review of notifications of congenital anomalies received as part of the England and Wales National Congenital Anomaly System, 2005 *MB3 no.20.* Office for National Statistics 2006.

283 **Mathews, T. and Hamilton, B.,** Trend analysis of the sex ratio at birth in the United States. *National Vital Statistics Report* 2005. **53.**

284 **Williams, F. and Ogston, S.,** Identifying populations at risk from environmental contamination from point sources of pollution. *Occupational and Environmental Medicine* 2002. **59**: 2-8.

Contents of Appendices

Figure A1.1: the ROC curve for Stone's test using 3, 6 and 12 exposure groups

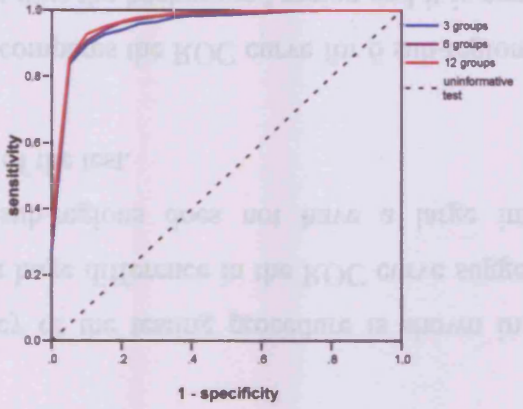
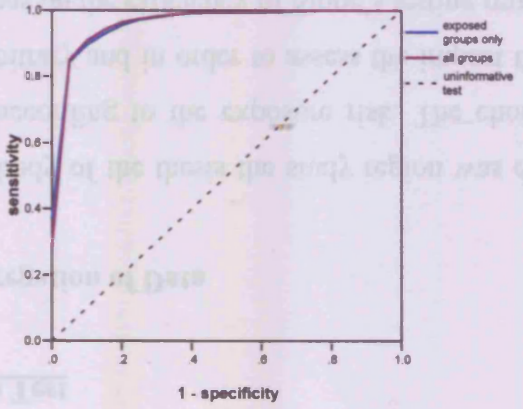
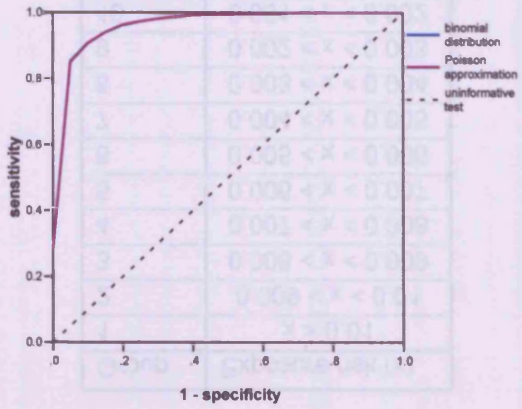


Figure A1.2: the ROC curve for Stone's test including and excluding the background region



A.1.1.2 Binomial Distribution and Poisson Approximation

Figure A1.3: the ROC curve for Stone's test using the binomial assumption and the Poisson approximation



Appendix One – ROC Curve Comparisons

A1.1 Stone's Test

A.1.1.1 Aggregation of Data

In the main body of the thesis the study region was divided into 6 sub-regions according to the exposure risk. The choice of 6 sub-regions is arbitrary and in order to assess the impact the number of sub-regions has on the efficiency of Stone's testing procedure, 3 and 12 sub-regions were considered as defined in Table A1.1 and Table A1.2.

The efficiency of the testing procedure is shown in Figure A1.1 there is not a huge difference in the ROC curve suggesting that the number of sub-regions does not have a large impact on the performance of the test.

Figure A1.2 compares the ROC curve for 6 sub-regions with that of 6 sub-regions plus the background region and it is apparent that the inclusion of the background region does not influence the efficiency of the test.

Table A1.1: the exposure groups for when the data is split into 3 groups

Group	Exposure risk (x)
1	$x > 0.006$
2	$0 < x < 0.006$
3	$x = 0$

Table A1.2: the exposure groups for when the data is split into 12 groups

Group	Exposure risk (x)
1	$x > 0.01$
2	$0.009 < x < 0.01$
3	$0.008 < x < 0.009$
4	$0.007 < x < 0.008$
5	$0.006 < x < 0.007$
6	$0.005 < x < 0.006$
7	$0.004 < x < 0.005$
8	$0.003 < x < 0.004$
9	$0.002 < x < 0.003$
10	$0.001 < x < 0.002$
11	$0 < x < 0.001$
12	$x = 0$

Figure A1.8: the ROC curve for testing using proportion and slope with smoothing parameters 1km, 0.5km and 0.25km

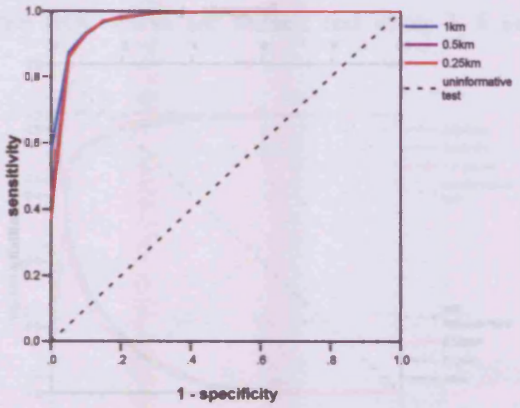


Figure A1.9: the ROC curve for testing using slope and background mean with smoothing parameters 1km, 0.5km and 0.25km

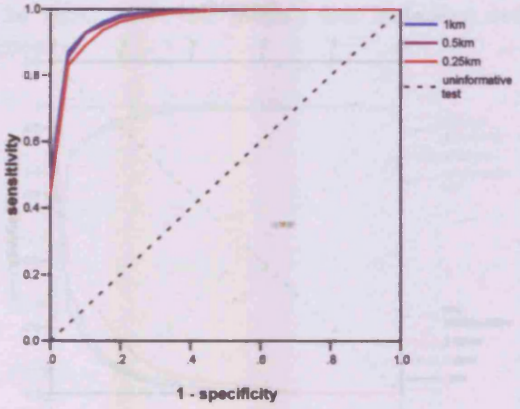
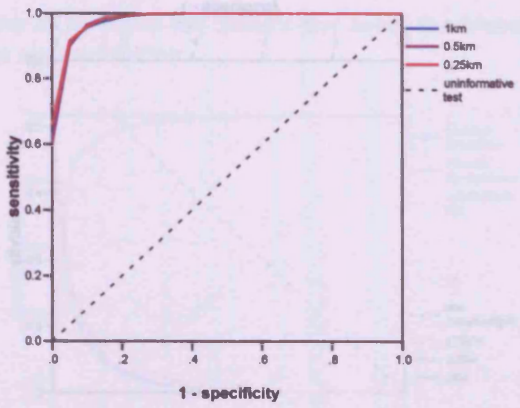


Figure A1.10: the ROC curve for testing using background mean, proportion and slope with smoothing parameters 1km, 0.5km and 0.25km



A1.2 Smoothing Parameters in the Relative Risk Surface

Figure A1.4: the ROC curve for testing using background mean with smoothing parameters 1km, 0.5km and 0.25km

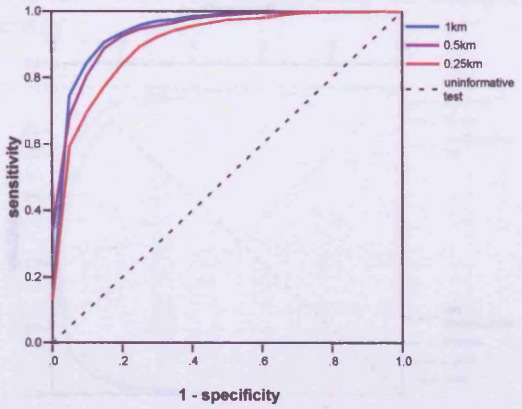


Figure A1.5: the ROC curve for testing using proportion with smoothing parameters 1km, 0.5km and 0.25km

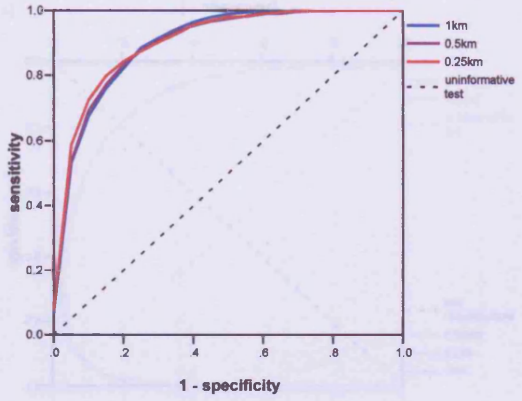


Figure A1.6: the ROC curve for testing using slope with smoothing parameters 1km, 0.5km and 0.25km

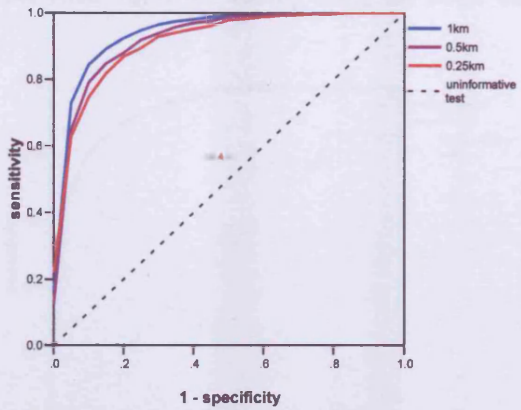


Figure A1.7: the ROC curve for testing using background mean and proportion with smoothing parameters 1km, 0.5km and 0.25km

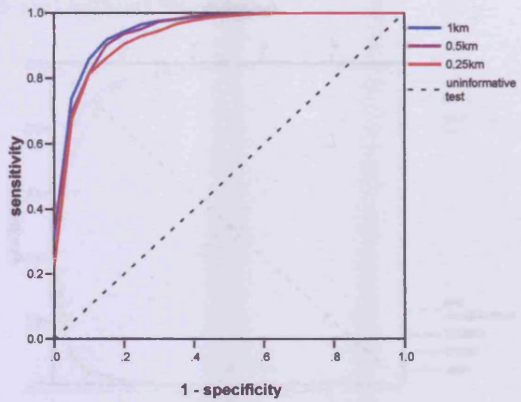


Figure A1.15: the ROC curve for testing using proportion and slope with and without filtering

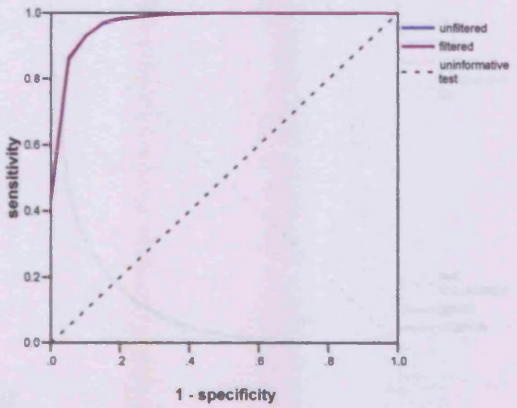


Figure A1.16: the ROC curve for testing using slope and background mean with and without filtering

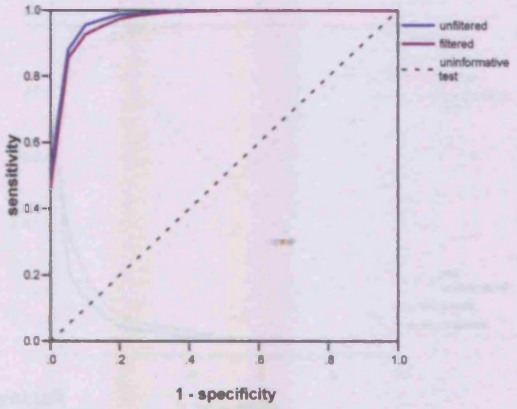
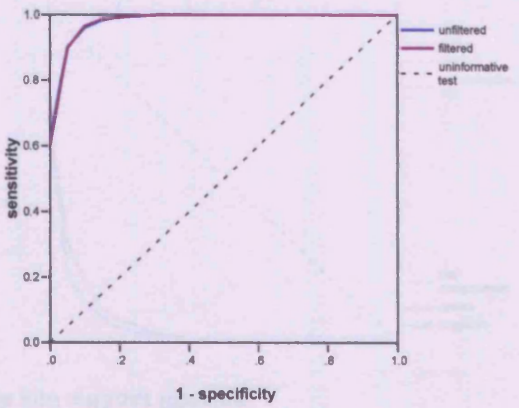


Figure A1.17: the ROC curve for testing using background mean, proportion and slope with and without filtering



A1.3 Filtering the Dataset

A1.3.1 Nant y Gwyddon

Figure A1.11: the ROC curve for testing using the background mean with and without filtering

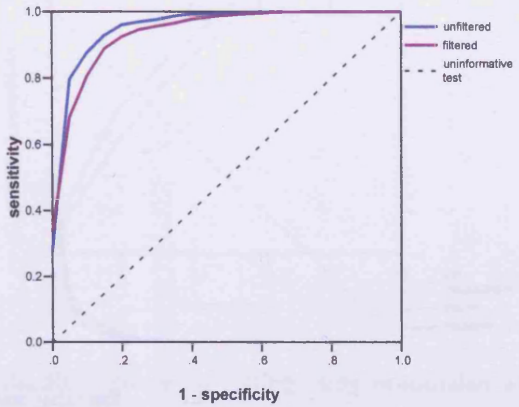


Figure A1.12: the ROC curve for testing using the proportion with and without filtering

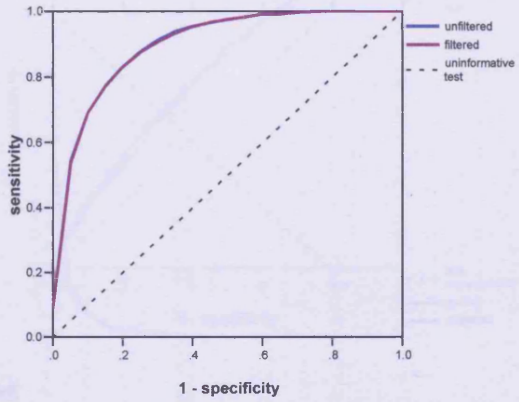


Figure A1.13: the ROC curve for testing using the slope with and without filtering

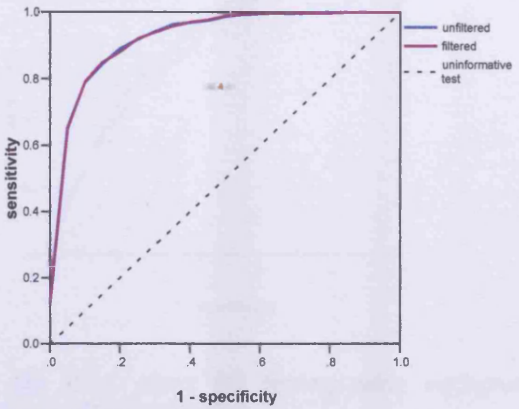


Figure A1.14: the ROC curve for testing using background mean and proportion with and without filtering

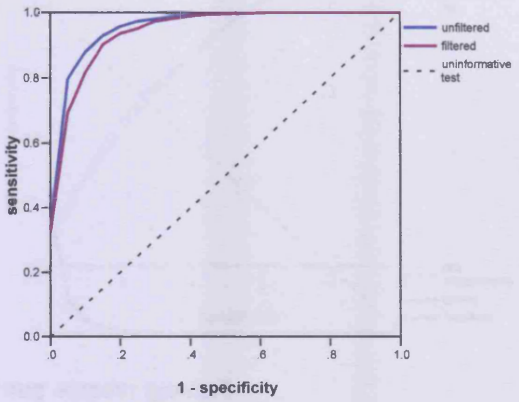


Figure A1.22: the ROC curve for testing using proportion and slope with and without filtering

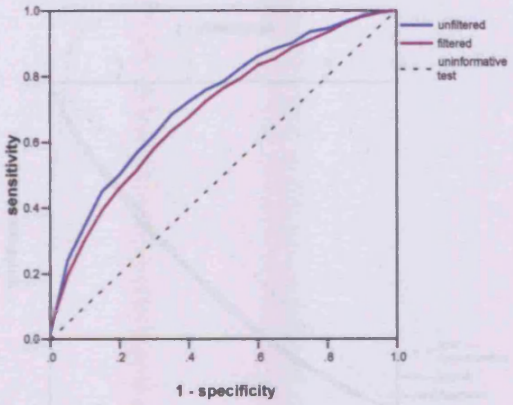


Figure A1.23: the ROC curve for testing using slope and background mean with and without filtering

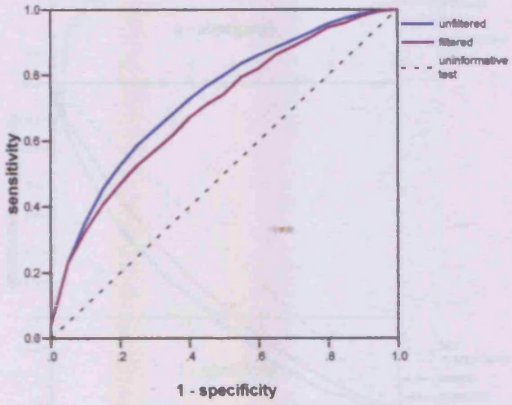
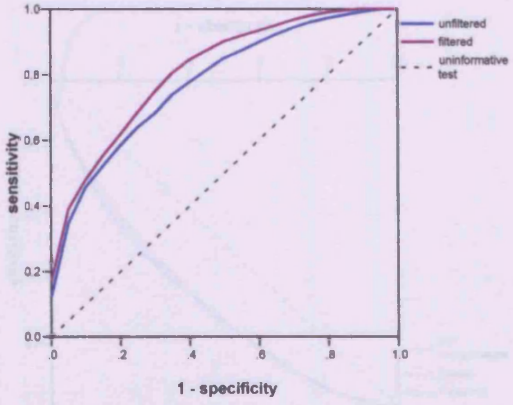


Figure A1.24: the ROC curve for testing using background mean, proportion and slope with and without filtering



A1.3.2 Abernant

Figure A1.18: the ROC curve for testing using background mean with and without filtering

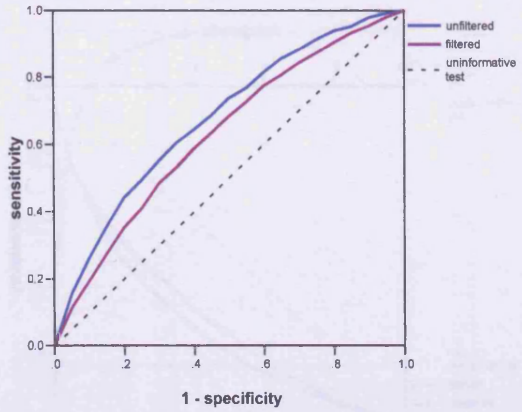


Figure A1.19: the ROC curve for testing using proportion with and without filtering

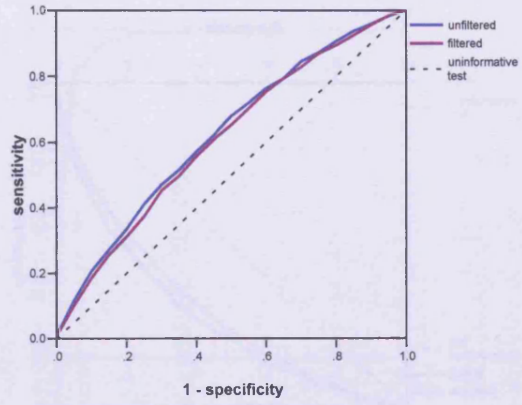


Figure A1.20: the ROC curve for testing using slope with and without filtering

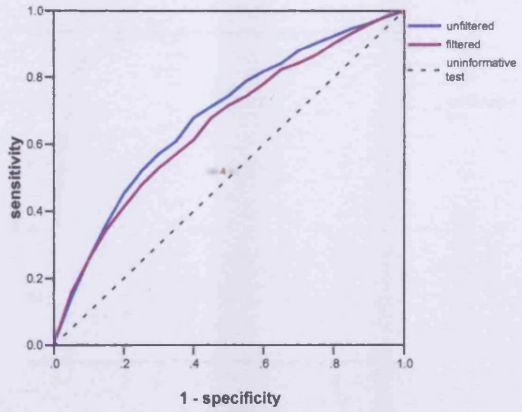
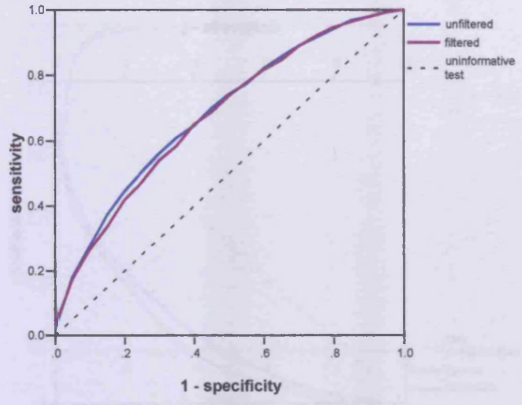


Figure A1.21: the ROC curve for testing using background mean and proportion with and without filtering



A1.5 Adjusting for Known Confounders

Figure A1.29: the ROC curve for Stone's test with no confounding and with adjustment for known confounding

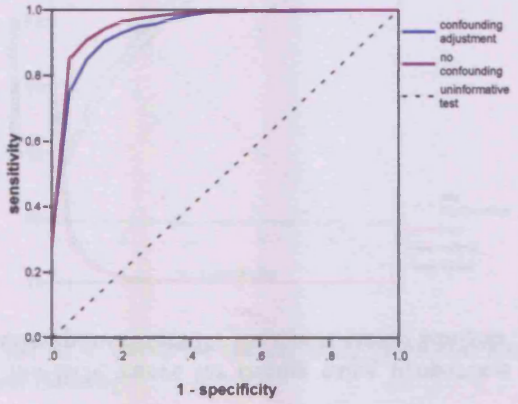


Figure A1.30: the ROC curve for the trend test with no confounding and with adjustment for known confounding

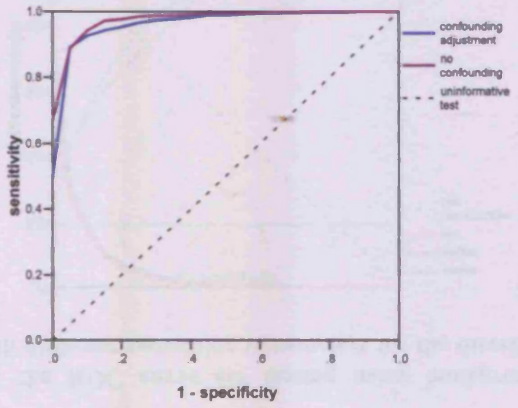


Figure A1.31: the ROC curve for testing using background mean with no confounding and with adjustment for known confounding

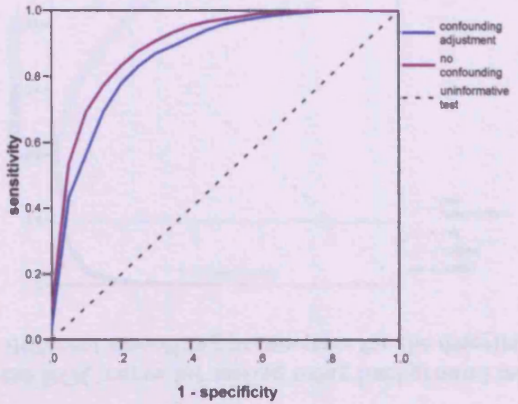
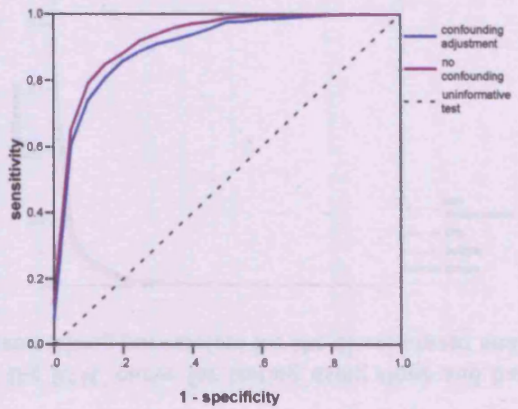


Figure A1.32: the ROC curve for testing using proportion with no confounding and with adjustment for known confounding



A1.4 Smoothing Parameters in Discriminant Analysis

Figure A1.25: the ROC curve for testing using background mean and proportion with different smoothing parameters for the discriminant analysis

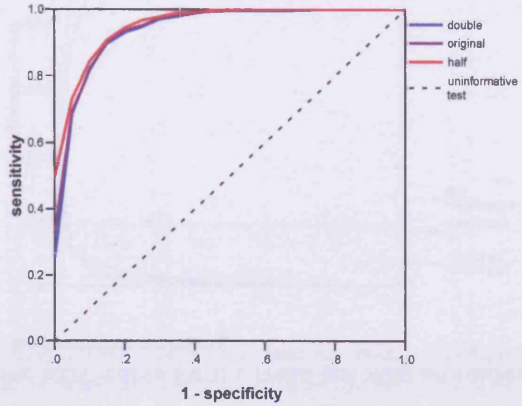


Figure A1.26: the ROC curve for testing using proportion and slope with different smoothing parameters for the discriminant analysis

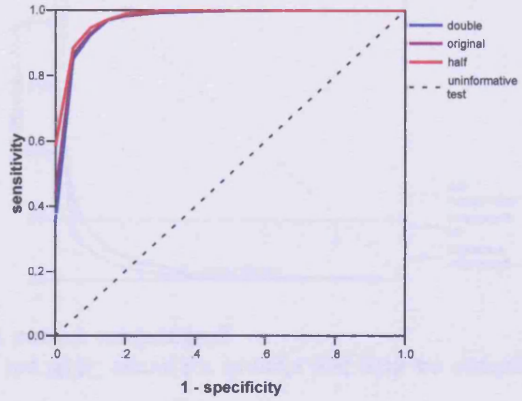


Figure A1.27: the ROC curve for testing using slope and background mean with different smoothing parameters for the discriminant analysis

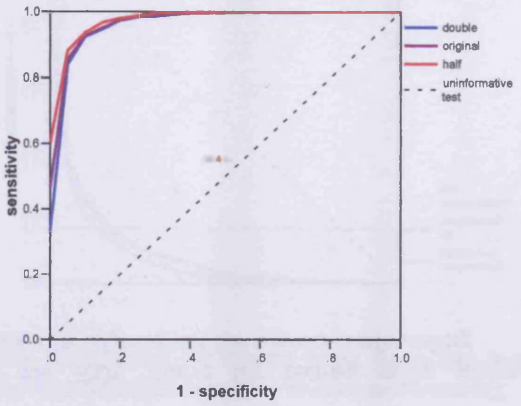


Figure A1.28: the ROC curve for testing using background mean, proportion and slope with different smoothing parameters for the discriminant analysis

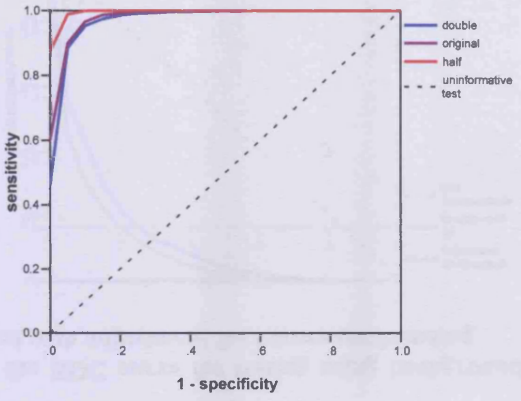
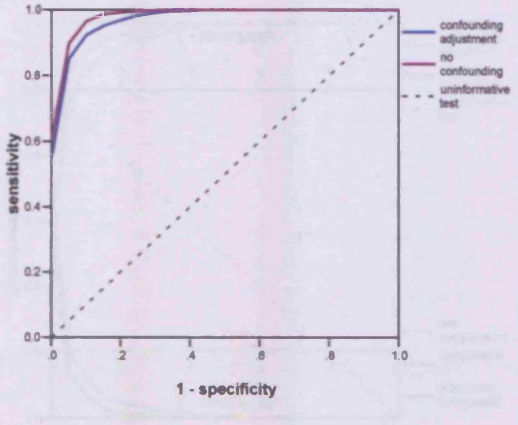


Figure A1.37: the ROC curve for testing using background mean, proportion and slope with no confounding and with adjustment for known confounding



A1.6 Effect Size

Figure A1.38: the ROC curve for Stone's test for different effect sizes

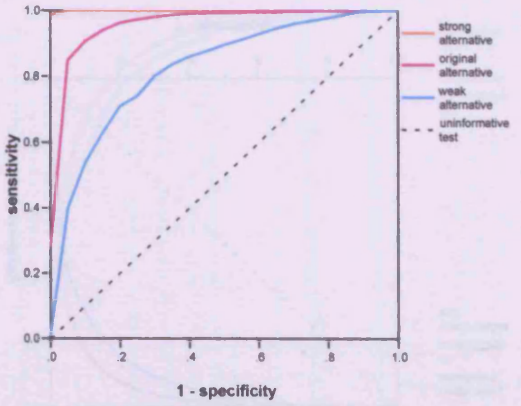


Figure A1.39: the ROC curve for the trend test for different effect sizes

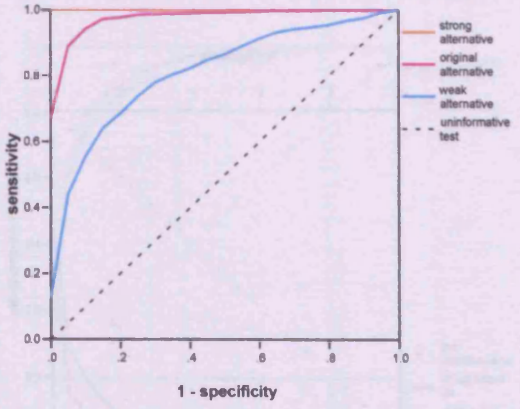


Figure A1.33: the ROC curve for testing using slope with no confounding and with adjustment for known confounding

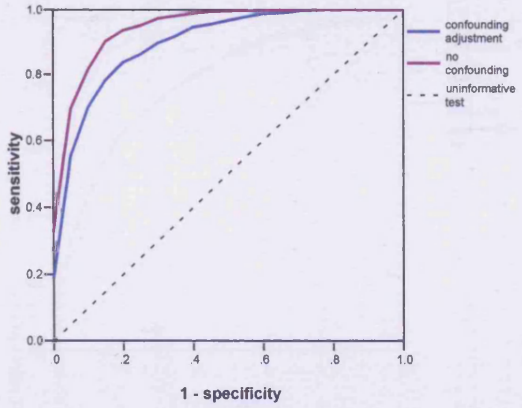


Figure A1.34: the ROC curve for testing using background mean and proportion with no confounding and with adjustment for known confounding

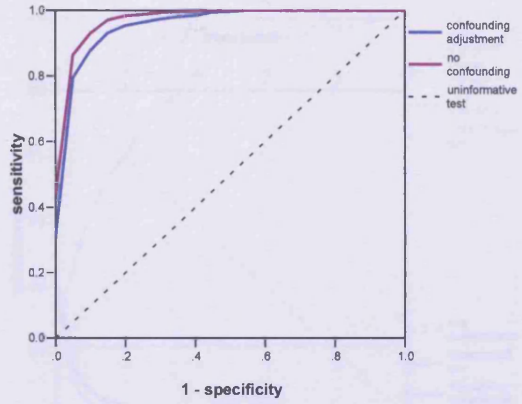


Figure A1.35: the ROC curve for testing using proportion and slope with no confounding and with adjustment for known confounding

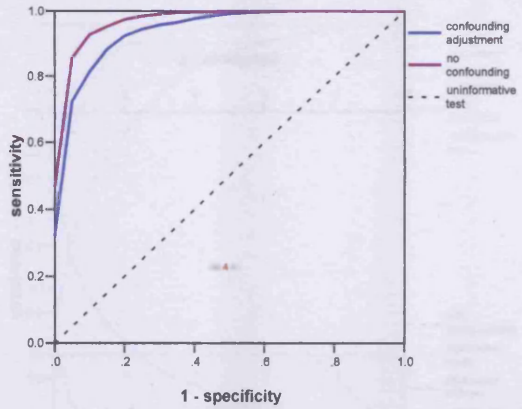


Figure A1.36: the ROC curve for testing using slope and background mean with no confounding and adjustment for known confounding

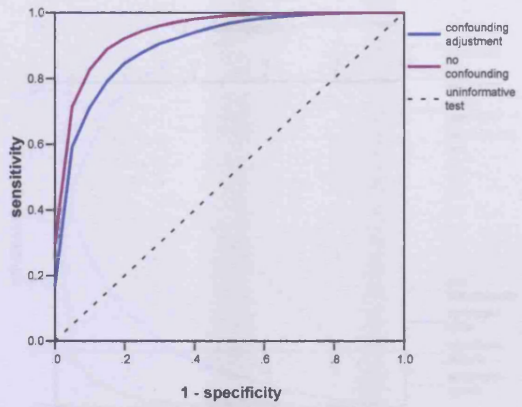


Figure A1.44: the ROC curve for testing using proportion and slope for different effect sizes

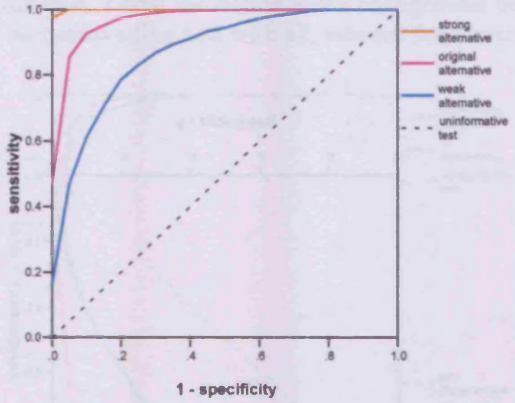


Figure A1.45: the ROC curve for testing using slope and background mean for different effect sizes

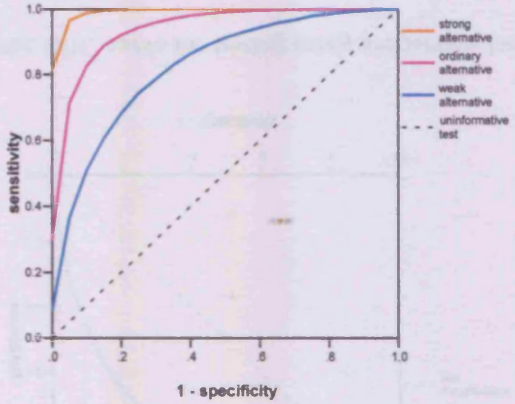


Figure A1.46: the ROC curve for testing using background mean, proportion and slope for different effect sizes

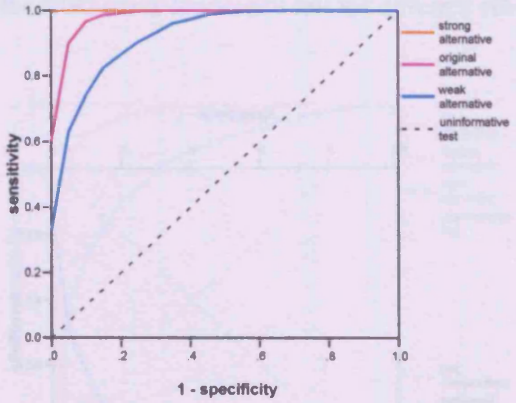


Figure A1.40: the ROC curve for testing using background mean for different effect sizes

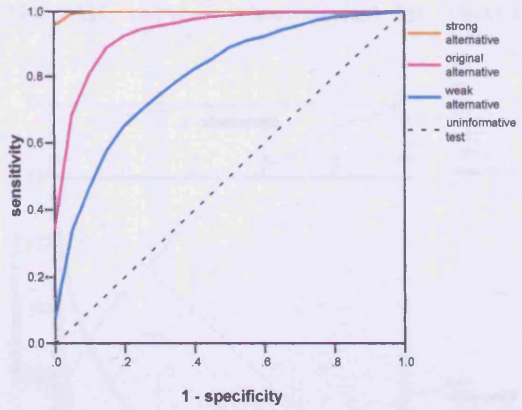


Figure A1.41: the ROC curve for testing using proportion for different effect sizes

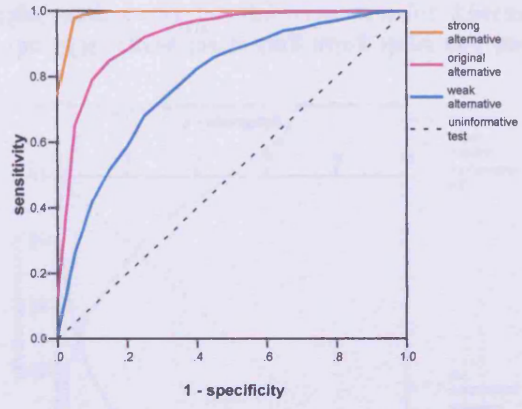


Figure A1.42: the ROC curve for testing using slope for different effect sizes

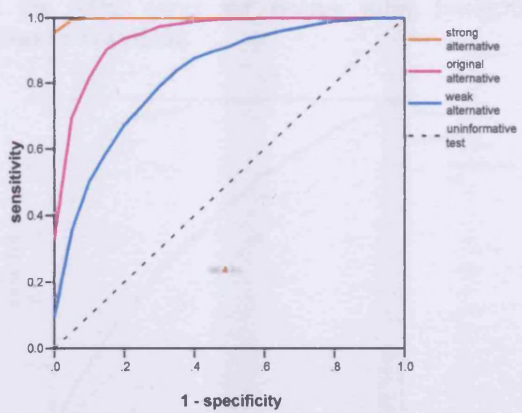


Figure A1.43: the ROC curve for testing using background mean and proportion for different effect sizes

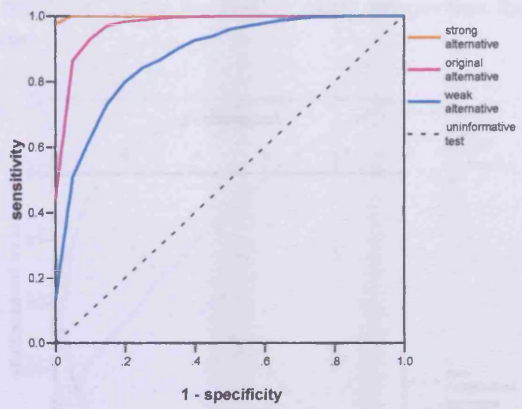


Figure A1.51: the ROC curve for testing using slope for Abernant and Nant y Gwyddon

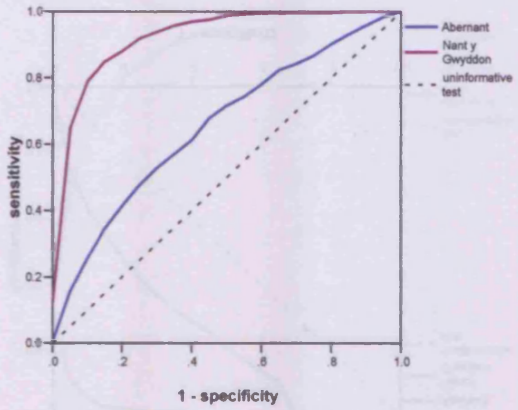


Figure A1.52: the ROC curve for testing using background mean and proportion for Abernant and Nant y Gwyddon

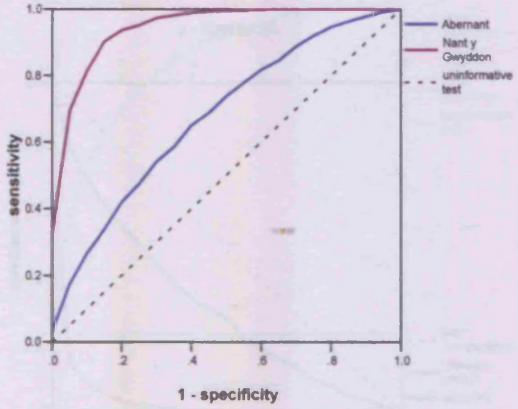


Figure A1.53: the ROC curve for testing using proportion and slope for Abernant and Nant y Gwyddon

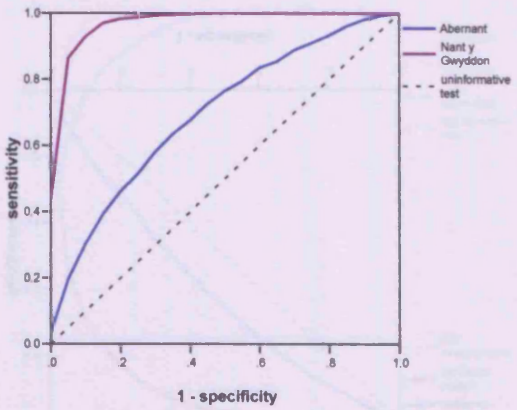
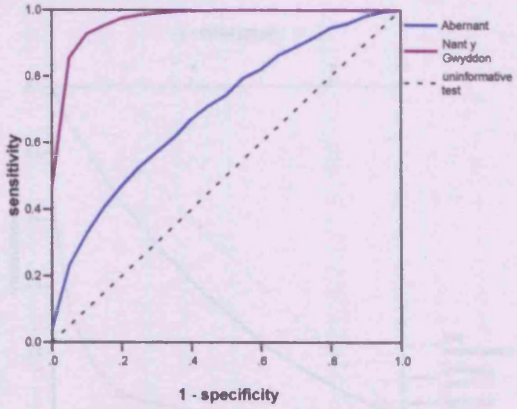


Figure A1.54: the ROC curve for testing using slope and background mean for Abernant and Nant y Gwyddon



A1.7 Population Size

Figure A1.47: the ROC curve for Stone's test for Abernant and Nant y Gwyddon

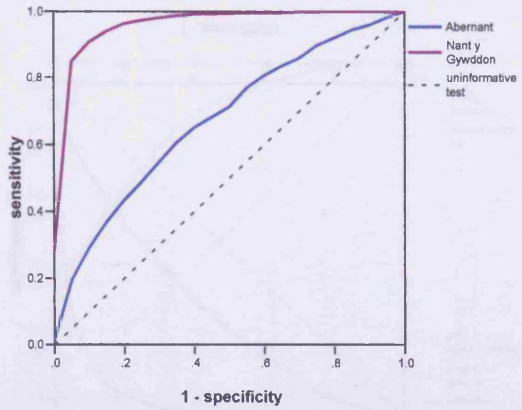


Figure A1.48: the ROC curve for the trend test for Abernant and Nant y Gwyddon

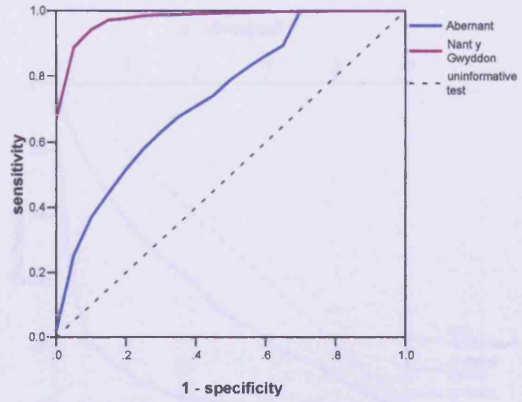


Figure A1.49: the ROC curve for testing using background mean for Abernant and Nant y Gwyddon

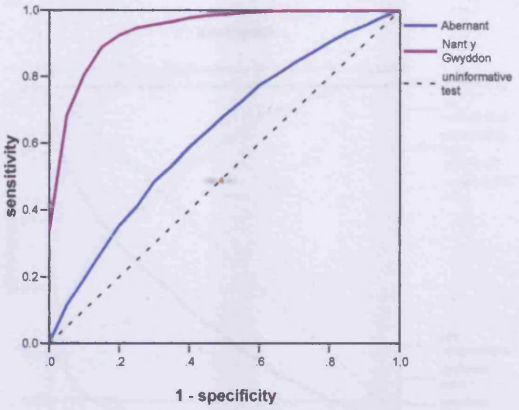


Figure A1.50: the ROC curve for testing using proportion for Abernant and Nant y Gwyddon

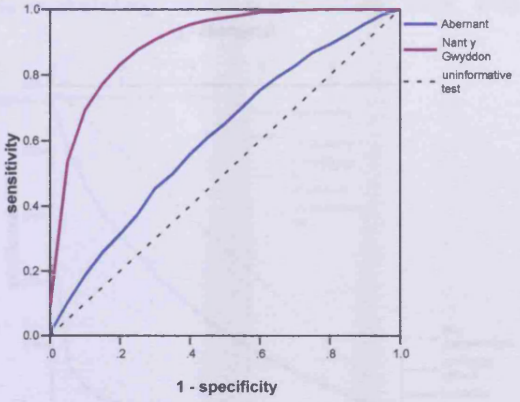


Figure A1.58: the ROC curve for testing using background mean comparing simulating with no confounding to simulating with confounding both with and without adjustment

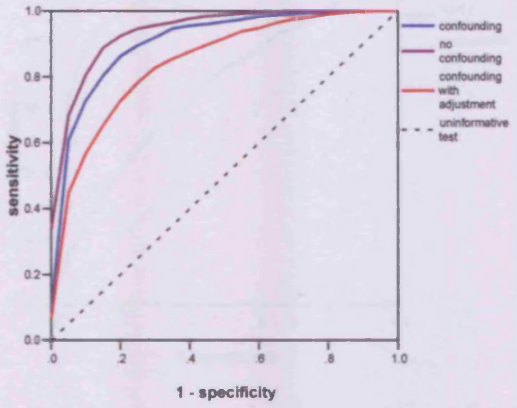


Figure A1.59: the ROC curve for testing using proportion comparing simulating with no confounding to simulating with confounding both with and without adjustment

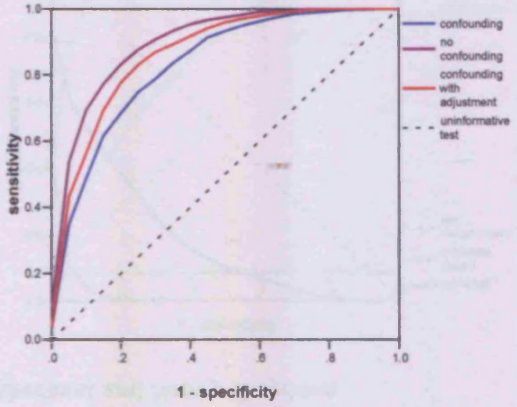


Figure A1.60: the ROC curve for testing using slope comparing simulating with no confounding to simulating with confounding both with and without adjustment

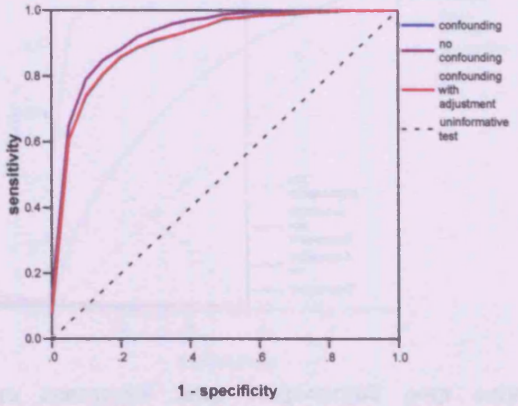


Figure A1.61: the ROC curve for testing using background mean and proportion comparing simulating with no confounding to simulating with confounding both with and without adjustment

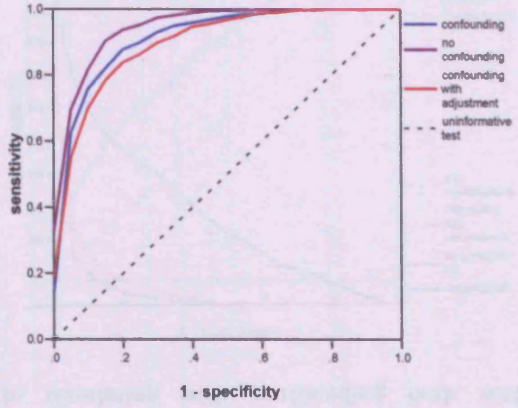
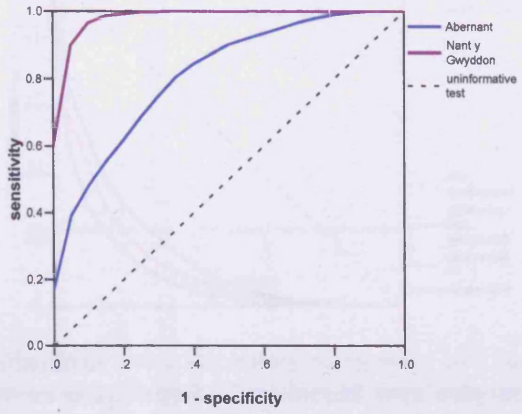


Figure A1.55: the ROC curve for testing using background mean, proportion and slope for Abernant and Nant y Gwyddon



A1.8 Unknown Confounding

Figure A1.56: the ROC curve for Stone's test comparing simulating with no confounding to simulating with confounding both with and without adjustment

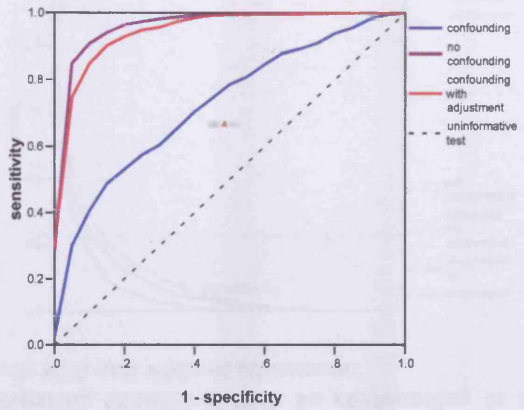
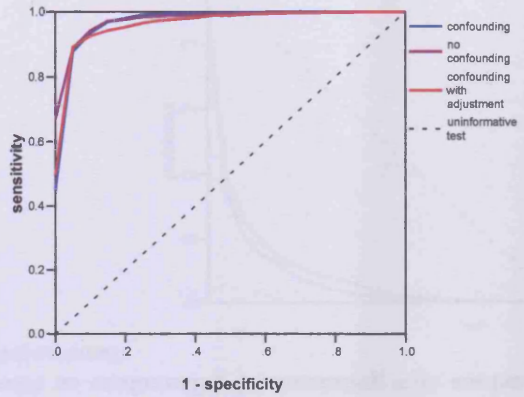


Figure A1.57: the ROC curve for the trend test comparing simulating with no confounding to simulating with confounding both with and without adjustment



A1.9 Misclassified Exposure

Figure A1.65: the ROC curve for Stone's test comparing correctly defined exposure to misclassified exposure

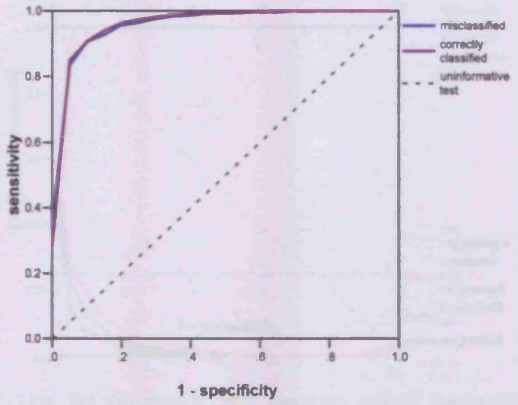


Figure A1.66: the ROC curve for the trend test comparing correctly defined exposure to misclassified exposure

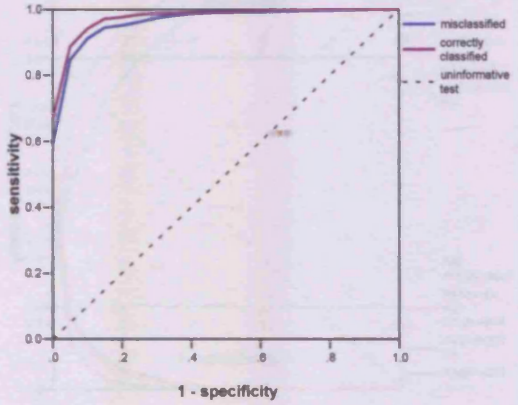


Figure A1.67: the ROC curve for testing using background mean comparing correctly defined exposure to misclassified exposure

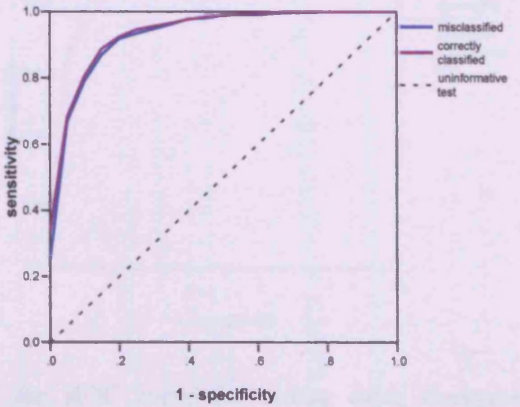


Figure A1.68: the ROC curve for testing using proportion comparing correctly defined exposure to misclassified exposure

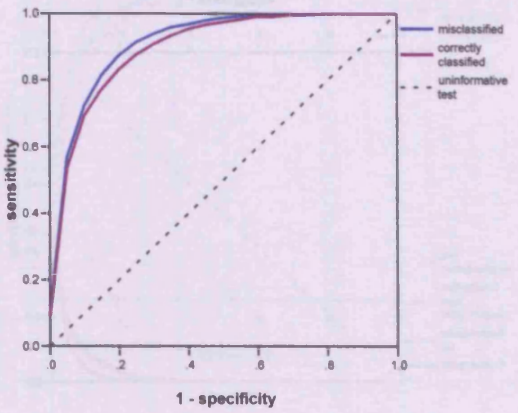


Figure A1.62: the ROC curve for testing using proportion and slope comparing simulating with no confounding to simulating with confounding both with and without adjustment

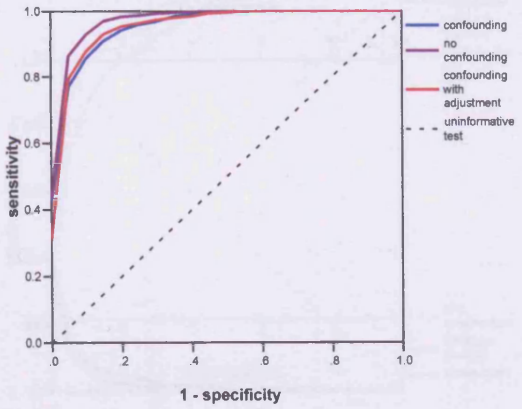


Figure A1.63: the ROC curve for testing using slope and background mean comparing simulating with no confounding to simulating with confounding both with and without adjustment

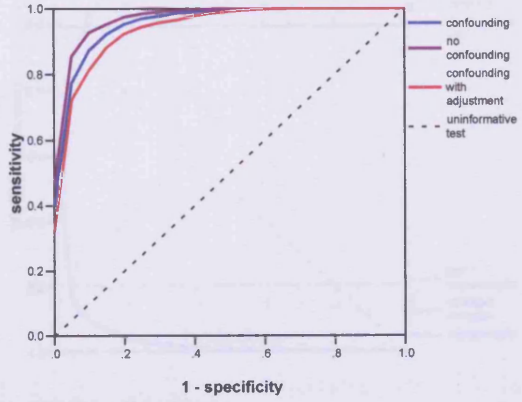
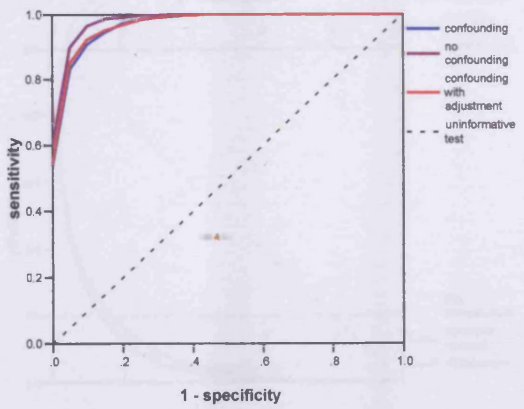


Figure A1.64: the ROC curve for testing using background mean, proportion and slope comparing simulating with no confounding to simulating with confounding both with and without adjustment



A1.9 Misclassified Exposure

Figure A1.65: the ROC curve for Stone's test comparing correctly defined exposure to misclassified exposure

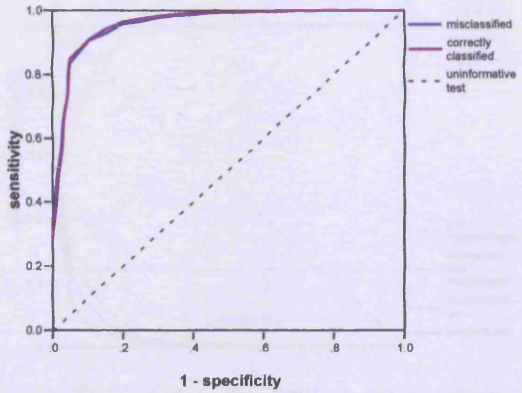


Figure A1.66: the ROC curve for the trend test comparing correctly defined exposure to misclassified exposure

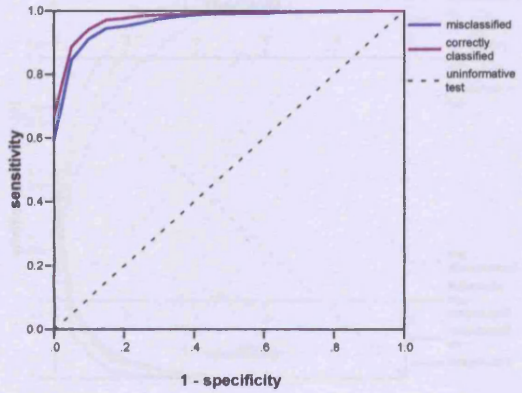


Figure A1.67: the ROC curve for testing using background mean comparing correctly defined exposure to misclassified exposure

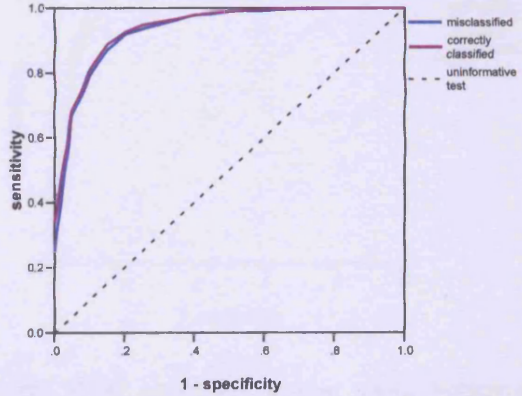


Figure A1.68: the ROC curve for testing using proportion comparing correctly defined exposure to misclassified exposure

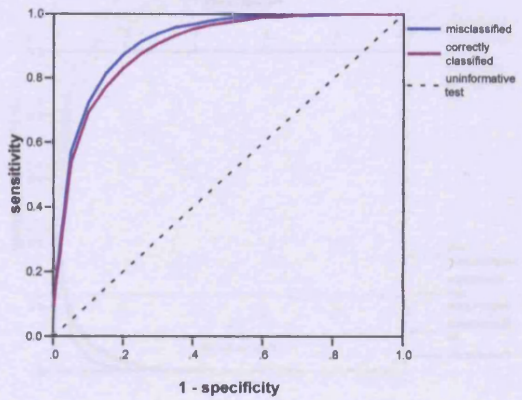


Figure A1.69: the ROC curve for testing using slope comparing correctly defined exposure to misclassified exposure

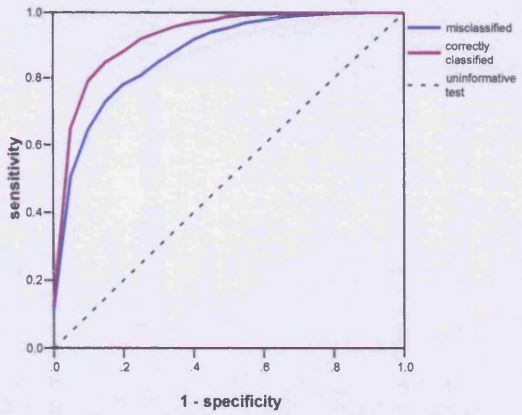


Figure A1.70: the ROC curve for testing using background mean and proportion comparing correctly defined exposure to misclassified exposure

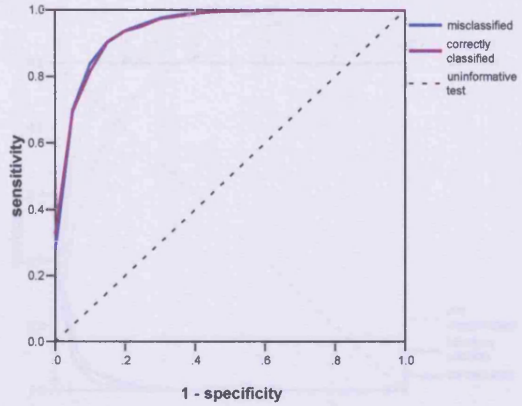


Figure A1.71: the ROC curve for testing using proportion and slope comparing correctly defined exposure to misclassified exposure

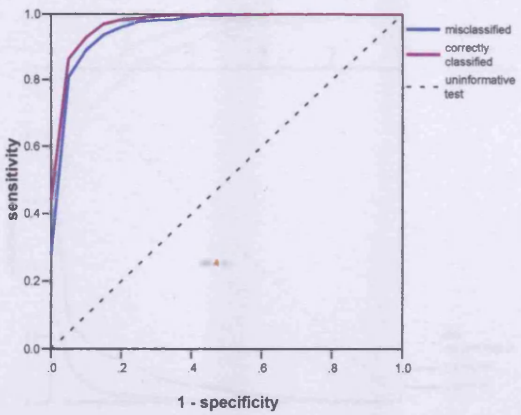


Figure A1.72: the ROC curve for testing using slope and background mean comparing correctly defined exposure to misclassified exposure

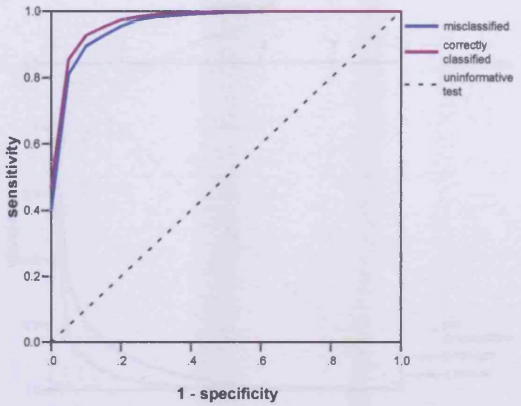
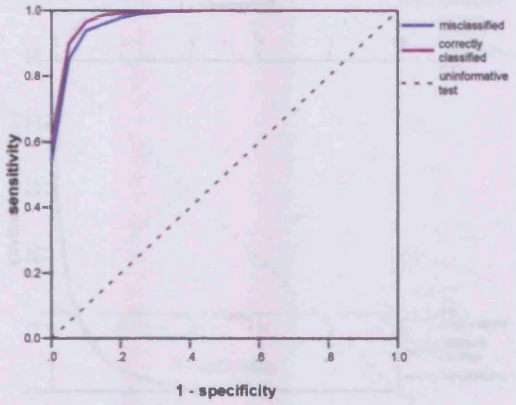


Figure A1.73: the ROC curve for testing using background mean, proportion and slope comparing correctly defined exposure to misclassified exposure



A1.10 Spatial Distribution of the Risk Factor

Figure A1.74: the ROC curve for Stone's test comparing two exposure sources to one exposure source

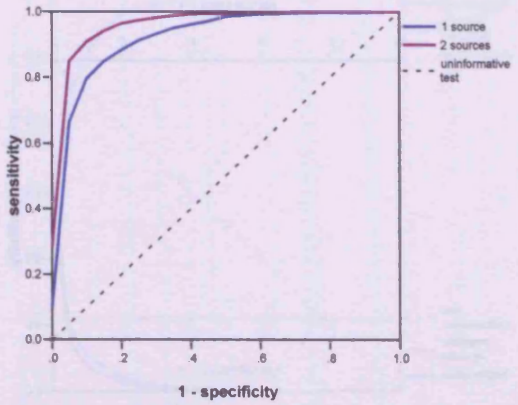


Figure A1.75: the ROC curve for the trend test comparing two exposure sources to one exposure source

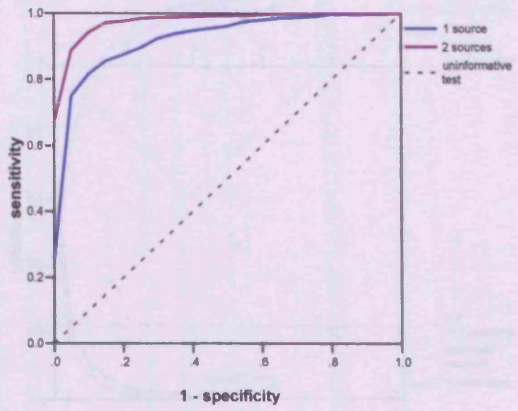


Figure A1.76: the ROC curve for testing using background mean comparing two exposure sources to one exposure source

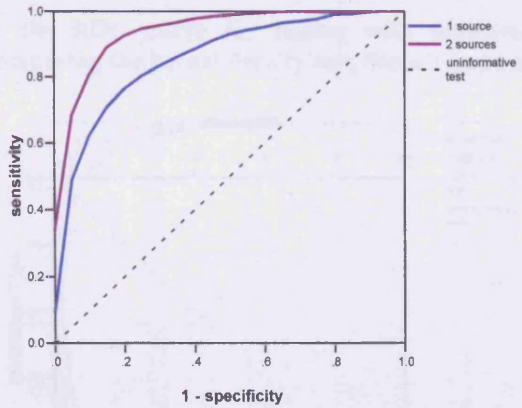


Figure A1.77: the ROC curve for testing using proportion comparing two exposure sources to one exposure source

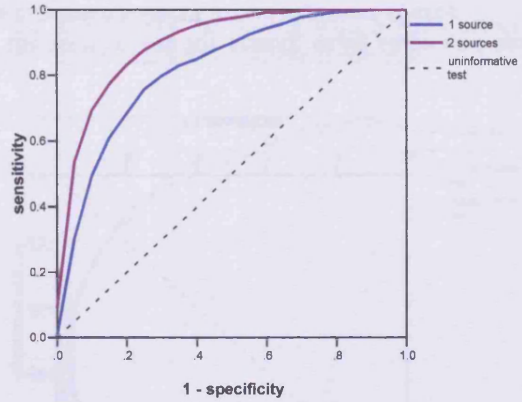


Figure A1.78: the ROC curve for testing using slope comparing two exposure sources to one exposure source

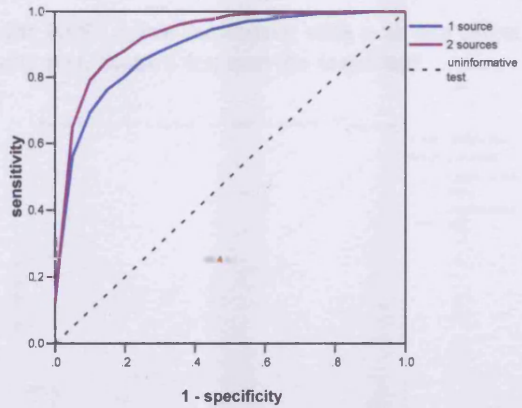


Figure A1.79: the ROC curve for testing using background mean and proportion comparing two exposure sources to one exposure source

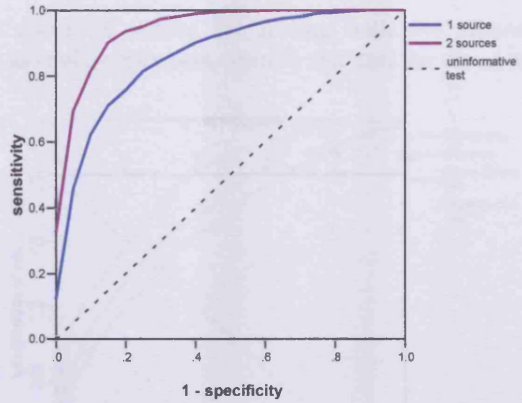


Figure A1.80: the ROC curve for testing using proportion and slope comparing two exposure sources to one exposure source

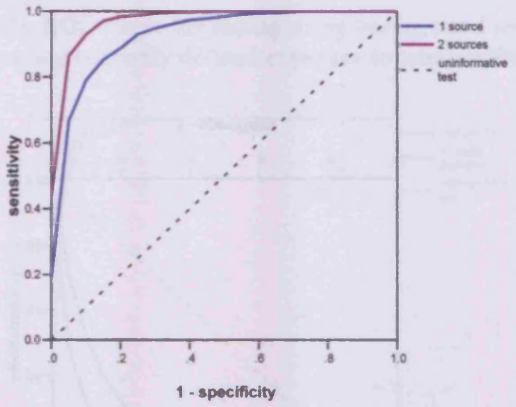


Figure A1.81: the ROC curve for testing using slope and background mean comparing two exposure sources to one exposure source

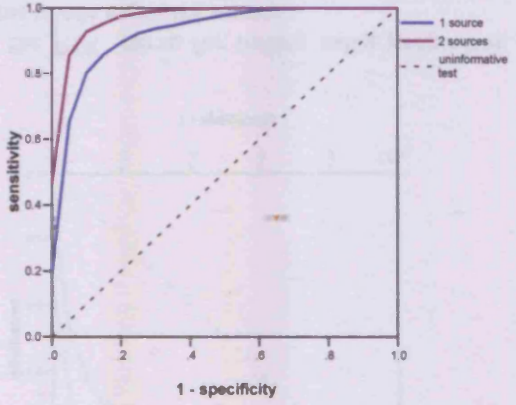


Figure A1.82: the ROC curve for testing using background mean, proportion and slope comparing two exposure sources to one exposure source

A1.11 Kernel Density Test, Stone's Test and the Trend Test

Figure A1.83: the ROC curve for testing with adjustment for known confounders comparing the kernel density test, Stone's test and the trend test

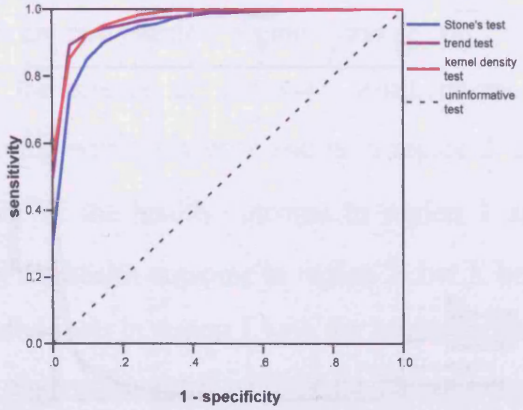


Figure A1.84: the ROC curve for testing with a weak effect size comparing the kernel density test, Stone's test and the trend test

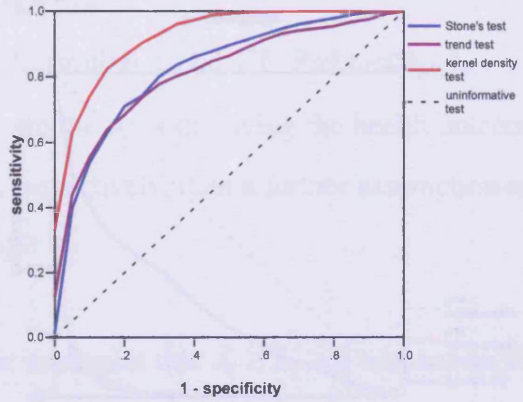


Figure A1.85: the ROC curve for testing with a strong effect size comparing the kernel density test, Stone's test and the trend test

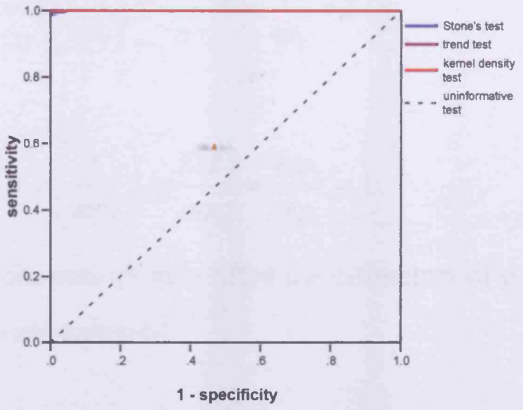


Figure A1.86: the ROC curve for testing with the Abernant population comparing the kernel density test, Stone's test and the trend test

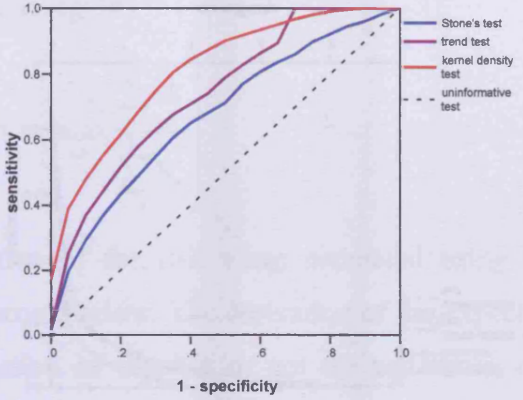


Figure A1.87: the ROC curve for testing with unknown confounding comparing the kernel density test, Stone's test and the trend test

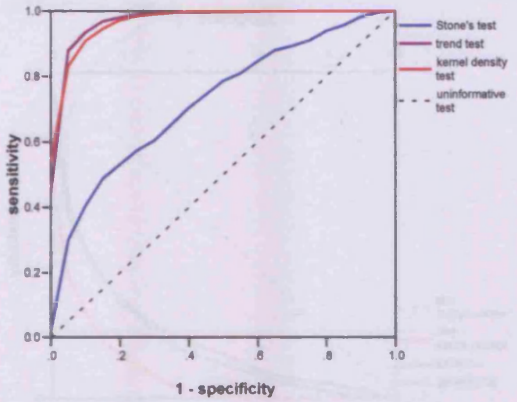


Figure A1.88: the ROC curve for testing with misclassified exposure contours comparing the kernel density test, Stone's test and the trend test

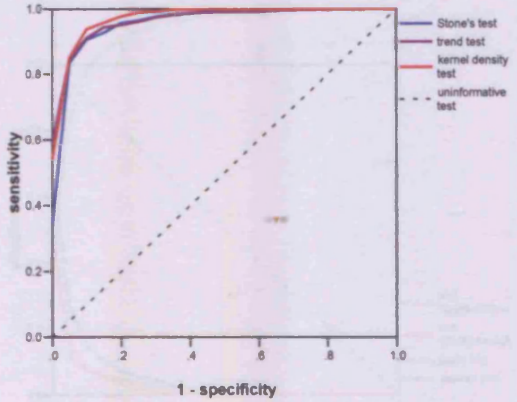
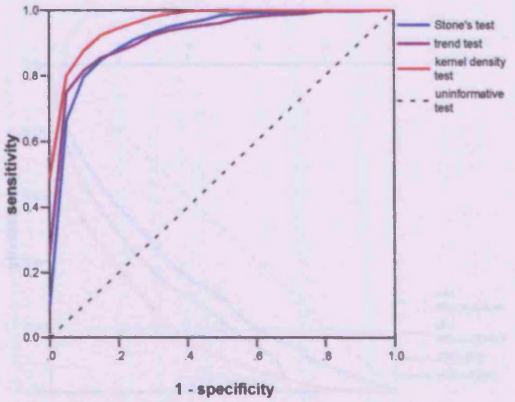


Figure A1.89: the ROC curve for testing with one exposure source comparing the kernel density test, Stone's test and the trend test



A1.12 Bias in Stone's Test

A1.12.1 Two Region Situation

Suppose there are two regions: region 1 and region 2 where region 1 is closer to the source of exposure. Furthermore, there are n individuals living within region 1 and m in region 2. Let p_1 denote the probability of the health outcome in region 1 and p_2 is the probability of the health outcome in region 2. Let X be the observed number of individuals in region 1 with the health outcome of interest and Y be the observed number with the health outcome in region 2.

In their test, Stone assumes that X and Y follow a Poisson distribution such that

$$X \sim Poisson(np_1) \text{ and } Y \sim Poisson(mp_2).$$

If λ_1 and λ_2 are the risks of having the health outcome in region 1 and region 2, respectively, then a further assumption made by Stone is that $\lambda_1 \geq \lambda_2 \geq 0$.

Supposing the constraint that $\lambda_1 \geq \lambda_2 \geq 0$ was not applied to the data then

$$\hat{\lambda}_1 = \frac{X}{np_1} \text{ and } \hat{\lambda}_2 = \frac{Y}{mp_2}.$$

X and Y are independent therefore

$$E(\hat{\lambda}_1) = E\left(\frac{X}{np_1}\right) = \frac{E(X)}{np_1} = \frac{np_1}{np_1} = 1$$

and

$$E(\hat{\lambda}_2) = E\left(\frac{Y}{mp_2}\right) = \frac{E(Y)}{mp_2} = \frac{mp_2}{mp_2} = 1.$$

Hence, when the constraint is lifted the estimators of the risk in each of the regions are unbiased.

However, in Stone's test the constraint that $\lambda_1 \geq \lambda_2 \geq 0$ is imposed and consequently the following minimax formula is used to estimate the risk for each region

$$\hat{\lambda}_i = \min_{s \leq i}\left(\max_{t \geq i} \frac{\sum_{r=s}^{t} O_r}{\sum_{r=s}^{t} E_r}\right).$$

The expectation of the risk when estimated using the minimax formula is derived below. The derivation of the expectation enables the determination of whether or not the estimation of the risk is biased.

For region 1, $i=1 \Rightarrow s=1$ and $t=1,2 \Rightarrow$

$$\hat{\lambda}_1 = \max\left(\frac{O_1}{E_1}, \frac{O_1+O_2}{E_1+E_2}\right) = \max\left(\frac{X}{np_1}, \frac{X+Y}{np_1+mp_2}\right).$$

For region 2, $i=2 \Rightarrow s=1,2$ and $t=2 \Rightarrow$

$$\hat{\lambda}_2 = \min\left(\frac{O_1+O_2}{E_1+E_2}, \frac{O_2}{E_2}\right) = \min\left(\frac{X+Y}{np_1+mp_2}, \frac{Y}{mp_2}\right).$$

There are two possible situations that can occur with only two regions: either the risk in region 1 is greater than or equal to the risk in region 2 or it is less than the risk in region 2.

Suppose the risk is higher in region 1 or the risk is equal in the two groups thus

$$\frac{X}{np_1} \geq \frac{Y}{mp_2} \Rightarrow X \geq \frac{Ynp_1}{mp_2} \text{ and } Y \leq \frac{Xmp_2}{np_1}.$$

Therefore,

$$\frac{X+Y}{np_1+mp_2} \leq \frac{X+Xmp_2/np_1}{np_1+mp_2} = \frac{X(np_1+mp_2)}{np_1(np_1+mp_2)} = \frac{X}{np_1}$$

and

$$\frac{X+Y}{np_1+mp_2} \geq \frac{Ynp_1/mp_2+Y}{np_1+mp_2} = \frac{Y(np_1+mp_2)}{mp_2(np_1+mp_2)} = \frac{Y}{mp_2}.$$

Hence,

$$\hat{\lambda}_1 = \max\left(\frac{X}{np_1}, \frac{X+Y}{np_1+mp_2}\right) = \frac{X}{np_1}$$

and

$$\hat{\lambda}_2 = \min\left(\frac{X+Y}{np_1+mp_2}, \frac{Y}{mp_2}\right) = \frac{Y}{mp_2}.$$

Suppose instead that the risk is higher in region 2 than in region 1 i.e.

$$\frac{X}{np_1} < \frac{Y}{mp_2} \Rightarrow X < \frac{Ynp_1}{mp_2} \text{ and } Y > \frac{Xmp_2}{np_1}.$$

Hence

$$\frac{X+Y}{np_1+mp_2} > \frac{X+Xmp_2/np_1}{np_1+mp_2} = \frac{X(np_1+mp_2)}{np_1(np_1+mp_2)} = \frac{X}{np_1}$$

and

$$\frac{X+Y}{np_1+mp_2} < \frac{Ynp_1/mp_2+Y}{np_1+mp_2} = \frac{Y(np_1+mp_2)}{mp_2(np_1+mp_2)} = \frac{Y}{mp_2}.$$

Consequently in this situation

$$\hat{\lambda}_1 = \max\left(\frac{X}{np_1}, \frac{X+Y}{np_1+mp_2}\right) = \frac{X+Y}{np_1+mp_2}$$

and

$$\hat{\lambda}_2 = \min\left(\frac{X+Y}{np_1+mp_2}, \frac{Y}{mp_2}\right) = \frac{X+Y}{np_1+mp_2}.$$

Combining these two possible situations enables the expectation to be calculated by conditioning on X

$$E(\hat{\lambda}_2) = E_X\left(E_Y(\hat{\lambda}_1|X)\right).$$

Let $p_y = P(Y=y)$, $p_x = P(X=x)$ and $u = \dfrac{Xmp_2}{np_1}$ so

$$E_Y(\hat{\lambda}_1|X) = \sum_{y=0}^{u}\frac{X}{np_1}p_y + \sum_{y=u+1}^{\infty}\frac{X+Y}{np_1+mp_2}p_y$$

$$= \frac{X}{np_1}\sum_{y=0}^{u}p_y + \frac{X}{np_1+mp_2}\sum_{y=u+1}^{\infty}p_y + \frac{1}{np_1+mp_2}\sum_{y=u+1}^{\infty}yp_y$$

Since $\sum_{y=0}^{\infty}yp_y = E(Y) = mp_2$

$$E_Y(\hat{\lambda}_1|X) = \frac{X}{np_1}\sum_{y=0}^{u}p_y + \frac{X}{np_1+mp_2}\left(1-\sum_{y=0}^{u}p_y\right)$$

$$+\frac{1}{np_1+mp_2}\left(mp_2-\sum_{y=0}^{u}yp_y\right)$$

$$= \frac{X}{np_1+mp_2} + X\left(\frac{1}{np_1}-\frac{1}{np_1+mp_2}\right)\sum_{y=0}^{u}p_y$$

$$+\frac{1}{np_1+mp_2}\left(mp_2-\sum_{y=0}^{u}yp_y\right)$$

Substituting this back in gives

$$E(\hat{\lambda}_1) = E_X\left(E_Y(\hat{\lambda}_1|X)\right)$$

$$= \frac{np_1}{np_1+mp_2} + \frac{mp_2}{(np_1+mp_2)np_1}\sum_{x=0}^{\infty}x\left(\sum_{y=0}^{u}p_y\right)p_x$$

$$+\frac{mp_2}{np_1+mp_2} - \frac{1}{np_1+mp_2}\sum_{x=0}^{\infty}\left(\sum_{y=0}^{u}yp_y\right)p_x$$

$$= \frac{np_1+mp_2}{np_1+mp_2} + \frac{mp_2}{np_1(np_1+mp_2)}\sum_{x=0}^{\infty}x\left(\sum_{y=0}^{u}p_y\right)p_x$$

$$-\frac{1}{np_1+mp_2}\sum_{x=0}^{\infty}\left(\sum_{y=0}^{u}yp_y\right)p_x$$

$$= 1 + \frac{mp_2}{np_1(np_1+mp_2)}\sum_{x=0}^{\infty}x\left(\sum_{y=0}^{u}p_y\right)p_x$$

$$-\frac{1}{np_1+mp_2}\sum_{x=0}^{\infty}\left(\sum_{y=0}^{u}yp_y\right)p_x$$

For $\hat{\lambda}_1 \geq \hat{\lambda}_2 \Rightarrow \dfrac{Xmp_2}{np_1} \geq Y$

$$\sum_{y=0}^{u}yp_y \leq \sum_{y=0}^{u}\frac{mp_2}{np_1}Xp_y = \frac{mp_2}{np_1}X\sum_{y=0}^{u}p_y.$$

Thus

$$E(\hat{\lambda}_1) \geq 1 + \frac{mp_2}{np_1(np_1+mp_2)}\sum_{x=0}^{\infty}x\left(\sum_{y=0}^{u}p_y\right)p_x$$

$$-\frac{mp_2}{np_1(np_1+mp_2)}\sum_{x=0}^{\infty}x\left(\sum_{y=0}^{u}p_y\right)p_x$$

$$=1.$$

Therefore, the expectation of the risk in region 1 is greater than or equal to 1 if the risk in region 1 is greater than or equal to that in region 2. Therefore, if there exists a simulation under the null hypothesis for which

$$\hat{\lambda}_1 > \hat{\lambda}_2$$

then the minimax formula used for estimating risk in Stone's test is biased. In the simulations under the null hypothesis performed to assess Stone's test n and m are 130 and 579, respectively, and $p_1 = p_2 = 0.01$.

In one simulation there were 3 cases in region 1 and 3 in region 2, thus,

$$\frac{Xmp_2}{np_1} = 13.362 > 3 = Y \Rightarrow \hat{\lambda}_1 > \hat{\lambda}_2.$$

Therefore, since the strict inequality is satisfied by at least one of the simulations

$$E(\hat{\lambda}_1) > 1$$

the minimax estimation of λ_1 is biased even under the null hypothesis.

A1.12.2 Three Region Situation

Introducing a third region creates a more complicated situation than with only two regions. If the same methodology was followed through for three regions as it was for two regions then the algebra becomes very complicated therefore an alternative approach has been considered for this situation.

Suppose the probability of the health outcome of interest was equal to p_3 in region 3, that q denotes the number of individuals in region 3 and that Z is the observed number of individuals with the health outcome.

In Stone's test it is assumed that

$$X \sim Poisson(np_1),\ Y \sim Poisson(mp_2) \text{ and}$$

$$Z \sim Poisson(qp_3).$$

Additionally, λ_1, λ_2 and λ_3 denote the risk of the health outcome in region 1, region 2 and region 3, respectively. Stone imposes a criterion upon these risks where $\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \hat{\lambda}_3$.

If this restriction is not enforced then

$$\hat{\lambda}_1 = \frac{X}{np_1},\ \hat{\lambda}_2 = \frac{Y}{mp_2} \text{ and } \hat{\lambda}_3 = \frac{Z}{qp_3}.$$

X, Y and Z are independent thus

$$E\left(\hat{\lambda}_1\right) = E\left(\frac{X}{np_1}\right) = \frac{E(X)}{np_1} = \frac{np_1}{np_1} = 1,$$

$$E\left(\hat{\lambda}_2\right) = E\left(\frac{Y}{mp_2}\right) = \frac{E(Y)}{mp_2} = \frac{mp_2}{mp_2} = 1$$

and

$$E\left(\hat{\lambda}_3\right) = E\left(\frac{Z}{qp_3}\right) = \frac{E(Z)}{qp_3} = \frac{qp_3}{qp_3} = 1.$$

Hence, without the restriction the estimators for λ_1, λ_2 and λ_3 are unbiased.

However, when performing Stone's test this constraint must be met therefore the above estimators of λ_1, λ_2 and λ_3 cannot be used and the minimax formula given by

$$\hat{\lambda}_i = \min_{s \leq i}\left(\max_{t \geq i}\frac{\sum_{r=s}^{t} O_r}{\sum_{r=s}^{t} E_r}\right)$$

is used instead.

In region 1 $i = 1 \Rightarrow s = 1$ and $t = 1, 2, 3 \Rightarrow$

$$\hat{\lambda}_1 = \max\left(\frac{O_1}{E_1}, \frac{O_1 + O_2}{E_1 + E_2}, \frac{O_1 + O_2 + O_3}{E_1 + E_2 + E_3}\right)$$

$$= \max\left(\frac{X}{np_1}, \frac{X + Y}{np_1 + mp_2}, \frac{X + Y + Z}{np_1 + mp_2 + qp_3}\right)$$

In region 2 $i = 2 \Rightarrow s = 1, 2$ and $t = 2, 3$

$$\hat{\lambda}_2 = \min\left(\max\left(\frac{O_1 + O_2}{E_1 + E_2}, \frac{O_1 + O_2 + O_3}{E_1 + E_2 + E_3}\right), \max\left(\frac{O_2}{E_2}, \frac{O_2 + O_3}{E_2 + E_3}\right)\right)$$

$$= \min\left(\max\left(\frac{X + Y}{np_1 + mp_2}, \frac{X + Y + Z}{np_1 + mp_2 + qp_3}\right), \max\left(\frac{Y}{mp_2}, \frac{Y + Z}{mp_2 + qp_3}\right)\right)$$

Finally in region 3 $i = 3 \Rightarrow s = 1, 2, 3$ and $t = 3$

$$\hat{\lambda}_3 = \min\left(\frac{O_1 + O_2 + O_3}{E_1 + E_2 + E_3}, \frac{O_2 + O_3}{E_2 + E_3}, \frac{O_3}{E_3}\right)$$

$$= \min\left(\frac{X+Y+Z}{np_1+mp_2+qp_3}, \frac{Y+Z}{mp_2+qp_3}, \frac{Z}{qp_3}\right).$$

When determining $\hat{\lambda}_1$ there are six possibilities

1. $\frac{X}{np_1} \le \frac{X+Y}{np_1+mp_2} \le \frac{X+Y+Z}{np_1+mp_2+qp_3}$
2. $\frac{X+Y}{np_1+mp_2} \le \frac{X}{np_1} \le \frac{X+Y+Z}{np_1+mp_2+qp_3}$
3. $\frac{X}{np_1} \le \frac{X+Y+Z}{np_1+mp_2+qp_3} \le \frac{X+Y}{np_1+mp_2}$
4. $\frac{X+Y+Z}{np_1+mp_2+qp_3} \le \frac{X}{np_1} \le \frac{X+Y}{np_1+mp_2}$
5. $\frac{X+Y+Z}{np_1+mp_2+qp_3} \le \frac{X+Y}{np_1+mp_2} \le \frac{X}{np_1}$
6. $\frac{X+Y}{np_1+mp_2} \le \frac{X+Y+Z}{np_1+mp_2+qp_3} \le \frac{X}{np_1}$

Each of these six cases is considered below.

Suppose that

$$\frac{X}{np_1} \le \frac{X+Y}{np_1+mp_2} \le \frac{X+Y+Z}{np_1+mp_2+qp_3}$$

then clearly

$$\hat{\lambda}_1 = \frac{X+Y+Z}{np_1+mp_2+qp_3}.$$

In region 2

$$\hat{\lambda}_2 = \min\left(\max\left(\frac{X+Y}{np_1+mp_2}, \hat{\lambda}_1\right), \max\left(\frac{Y}{mp_2}, \frac{Y+Z}{mp_2+qp_3}\right)\right).$$

Since $\hat{\lambda}_1 \ge \frac{X+Y}{np_1+mp_2}$

$$\hat{\lambda}_2 = \min\left(\hat{\lambda}_1, \max\left(\frac{Y}{mp_2}, \frac{Y+Z}{mp_2+qp_3}\right)\right) \le \hat{\lambda}_1.$$

Hence, $\hat{\lambda}_1 \ge \hat{\lambda}_2$.

For region 3

$$\hat{\lambda}_3 = \min\left(\hat{\lambda}_1, \frac{Y+Z}{mp_2+qp_3}, \frac{Z}{qp_3}\right)$$

If $\frac{Y+Z}{mp_2+qp_3} \ge \frac{Y}{mp_2}$ then

$$\hat{\lambda}_2 = \min\left(\hat{\lambda}_1, \frac{Y+Z}{mp_2+qp_3}\right) \ge \min\left(\hat{\lambda}_1, \frac{Y+Z}{mp_2+qp_3}, \frac{Z}{qp_3}\right) = \hat{\lambda}_3$$

If $\frac{Y}{mp_2} \ge \frac{Y+Z}{mp_2+qp_3}$ then

$$\hat{\lambda}_2 = \min\left(\hat{\lambda}_1, \frac{Y}{mp_2}\right) \ge \min\left(\hat{\lambda}_1, \frac{Y+Z}{mp_2+qp_3}\right) \ge \hat{\lambda}_3$$

Thus $\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \hat{\lambda}_3$ for case 1.

Suppose that

$$\frac{X+Y}{np_1+mp_2} \leq \frac{X}{np_1} \leq \frac{X+Y+Z}{np_1+mp_2+qp_3}$$

then again

$$\hat{\lambda}_1 = \frac{X+Y+Z}{np_1+mp_2+qp_3}.$$

Hence

$$\hat{\lambda}_2 = \min\left(\max\left(\frac{X+Y}{np_1+mp_2}, \hat{\lambda}_1\right), \max\left(\frac{Y}{mp_2}, \frac{Y+Z}{mp_2+qp_3}\right)\right)$$

and because $\hat{\lambda}_1 \geq \frac{X+Y}{np_1+mp_2}$ this can be simplified to

$$\hat{\lambda}_2 = \min\left(\hat{\lambda}_1, \max\left(\frac{Y}{mp_2}, \frac{Y+Z}{mp_2+qp_3}\right)\right) \leq \hat{\lambda}_1$$

so $\hat{\lambda}_1 \geq \hat{\lambda}_2$.

For region 3

$$\hat{\lambda}_3 = \min\left(\hat{\lambda}_1, \frac{Y+Z}{mp_2+qp_3}, \frac{Z}{qp_3}\right)$$

as before so $\hat{\lambda}_2 \geq \hat{\lambda}_3$.

Hence, for case 2 $\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \hat{\lambda}_3$.

Assume that

$$\frac{X}{np_1} \leq \frac{X+Y+Z}{np_1+mp_2+qp_3} \leq \frac{X+Y}{np_1+mp_2}$$

so

$$\hat{\lambda}_1 = \frac{X+Y}{np_1+mp_2}.$$

Therefore, in region 2

$$\hat{\lambda}_2 = \min\left(\max\left(\hat{\lambda}_1, \frac{X+Y+Z}{np_1+mp_2+qp_3}\right), \max\left(\frac{Y}{mp_2}, \frac{Y+Z}{mp_2+qp_3}\right)\right).$$

Clearly

$$\hat{\lambda}_1 \geq \frac{X+Y+Z}{np_1+mp_2+qp_3}$$

thus

$$\hat{\lambda}_2 = \min\left(\hat{\lambda}_1, \max\left(\frac{Y}{mp_2}, \frac{Y+Z}{mp_2+qp_3}\right)\right) \leq \hat{\lambda}_1$$

and consequently $\hat{\lambda}_1 \geq \hat{\lambda}_2$.

By definition

$$\hat{\lambda}_3 = \min\left(\frac{X+Y+Z}{np_1+mp_2+qp_3}, \frac{Y+Z}{mp_2+qp_3}, \frac{Z}{qp_3}\right)$$

If $\frac{Y+Z}{mp_2+qp_3} \geq \frac{Y}{mp_2}$ then

$$\hat{\lambda}_2 = \min\left(\hat{\lambda}_1, \frac{Y+Z}{mp_2+qp_3}\right) \geq \min\left(\frac{X+Y+Z}{np_1+mp_2+qp_3}, \frac{Y+Z}{mp_2+qp_3}\right)$$

$$\geq \min\left(\frac{X+Y+Z}{np_1+mp_2+qp_3}, \frac{Y+Z}{mp_2+qp_3}, \frac{Z}{qp_3}\right) = \hat{\lambda}_3.$$

If $\frac{Y}{mp_2} \geq \frac{Y+Z}{mp_2+qp_4}$ then

$$\hat{\lambda}_2 = \min\left(\hat{\lambda}_1, \frac{Y}{mp_2}\right) \geq \min\left(\hat{\lambda}_1, \frac{Y+Z}{mp_2+qp_3}\right) \geq \hat{\lambda}_3$$

Therefore, $\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \hat{\lambda}_3$ always holds for case 3.

Suppose that

$$\frac{X+Y+Z}{np_1+mp_2+qp_3} \leq \frac{X}{np_1} \leq \frac{X+Y}{np_1+mp_2}$$

which means that

$$\hat{\lambda}_1 = \frac{X+Y}{np_1+mp_2}.$$

Thus

$$\hat{\lambda}_2 = \min\left(\max\left(\hat{\lambda}_1, \frac{X+Y+Z}{np_1+mp_2+qp_3}\right), \max\left(\frac{Y}{mp_2}, \frac{Y+Z}{mp_2+qp_3}\right)\right).$$

and since

$$\hat{\lambda}_1 \geq \frac{X+Y+Z}{np_1+mp_2+qp_3}$$

$$\hat{\lambda}_2 = \min\left(\hat{\lambda}_1, \max\left(\frac{Y}{mp_2}, \frac{Y+Z}{mp_2+qp_3}\right)\right) \leq \hat{\lambda}_1.$$

For region 3

$$\hat{\lambda}_3 = \min\left(\frac{X+Y+Z}{np_1+mp_2+qp_3}, \frac{Y+Z}{mp_2+qp_3}, \frac{Z}{qp_3}\right)$$

which similarly to in case 3 implies that $\hat{\lambda}_2 \geq \hat{\lambda}_3$.

Consequently, in case 4 $\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \hat{\lambda}_3$ holds.

Assume

$$\frac{X+Y+Z}{np_1+mp_2+qp_3} \leq \frac{X+Y}{np_1+mp_2} \leq \frac{X}{np_1}$$

then

$$\hat{\lambda}_1 = \frac{X}{np_1}.$$

In region 2

$$\hat{\lambda}_2 = \min\left(\max\left(\frac{X+Y}{np_1+mp_2}, \frac{X+Y+Z}{np_1+mp_2+qp_3}\right), \max\left(\frac{Y}{mp_2}, \frac{Y+Z}{mp_2+qp_3}\right)\right)$$

$$= \min\left(\frac{X+Y}{np_1+mp_2}, \max\left(\frac{Y}{mp_2}, \frac{Y+Z}{mp_2+qp_3}\right)\right)$$

$$\le \min\left(\hat{\lambda}_1, \max\left(\frac{Y}{mp_2}, \frac{Y+Z}{mp_2+qp_3}\right)\right) \le \hat{\lambda}_1 .$$

Here in region 3

$$\hat{\lambda}_3 = \min\left(\frac{X+Y+Z}{np_1+mp_2+qp_3}, \frac{Y+Z}{mp_2+qp_3}, \frac{Z}{qp_3}\right)$$

If $\frac{Y+Z}{mp_2+qp_3} \ge \frac{Y}{mp_2}$ then

$$\hat{\lambda}_2 = \min\left(\frac{X+Y}{np_1+mp_2}, \frac{Y+Z}{mp_2+qp_3}\right)$$

$$\ge \min\left(\frac{X+Y+Z}{np_1+mp_2+qp_3}, \frac{Y+Z}{mp_2+qp_3}\right) \ge \hat{\lambda}_3$$

If $\frac{Y}{mp_2} \ge \frac{Y+Z}{mp_2+qp_4}$ then

$$\hat{\lambda}_2 = \min\left(\frac{X+Y}{np_1+mp_2}, \frac{Y}{mp_2}\right) \ge \min\left(\frac{X+Y}{np_1+mp_2}, \frac{Y+Z}{mp_2+qp_3}\right)$$

$$\ge \hat{\lambda}_3 .$$

Hence for case 5 $\hat{\lambda}_1 \ge \hat{\lambda}_2 \ge \hat{\lambda}_3$.

Finally assume that

$$\frac{X+Y}{np_1+mp_2} \le \frac{X+Y+Z}{np_1+mp_2+qp_3} \le \frac{X}{np_1}$$

then as before

$$\hat{\lambda}_1 = \frac{X}{np_1} .$$

In region 2

$$\hat{\lambda}_2 = \min\left(\max\left(\frac{X+Y}{np_1+mp_2}, \frac{X+Y+Z}{np_1+mp_2+qp_3}\right), \max\left(\frac{Y}{mp_2}, \frac{Y+Z}{mp_2+qp_3}\right)\right)$$

$$= \min\left(\frac{X+Y+Z}{np_1+mp_2+qp_3}, \max\left(\frac{Y}{mp_2}, \frac{Y+Z}{mp_2+qp_3}\right)\right)$$

$$\le \min\left(\hat{\lambda}_1, \max\left(\frac{Y}{mp_2}, \frac{Y+Z}{mp_2+qp_3}\right)\right) \le \hat{\lambda}_1 .$$

For region 3

$$\hat{\lambda}_3 = \min\left(\frac{X+Y+Z}{np_1+mp_2+qp_3}, \frac{Y+Z}{mp_2+qp_3}, \frac{Z}{qp_3}\right)$$

If $\frac{Y+Z}{mp_2+qp_3} \ge \frac{Y}{mp_2}$ then

$$\hat{\lambda}_2 = \min\left(\frac{X+Y+Z}{np_1+mp_2+qp_3}, \frac{Y+Z}{mp_2+qp_3}\right) \ge \hat{\lambda}_3 .$$

If $\frac{Y}{mp_2} \geq \frac{Y+Z}{mp_2+qp_4}$ then

$$\hat{\lambda}_2 = \min\left(\frac{X+Y+Z}{np_1+mp_2+qp_3}, \frac{Y}{mp_2}\right)$$

$$\geq \min\left(\frac{X+Y+Z}{np_1+mp_2+qp_3}, \frac{Y+Z}{mp_2+qp_3}\right) \geq \hat{\lambda}_3.$$

Thus for case 6 and all other cases $\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \hat{\lambda}_3$ holds.

The criterion holds for all possible combinations for λ_1 therefore by using the minimax formula this criterion is enforced. If there is a situation for which the criterion becomes a strict inequality then the estimators used in Stone's test for the risk in each of the groups is biased.

In one of the simulations under the null hypothesis

$$\frac{X}{np_1} < \frac{X+Y}{np_1+mp_2} < \frac{X+Y+Z}{np_1+mp_2+qp_3}$$

since n = 130, m = 579 and q = 1542 with $p_1 = p_2 = p_3 = 0.01$ and there are 0 simulated cases in region 1, 2 in region 2 and 18 in region 3. Therefore,

$$\frac{X}{np_1} = 0, \quad \frac{X+Y}{np_1+mp_2} = 0.282 \text{ and } \frac{X+Y+Z}{np_1+mp_2+qp_3} = 0.888.$$

Thus it follows that

$$\frac{X}{np_1} < \frac{X+Y}{np_1+mp_2} < \frac{X+Y+Z}{np_1+mp_2+qp_3}$$

and consequently

$$\hat{\lambda}_1 > \hat{\lambda}_2 > \hat{\lambda}_3$$

for at least one simulation thus the estimators are biased.

Appendix Two: Density Estimation

A2.1 Density Estimation

A2.1.1 Definition of a Probability Density Function

Probability density functions denoted by $f(x)$ summarise the properties of the random variable X. The probability that the random variable X lies in the interval $[A, B]$, where A and B are constants, is $P(A < X < B) = \int_A^B f(x)dx$. If $f(x)$ is known then so are the properties of the random variable X.

A2.1.2 Mathematical Definition of Histograms

Suppose the origin is x_0 and the sample space is split into intervals of length h called bins then the bins are defined by the intervals $[x_0 + mh, x_0 + (m+1)h)$ for positive and negative integers, m.

The histogram is defined by

$$\hat{f}(x) = \frac{1}{nh}\{\text{number of } X_i \text{ 's in the same bin as x}\}$$

$$= \frac{1}{nh}\sum_{i=1}^{n} I(x - Xm_i, h)$$

where Xm_i is the midpoint of the bin containing X_i and $I(y;h)$ is the indicator function defined as

$$I(y;h) = \begin{cases} 1 & \text{if } |y| < h \\ 0 & \text{otherwise} \end{cases}.$$

If there is a low density of data over some areas of the range this definition of a histogram can be generalised for unequal bandwidths (bin widths) where

$$\hat{f}(x) = \frac{1}{n} \times \left\{ \frac{\text{number of } X_i \text{ 's in the same bin as x}}{\text{width of bin containing x}} \right\}$$

$$= \frac{1}{n} \sum_{i=1}^{n} \frac{I(x - Xm_i, h_i)}{h_i}.$$

A2.1.3 Properties of Histograms

Figure A2.1: the histogram for the distance of births from Nant y Gwyddon in the circular region with a 7km radius centred over the Nant y Gwyddon site (origin = 0km, number of bins = 14)

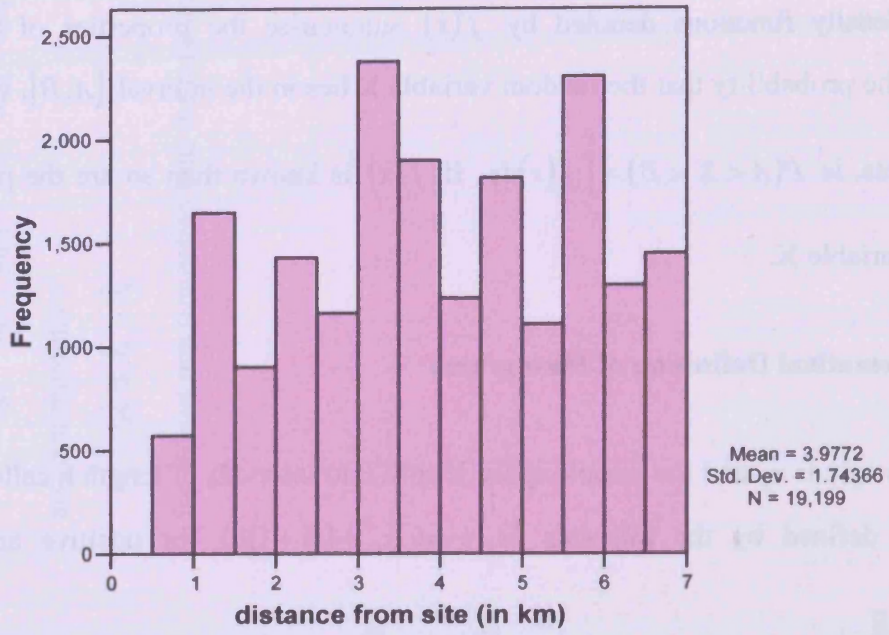


Figure A2.2: the histogram for the distance of births from Nant y Gwyddon in the circular region with radius 7km centred over the Nant y Gwyddon site (origin = 0.25km, number of bins = 14)

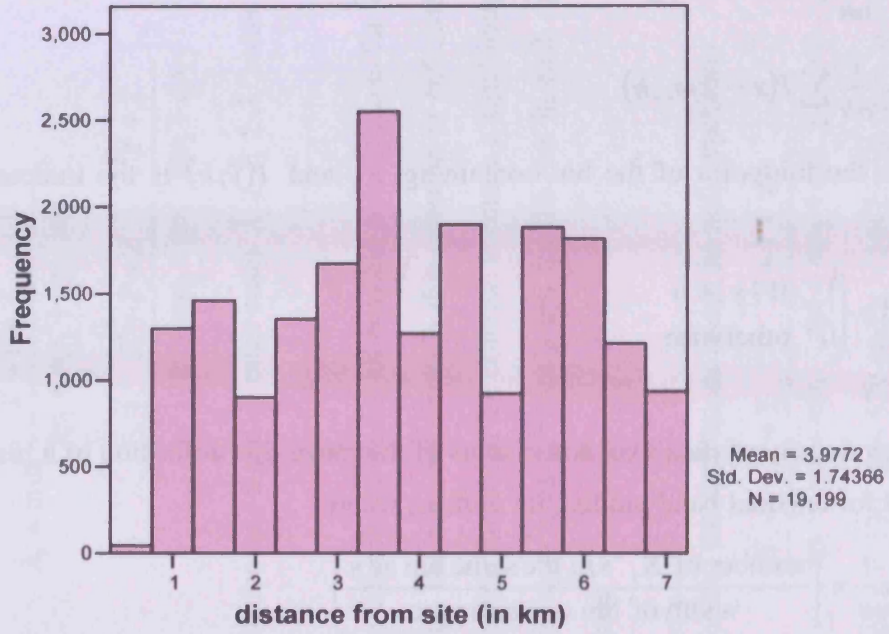
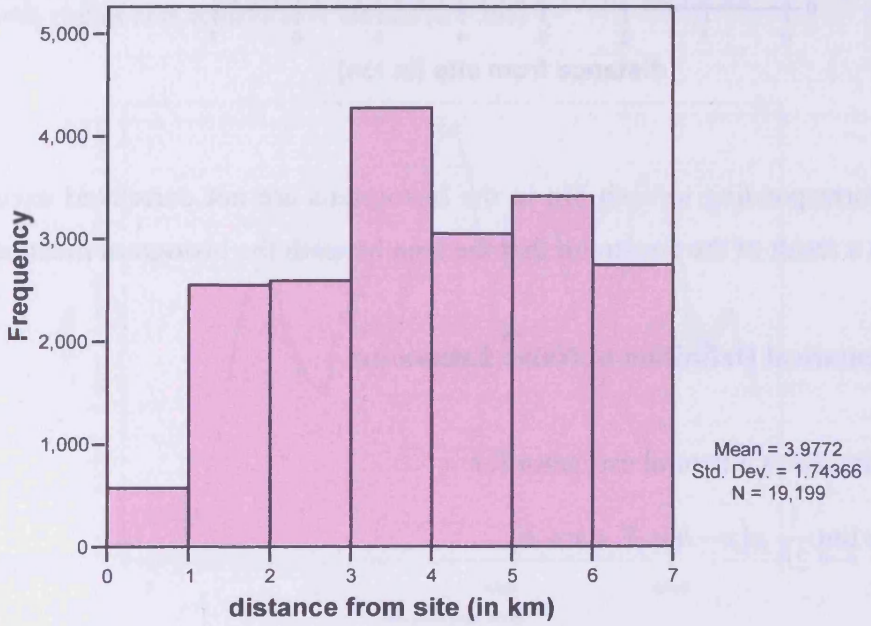


Figure A2.1 and Figure A2.2 are examples of histograms based on the dataset of all births between 1983 and 1997 which resided in the circular region with a radius of 7km centred over the Nant y Gwyddon site. Note that rather than the relative frequencies the values represented on the y-axis are the frequencies this is the only option available in SPSS. The relative frequency can be calculated by dividing the frequency by n. The variable is the distance of the residence of the births from the centre of the site in metres and the number of bins is 14 in both figures. In Figure A2.1 the origin is 0km and in Figure A2.2 the origin is 0.25km. These figures indicate the impact of the choice of origin on the shape of the histogram.

Histograms are biased as they display the average density of all the values within the same bin as X_i rather than the true density at X_i. The histogram becomes unbiased as the bandwidth tends to 0 however the smaller the bandwidth the greater the imprecision becomes as the number of observations within each band, upon which the calculation of the density is based, becomes smaller. This concept is illustrated in Figure A2.3 and Figure A2.4 which are also histograms of the distance of births from the Nant y Gwyddon landfill site with origin 0km.

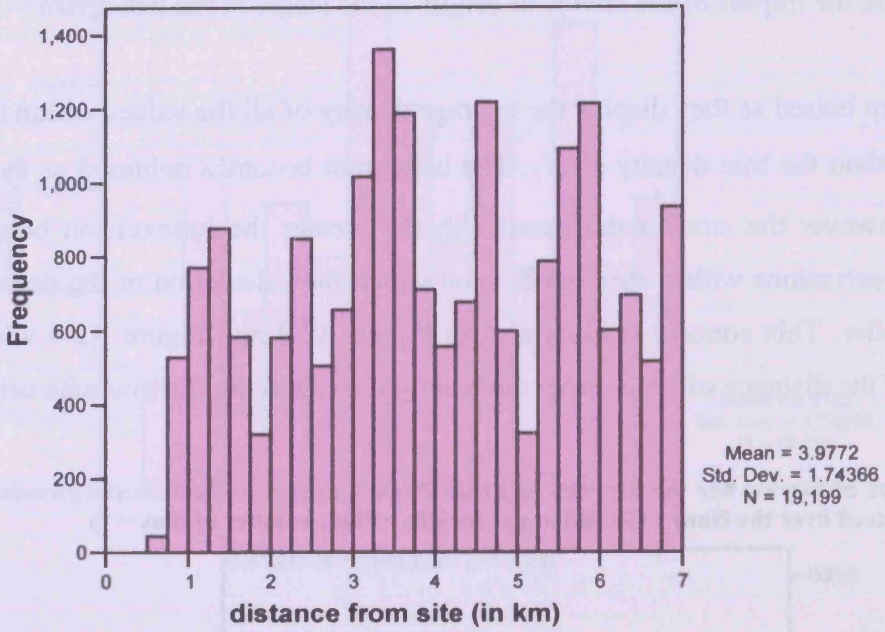
Figure A2.3: the histogram for the distance of births from Nant y Gwyddon in the circular region with radius 7km centred over the Nant y Gwyddon site (origin = 0km, number of bins = 7)



In Figure A2.3 the number of bins is halved to 7 and in Figure A2.4 it is doubled to 28. The detail in the histogram is dramatically reduced with the fall in the number of bins used. On the other hand the rise in the number of bins increases the detail shown by the histogram and the

variability of the density estimator. This makes it difficult to distinguish between genuine features of the probability density and random noise. Therefore, it is best to find an optimum bandwidth that displays the important features of the density but also hides this random noise, however, the lack of mathematical form means that histograms are not conducive to achieving this.

Figure A2.4: the histogram for the distance of births from Nant y Gwyddon in the circular region with radius 7km centred over the Nant y Gwyddon site (origin = 0km, number of bins = 28)



The heights corresponding to each bin in the histograms are not correlated except in only a weak sense as a result of the constraint that the area beneath the histogram must be equal to 1.

A2.1.4 Mathematical Definition of Naïve Estimators

The Naïve estimator is a natural estimator for

$$f(x) = \lim_{h \to 0} \frac{1}{2h} p(x - h < X < x + h).$$

Hence, for any given h, $f(x)$ can be estimated as the proportion of sample values falling within the interval (x – h, x + h). In other words,

$$\hat{f}(x) = \frac{1}{2nh} \times \{\text{number of } X_i\text{'s falling in } (x - h, x + h)\}$$

$$= \frac{1}{n}\sum_{i=1}^{n}\frac{1}{h}w\left(\frac{x - X_i}{h}\right)$$

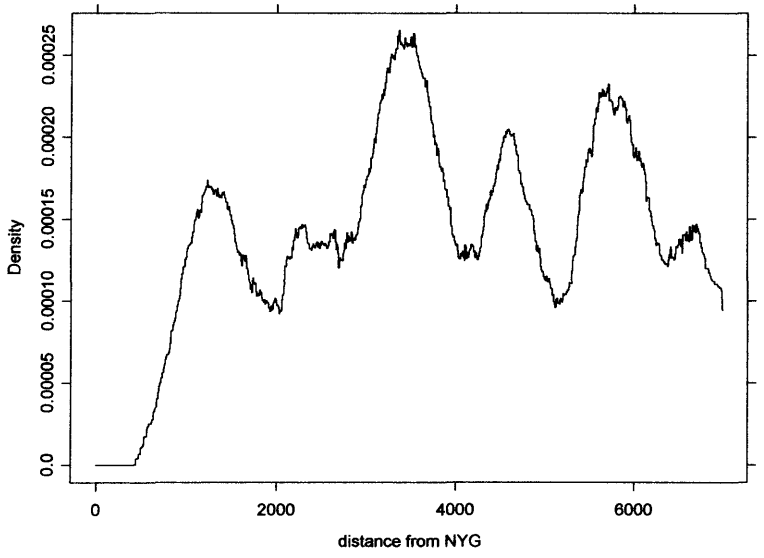
where the weight function, $w(x) = \begin{cases} \frac{1}{2} & \text{if } |x| < 1 \\ 0 & \text{otherwise} \end{cases}$.

Therefore the Naïve estimator is similar to a histogram except that the estimate of the density is evaluated at every value on the x-axis whereas in the histogram it is only determined at the centre of each bin which is then used for all other values within that bin.

A2.1.5 Properties of Naïve Estimators

Naïve estimators are less biased than histograms as a result of the evaluation of the density at every point however it is not an unbiased estimator because the use of moving bins means that some density is moved from the peaks into the tails. Unlike the histogram there is high correlation in the heights of the estimator at locations on the x-axis that are close together.

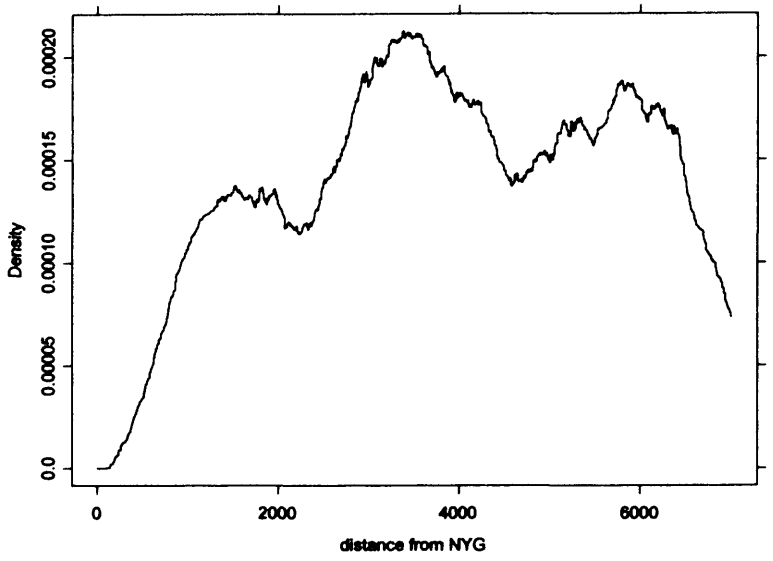
Figure A2.5: the naïve estimator for the distance of births from the centre of Nant y Gwyddon in the circular area with radius 7km centred over the site (h = 300)



The presence of bias can be seen by comparing Figure A2.5 and Figure A2.6 which are examples of naïve estimators for the distance of births from the centre of the Nant y Gwyddon

landfill site. In Figure A2.5 the value chosen for h is 300m and the underlying shape revealed is very similar to that seen in Figure A2.4.

Figure A2.6: the naïve estimator for the distance of births from the centre of Nant y Gwyddon in the circular area with radius 7km centred over the site (h = 600)



The value of h is 600m in Figure A2.6 and it is clear that some of the peaks in Figure A2.5 are not revealed in Figure A2.6, but Figure A2.5 shows far more variability in its appearance. This is because by increasing the value of h the bandwidth increases so the boxes over each observation are shorter and wider. The boxes over each observation overlap more, so the appearance of the density is smoother reducing the variation. The heights of the peaks are reduced and the heights at the tails are increase reflecting an increase in the bias. A value of h which shows the genuine features of the density but reduces the variation caused by random noise should be selected.

A2.2 Kernel Density Estimators

A2.2.1 Mathematical Definition of Kernel Density Estimators

The kernel density estimator is similar to the naïve estimator but with the weight function replaced with the symmetric kernel function k, which satisfies the equation

$$\int_{-\infty}^{\infty} k(x)dx = 1.$$

The univariate kernel density is defined as

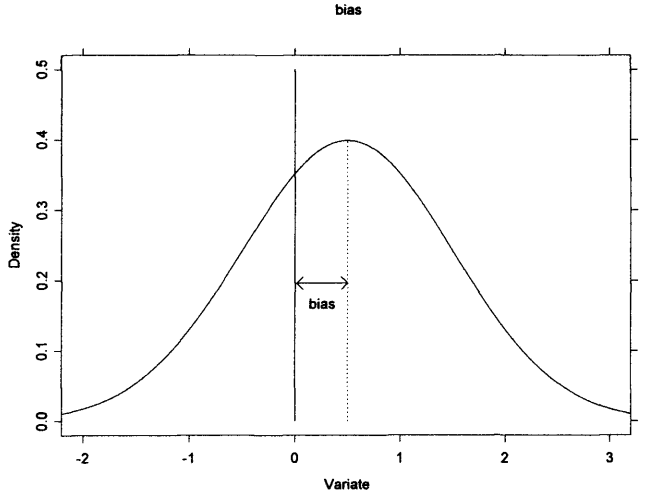
$$\hat{f}(x) = \frac{1}{nh}\sum_{i=1}^{n} k\left(\frac{x - X_i}{h}\right)$$

where h is the smoothing parameter.

A2.2.2 Definition of Bias

Suppose $\hat{\theta}$ is an estimator for the parameter θ and the expectation of $\hat{\theta}$ is $\theta + b$, where $b \neq 0$, then $\hat{\theta}$ is a biased estimator of θ where the bias is given by $b = E(\hat{\theta}) - \theta$. Figure A2.7 illustrates this point, the estimator has an expectation of 0.5 and the parameter's actual value is 0 so the bias is 0.5 and quantifies the shift to the right in this case.

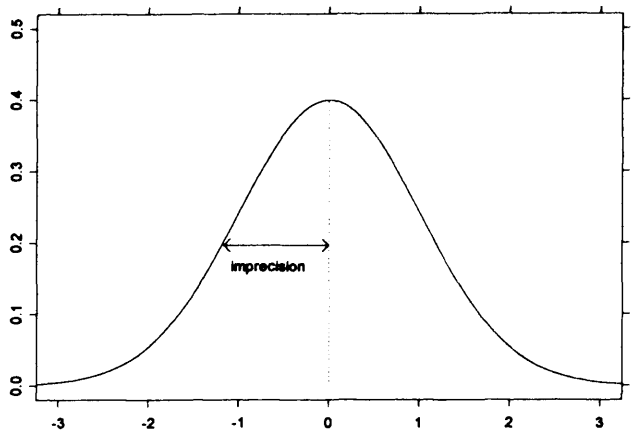
Figure A2.7: an illustration of the bias of the estimator of the parameter



A2.2.3 Definition of Imprecision

Suppose $\hat{\theta}$ is an estimator for the parameter θ and the spread of the distribution of $\hat{\theta}$ is very large about the expectation then there is a large imprecision. Therefore, there is more uncertainty about using $\hat{\theta}$ as an estimator of the true value of the parameter θ. Figure A2.8 illustrates this, the imprecision is determined by the spread of the distribution, the larger the spread the larger the imprecision of $\hat{\theta}$.

Figure A2.8: an illustration of the imprecision of the estimator for the parameter



A2.2.4 Definition of Mean Squared Error

The mean squared error (MSE) is a statistic that combines the two error types, imprecision (quantified by variance) and bias, and is defined as

$$E\left[\left\{\hat{\theta}-\theta\right\}^2\right]=E\left[\left\{\hat{\theta}-E\left(\hat{\theta}\right)\right\}^2\right]+\left\{E\left(\hat{\theta}\right)-\theta\right\}^2=\text{var}\left(\hat{\theta}\right)+bias^2$$

where $E\left(\hat{\theta}\right)$ and $\text{var}\left(\hat{\theta}\right)$ are the expectation and variance of $\hat{\theta}$, respectively.

In kernel density estimation the height of the density at each value of x is treated as the parameter for estimation. Therefore, if the chosen smoothing parameter is too large then because the peaks are under estimated and the tails are over estimated, as shown in Figure 3.5, so there is more bias in the estimation of the height of the density. In other words, there is a shift in the estimation above or below the actual height. Additionally, the estimate of the density height has a lower imprecision because the Gaussian curves at each observation overlap more so the kernel density estimate is less undulating and smoother.

In contrast, when the smoothing parameter is too low the overlap of the Gaussian curves over each observation is smaller so the kernel density is more erratic, as shown in Figure 3.6, therefore there is more imprecision. Also, there is less density shifted from the tails to the peaks so the estimate of the density height at each point has less bias because the shift away from the actual height is smaller.

The MSE of a kernel density can be calculated in a similar way. Suppose $\hat{f}(x)$ is the kernel density estimator of $f(x)$ then the MSE of $\hat{f}(x)$ when x = y is defined as

$$\begin{aligned} MSE(\hat{f}(y)) &= E\left[\{\hat{f}(y) - f(y)\}^2\right] \\ &= E\left[\{\hat{f}(y) - E(\hat{f}(y))\}^2\right] + \{E(\hat{f}(y)) - f(y)\}^2 \\ &= \text{var}(\hat{f}(y)) + \{bias\}^2 \end{aligned}$$

where $E(\hat{f}(y))$ and $\text{var}(\hat{f}(y))$ are the expectation and the variance of $\hat{f}(x)$ when x = y, respectively.

One method of estimating a good smoothing parameter is to minimise the MSE so that it is as small as possible over all values of x. Therefore, in order to do this the expectation and variance of $\hat{f}(x)$ when x = y are required, which will then enable the MSE to be calculated, this is shown below.

A2.2.5 Expectation of the Kernel Density Estimator

$$\begin{aligned} E[\hat{f}(y)] &= \int_{-\infty}^{\infty} \hat{f}(y) f(\mathbf{x}) d\mathbf{x} \\ &= \frac{1}{nh} \sum_{i=1}^{n} \int_{-\infty}^{\infty} k\left(\frac{y - x_i}{h}\right) f(x_i) dx_i \\ &= \frac{1}{nh} \sum_{n=1}^{n} \int_{-\infty}^{\infty} k\left(\frac{y - z}{h}\right) f(z) dz \quad \text{(independent of i since } X_i \text{ is a random variable)} \\ &= \frac{1}{h} \int_{-\infty}^{\infty} k\left(\frac{y - z}{h}\right) f(z) dz . \end{aligned}$$

This is a convolution of the true density f with the function k. This form enables us to conclude that $\hat{f}$ is biased. The bias is quantified in the method shown below.

Define $s = \frac{y - z}{h} \Rightarrow z = y - hs$.

$$E[\hat{f}(y)] = \frac{1}{h} \int_{\infty}^{-\infty} k(s) f(y - hs)(-hds)$$

$$= \int_{-\infty}^{\infty} k(s) f(y - hs) ds .$$

$f(y - hs)$ can be expanded about y by the Taylor's series

$$f(y - hs) \approx f(y) - hsf'(y) + \frac{h^2 s^2}{2!} f''(y) - \ldots$$

$$E\left[\hat{f}(y)\right] = f(y) \int_{-\infty}^{\infty} k(s) ds - hf'(y) \int_{-\infty}^{\infty} sk(s) ds + \frac{h^2 f''(y)}{2} \int_{-\infty}^{\infty} s^2 k(s) ds - \ldots$$

Now since k(s) is a symmetric probability density function

$$\Rightarrow \int_{-\infty}^{\infty} sk(s) ds = 0 , \int_{-\infty}^{\infty} s^3 k(s) ds = 0 , \ldots$$

in other words the odd moments are zero and $\int_{-\infty}^{\infty} k(s) ds = 1$.

Hence, $E\left[\hat{f}(y)\right] = f(y) + \frac{h^2 f''(y)}{2} \sigma_k^2 + \ldots$ where σ_k^2 is the variance of k(s), so

$\sigma_k^2 = \int_{-\infty}^{\infty} s^2 k(s) ds$.

This implies that $\hat{f}$ is biased and that the bias is $\frac{h^2}{2} f''(y) \sigma_k^2$ to second order in h, which is independent of n. The choice of h often depends on n, thus the bias becomes an implicit function of n. The expression for the bias clearly indicates that the bias increases with higher values of h therefore to minimise the bias h should be as small as possible. In addition the presence of $f''(y)$ in this expression indicates that there is negative bias in the peaks resulting in underestimation and in the troughs there is overestimation as a result of positive bias.

A2.2.6 Variance of the Kernel Density Estimator

The values chosen for X_i are from a random sample therefore there are no covariances involved in the calculation of the variance. Hence

$$\text{var}\left(\hat{f}(y)\right) = \text{var}\left\{ \frac{1}{nh} \sum_{i=1}^{n} k\left(\frac{y - X_i}{h} \right) \right\}$$

$$= \frac{1}{n^2 h^2} \sum_{i=1}^{n} \text{var}\left\{ k\left(\frac{y - X_i}{h} \right) \right\}$$

$$= \frac{n}{n^2 h^2} \text{var}\left\{ k\left(\frac{y - X}{h} \right) \right\} \quad \text{(independent of i)}$$

$$= \frac{1}{nh^2} \text{var}\left\{ k\left(\frac{y - X}{h} \right) \right\}$$

$$\text{var}\left\{ k\left(\frac{y - X}{h} \right) \right\} = \int_{-\infty}^{\infty} k^2\left(\frac{y - x}{h} \right) f(x) dx - \left[E\left\{ k\left(\frac{y - x}{h} \right) \right\} \right]^2$$

Let $s = \frac{y - x}{h} \Rightarrow x = y - sh$ and $dx = -hds$. Therefore,

$$\int_{-\infty}^{\infty} k^2\left(\frac{y - x}{h} \right) f(x) dx = \int_{\infty}^{-\infty} k^2(s) f(y - hs)(-hds)$$

$$= h \int_{-\infty}^{\infty} k^2(s) f(y - hs) ds$$

By Taylor's expansion $f(y - hs) = f(y) - hsf'(y) + \frac{h^2 s^2}{2!} f''(y) - \ldots$

Hence,

$$\int_{-\infty}^{\infty} k^2\left(\frac{y - x}{h} \right) f(x) dx = h\left\{ f(y) \int_{-\infty}^{\infty} k^2(s) ds - hf'(y) \int_{-\infty}^{\infty} sk^2(s) ds + \frac{h^2}{2!} f''(y) \int_{-\infty}^{\infty} s^2 k^2(s) ds - \ldots \right\}$$

$\int_{-\infty}^{\infty} sk^2(s) ds = 0$, $\int_{-\infty}^{\infty} s^3 k^2(s) ds = 0$,… because they are odd functions.

Combining these results gives

$$\int_{-\infty}^{\infty} k^2\left(\frac{y - x}{h} \right) f(x) dx = hf(y) \int_{-\infty}^{\infty} k^2(s) ds + \frac{h^3}{2!} f''(y) \int_{-\infty}^{\infty} s^2 k^2(s) ds + \ldots$$

$$\text{var}\left[k\left(\frac{y - X}{h} \right) \right] = hf(y) \int_{-\infty}^{\infty} k^2(s) ds + \frac{h^3}{2!} f''(y) \int_{-\infty}^{\infty} s^2 k^2(s) ds + \ldots - [E(k)]^2$$

Similarly,

$$E\left\{k\left(\frac{y-x}{h}\right)\right\} = \int_{-\infty}^{\infty} k\left(\frac{y-x}{h}\right) f(x)dx$$

$$= h\int_{-\infty}^{\infty} k(s)f(y-hs)ds$$

$$= h\left\{f(y) + \frac{h^2}{2!}f''(y)\int_{-\infty}^{\infty} s^2 k(s)ds + \ldots\right\}$$

Therefore,

$$\text{var}\left[k\left(\frac{y-X}{h}\right)\right] = hf(y)\int_{-\infty}^{\infty} k^2(s)ds + \frac{h^3}{2!}f''(y)\int_{-\infty}^{\infty} s^2 k^2(s)ds + \ldots$$

$$- h^2\left[f(y) + \frac{h^2}{2!}f''(y)\int_{-\infty}^{\infty} s^2 k(s)ds\right]^2$$

$$\approx hf(y)\int_{-\infty}^{\infty} k^2(s)ds \quad \text{for small h.}$$

Combining the above results gives

$$\text{var}\left(\hat{f}(y)\right) = \frac{f(y)}{nh}\int_{-\infty}^{\infty} k^2(s)ds \quad \text{for small h.}$$

This expression for the variance of $\hat{f}(y)$ indicates that the imprecision is reduced by increasing the value of h which conflicts with the need for a small value of h to decrease the bias.

A2.2.7 Mean Squared Error of the Kernel Density Estimator

The mean squared error of $\hat{f}(x)$ when x = y is defined as

$$MSE\left\{\hat{f}(y)\right\} = E\left(\left[\hat{f}(y) - f(y)\right]^2\right)$$

$$= E\left[\left\{\left(\hat{f}(y) - E\left(\hat{f}(y)\right)\right) - \left(f(y) - E\left(\hat{f}(y)\right)\right)\right\}^2\right]$$

$$= \text{var}\left(\hat{f}(y)\right) + bias^2\left(\hat{f}(y)\right)$$

$$= \frac{1}{nh} f(y) \int_{-\infty}^{\infty} k^2(s) ds + \left[\frac{h^2}{2!} f''(y) \sigma_k^2 \right]^2$$

Clearly the MSE of $\hat{f}(x)$ cannot be evaluated at x = y since $f(y)$ is unknown.

A2.2.8 Mean Integrated Squared Error

The mean squared error of $\hat{f}(x)$ when x = y represents the error from both bias and imprecision in the estimated height at the point x = y. However, this should be minimised over all points on the x-axis. In order to achieve this, the mean integrated squared error has been suggested.

The mean integrated squared error (MISE) is the integral of the mean squared error (MSE). The MISE measures the total area under the curve of the graph for the MSE over all values of x. The height of the MSE curve at a value of x = y is the total error at that point. Therefore, by taking the integral over all values of x the result is the overall measure of error in $\hat{f}(x)$.

The mean integrated squared error denoted by MISE is defined as

$$MISE = \int_{-\infty}^{\infty} MSE\{\hat{f}(y)\} dy$$

$$= \frac{h^4 \sigma_k^4}{4} \int_{-\infty}^{\infty} f''(y)^2 dy + \frac{1}{nh} \int_{-\infty}^{\infty} f(y) \left\{ \int_{-\infty}^{\infty} k^2(s) ds \right\} dy$$

$$= \frac{h^4 \sigma_k^4}{4} \int_{-\infty}^{\infty} f''(y)^2 dy + \frac{1}{nh} \int_{-\infty}^{\infty} k^2(s) ds \quad \text{since} \int_{-\infty}^{\infty} f(y) dy = 1$$

The definition of the MISE estimates error globally and not locally over the sub-region of interest. This weakness in using the MISE can be overcome by implementing some form of weighting function so that the regions of interest have greater influence over the MISE than those beyond that.

A2.2.9 Optimal Smoothing

In this method of determining the optimal smoothing parameter h the aim is to determine the value which minimises the MISE. Therefore, differentiating the MISE with respect to h and equating it to zero gives

$$\frac{d}{dh}MISE = h^3\sigma_k^4\int_{-\infty}^{\infty} f''(y)^2\,dy - \frac{1}{nh^2}\int_{-\infty}^{\infty} k^2(s)ds = 0$$

$$h^5 = \frac{1}{n}\frac{\int_{-\infty}^{\infty} k^2(s)ds}{\sigma_k^4\int_{-\infty}^{\infty} f''(y)^2\,dy}$$

$$h_{opt} \propto n^{-1/5} \Rightarrow \qquad \text{(A2.1)}$$

The smoothing parameter decreases with increasing sample size. In order to determine the constant of proportionality the form of both k and f must be known. This can be investigated when the Gaussian form is assumed for both functions.

A2.2.10 Normal Optimal Smoothing

The above optimal smoothing method relies on $f(y)$ being known, hence, in this case $f(y)$ is assumed to be normal in other words $f(y) \sim N(\mu, \sigma^2)$ and it is further assumed that $k(s) \sim N(0,1)$. The following results are generated from these assumptions.

$\sigma_k^2 = \int_{-\infty}^{\infty} s^2 k(s)ds = 1$ is the variance of $k(s)$

$$\Rightarrow \sigma_k^4 = 1$$

Substitute $z = \frac{x}{\sqrt{2}} \Rightarrow dz = \frac{dx}{\sqrt{2}}$ to obtain

$$\int_{-\infty}^{\infty} k^2(s)ds = \frac{1}{2\pi}\int_{-\infty}^{\infty}\exp(-z^2)dz = \frac{1}{2\pi}\int_{-\infty}^{\infty}\exp\left(\frac{-x^2}{2}\right)\frac{dx}{\sqrt{2}} = \frac{1}{2\sqrt{\pi}}$$

since $\int_{-\infty}^{\infty}\exp\left(\frac{-x^2}{2}\right)dx = \sqrt{2\pi}$ according to the definition of a standard normal.

$$f(y)=\frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2\right\}$$

Therefore,

$$f'(y)=\frac{1}{\sqrt{2\pi}\sigma}\left(-\frac{y-\mu}{\sigma^2}\right)\exp\left\{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2\right\}=-\frac{y-\mu}{\sigma^2}f(y)$$

$$f''(y)=-\frac{f(y)}{\sigma^2}+\frac{(y-\mu)^2}{\sigma^4}f(y)$$

$$\int_{-\infty}^{\infty}f''(y)^2\,dy=\frac{1}{\sigma^8}\int_{-\infty}^{\infty}\left\{(y-\mu)^2-\sigma^2\right\}^2 f^2(y)dy$$

$$f^2(y)=\frac{1}{2\pi\sigma^2}\exp\left\{-\left(\frac{y-\mu}{\sigma}\right)^2\right\}$$

$$z=\sqrt{2}\left(\frac{y-\mu}{\sigma}\right)\Rightarrow dz=\frac{\sqrt{2}dy}{\sigma}$$

$$f^2(y)dy=\frac{1}{2\pi\sigma^2}\exp\left\{-\frac{z^2}{2}\right\}\frac{\sigma}{\sqrt{2}}dz$$

$$=\frac{1}{2\sigma\sqrt{\pi}}\phi(z)dz$$

Therefore,

$$\int_{-\infty}^{\infty}f''(y)^2\,dy=\frac{1}{\sigma^8}\int_{-\infty}^{\infty}\left[\frac{\sigma^2z^2}{2}-\sigma^2\right]^2\frac{1}{\sigma2\sqrt{\pi}}\phi(z)dz$$

$$=\frac{1}{\sigma^9 2\sqrt{\pi}}\int_{-\infty}^{\infty}\left[\frac{\sigma^4z^4}{4}-\sigma^4z^2+\sigma^4\right]\phi(z)dz$$

$$=\frac{1}{\sigma^5 2\sqrt{\pi}}\int_{-\infty}^{\infty}\left[\frac{z^4}{4}-z^2+1\right]\phi(z)dz$$

By definition $\int_{-\infty}^{\infty}\phi(z)dz=1$ and

$$\int_{-\infty}^{\infty}z^2\phi(z)dz=\mathrm{var}(z)+[E(z)]^2=1$$

Hence, $\int_{-\infty}^{\infty}f''(y)^2\,dy=\frac{1}{\sigma^5 2\sqrt{\pi}}\left[\frac{1}{4}M_4-1+1\right]$ where $M_4=\int z^4\phi(z)dz=E(Z^4)$.

The moment generating function of a standard normal density is

$$M_z(t) = E[\exp(Zt)] = \exp\left(\mu t + \frac{1}{2}\sigma^2 t^2\right) = \exp\left(\frac{t^2}{2}\right) = \left[1 + \frac{t^2}{2} + \frac{1}{2} \times \frac{t^4}{4} + \ldots\right] \text{ and also}$$

$$M_z(t) = E\left[1 + Zt + \frac{Z^2 t^2}{2!} + \ldots\right] = 1 + E[Z] + \frac{t^2}{2!}E[Z^2] + \frac{t^3}{3!}E[Z^3] + \frac{t^4}{4!}E[Z^4] + \ldots$$

Hence

$$M_4 = E(Z^4) = coefficient\left(\frac{t^4}{4!}\right) = \frac{4!}{2 \times 4} = 3$$

Thus,

$$\int_{-\infty}^{\infty} f''(y)^2 dy = \frac{3}{4\sigma^5 2\sqrt{\pi}}$$

Therefore,

$$h_{opt}^5 = \frac{1}{n}\frac{1}{2\sqrt{\pi}}\frac{1}{1}\frac{4\sigma^5 2\sqrt{\pi}}{3} = \frac{1}{n}\frac{4}{3}\sigma^5$$

Hence,

$$h_{opt} = \left(\frac{4}{3n}\right)^{1/5}\sigma \text{ for a Gaussian density.}$$

The optimal smoothing parameter resulting from normal optimal smoothing usually overestimates the value of h and is only a guideline because it is calculated by assuming that the functions f and k are normally distributed, which is not the case in general. If both functions were normally distributed then using non-parametric methods to estimate the probability density function for f would not make sense and a parametric method should be used. Therefore, if the assumption is incorrect then any value for the optimal smoothing parameter, which provides a compromise between bias and variability, cannot be accurate and may only be a rough guideline.

A2.2.11 Cross Validation Method

The cross validation is an alternative method for calculating the optimal smoothing parameter h. In contrast to the normal optimal smoothing method the integrated squared error (ISE) is used rather than the MISE. The ISE is easier to deal with because the MISE involves another expectation which increases the need to know the unknown function $f(y)$. The ISE is defined as

$$ISE = \int_{-\infty}^{\infty} \left\{\hat{f}(y) - f(y)\right\}^2 dy = \int_{-\infty}^{\infty} \hat{f}(y)^2 dy - 2\int_{-\infty}^{\infty} f(y)\hat{f}(y)dy + \int_{-\infty}^{\infty} f(y)^2 dy \geq 0$$

which like the MISE is a global not local estimate of error. The results using this method differ from those of the normal optimal smoothing approach because of the use of the ISE as opposed to the MISE.

If the overall error is small then the ISE is small so the optimal value of h will give the lowest ISE value. Therefore, the optimal smoothing parameter is determined by minimising the ISE. The smoothing parameter does not appear in the function $f(y)$ so for minimisation the last term can be disregarded. Hence the remaining obstacle is to evaluate the first two terms which can be achieved using the principle explained below.

Suppose a sample $x_1, x_2, ..., x_n$ has been taken then $E[h(X)]$ can be evaluated by calculating $\sum_{i=1}^{n} \frac{h(x_i)}{n}$ therefore $\int_{-\infty}^{\infty} h(x)f(x)dx \approx \sum_{i=1}^{n} \frac{h(x_i)}{n}$ and $f(x)$ does not have to be known. The same principle for the second term can be used by estimating $\int_{-\infty}^{\infty} f(y)\hat{f}(y)dy$ using the mean density estimator over all observations in the population, i.e. $\frac{1}{n}\sum_{i=1}^{n} \hat{f}(X_i)$. However this expression requires some manipulation since there are problems in evaluating $\hat{f}(X_i)$ since one of the terms in the calculation is $\frac{1}{h\sqrt{2\pi}} \exp\left\{-\frac{1}{2h^2}(x_j - x_j)\right\} = \frac{1}{h\sqrt{2\pi}}$ regardless of the value of x_j. This difficulty is overcome by using $\frac{1}{n}\sum_{i=1}^{n} \hat{f}_{-i}(X_i)$, where $\hat{f}_{-i}(X_i)$ is the density estimator constructed without the i^{th} observation X_i but evaluated at X_i, as an estimator for

the second term instead. The exclusion of the i^{th} observation has little impact on the results and is it therefore acceptable to make the substitution. This estimator removes the need to know $f(y)$ so it is possible to evaluate it.

It is shown below that the first two terms of the ISE can be estimated by

$$\int_{-\infty}^{\infty} \hat{f}(y)^2 dy - \frac{2}{n}\sum_{i=1}^{n} \hat{f}_{-i}(X_i).$$

By definition

$$E(X) = \int_x xf(x)dx .$$

The conditional expectation

$$E_X(H(X,Y)|Y) = \int_x h(x,y)f(x|y)dx \qquad \text{(A2.2)}$$

where $E_X(H(X,Y)|Y)$ is the expectation over X of the function H(X,Y) given Y.

By definition

$$\begin{aligned}
E_{(X,Y)}(H(X,Y)) &= \int_x \int_y h(x,y)f(x,y)dydx \\
&= \int_x \int_y h(x,y)\frac{f(x,y)}{f(x)}f(x)dydx \\
&= \int_x \int_y h(x,y)f(y|x)f(x)dxdy \\
&= \int_x f(x)\left\{\int_y h(x,y)f(y|x)dy\right\}dx \\
&= \int_x f(x)E_Y(h(X,Y)|X=x)dx \qquad \text{by equation (A2.2)} \\
&= E_X(E_Y(H(X,Y)|X))
\end{aligned}$$

Now consider the expectation of $\frac{1}{n}\sum_{i=1}^{n} \hat{f}_{-i}(X_i)$, this is the same as the expectation of $\hat{f}_{-n}(X_n)$ as the expectation of $\hat{f}_{-i}(X_i)$ is the same for all i. Therefore

$$E_{\mathbf{x}}\left\{\frac{1}{n}\sum_{i=1}^{n}\hat{f}_{-i}(X_i)\right\}=E_{\mathbf{x}}\left\{\hat{f}_{-n}(X_n)\right\}$$

It follows that

$$E_{\mathbf{x}}\left\{\hat{f}_{-n}(X_n)\right\}=E_{\mathbf{x}_{-n}}\left\{E_{x_n}\left(\hat{f}_{-n}(X_n)|\mathbf{X}_{-n}\right)\right\}$$

$$=E_{\mathbf{x}_{-n}}\left\{\int_{\theta}\hat{f}_{-n}(\theta)f(\theta)d\theta\right\}$$

As shown in section A2.2.5, to second order in h $E\left(\hat{f}(y)\right)$ has a bias of $\frac{h^2}{2}f''(y)\sigma_k^2$ which is independent of n, so as n changes the bias remains constant.

Therefore,

$$E_{\mathbf{x}_{-n}}\left(\hat{f}_{-n}(x)\right)\approx E_{\mathbf{x}}\left(\hat{f}(x)\right).$$

Hence,

$$E\left\{\int_{-\infty}^{\infty}\hat{f}(y)^2dy-\frac{2}{n}\sum_{i=1}^{n}\hat{f}_{-i}(X_i)\right\}\approx E\left\{\int_{-\infty}^{\infty}\hat{f}(y)^2dy-2\int_{-\infty}^{\infty}f(y)\hat{f}(y)dy\right\}$$

It follows from the above results that

$$\int_{-\infty}^{\infty}\hat{f}(y)^2dy-2\int_{-\infty}^{\infty}\hat{f}(y)f(y)dy$$

can be estimated by

$$\int_{-\infty}^{\infty}\hat{f}(y)^2dy-\frac{2}{n}\sum_{i=1}^{n}\hat{f}_{-i}(X_i).$$

Hence the optimal smoothing parameter that minimises the ISE can be estimated by minimising

$$\int_{-\infty}^{\infty}\hat{f}(y)^2dy-\frac{2}{n}\sum_{i=1}^{n}\hat{f}_{-i}(X_i)$$

with respect to the smoothing parameter h.

The minimal point of $\int_{-\infty}^{\infty}\hat{f}(y)^2dy-\frac{2}{n}\sum_{i=1}^{n}\hat{f}_{-i}(X_i)$ is determined by plotting the estimator against values of h and selecting the minimal point on the curve otherwise a minimal search algorithm can be used. It should be noted that $\int_{-\infty}^{\infty}\hat{f}(y)^2dy-\frac{2}{n}\sum_{i=1}^{n}\hat{f}_{-i}(X_i)$ can be negative since the third term $\int_{-\infty}^{\infty}f(y)^2dy$ is ignored.

As shown below $\int_{-\infty}^{\infty}\hat{f}(y)^2dy-\frac{2}{n}\sum_{i=1}^{n}\hat{f}_{-i}(X_i)$ can be expressed as a series of Gaussian distributions therefore it can be easily derived for different values of h.

Consider the second term

$$-\frac{2}{n}\sum_{i=1}^{n}\hat{f}_{-i}(X_i)=-\frac{2}{n}\frac{1}{n-1}\sum_{i}\sum_{\substack{j\\ i\neq j}}\frac{1}{h\sqrt{2\pi}}\exp\left\{-\frac{1}{2h^2}(x_i-x_j)^2\right\}$$

The first term can be written as a series of normal distributions since $\hat{f}(y)^2$ is known.

Now

$$n^2\hat{f}(y)^2=\sum_{i=1}^{n}\frac{1}{h\sqrt{2\pi}}\exp\left\{-\frac{1}{2h^2}(y-x_i)^2\right\}\sum_{j=1}^{n}\frac{1}{h\sqrt{2\pi}}\exp\left\{-\frac{1}{2h^2}(y-x_j)^2\right\}$$

$$=\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{1}{2\pi h^2}\exp\left\{-\frac{1}{2h^2}\left[(y-x_i)^2+(y-x_j)^2\right]\right\}$$

$$n^2\int_{-\infty}^{\infty}\hat{f}(y)^2dy=\frac{1}{2\pi h^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\int_{-\infty}^{\infty}\exp\left\{-\frac{1}{2h^2}\left[(y-x_i)^2+(y-x_j)^2\right]\right\}dy$$

$$=\frac{1}{2\pi h^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\int_{-\infty}^{\infty}\exp\left\{-\frac{1}{2h^2}\left[2y^2-2y(x_i+x_j)+(x_i^2+x_j^2)\right]\right\}dy$$

$$=\frac{1}{2\pi h^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\int_{-\infty}^{\infty}\exp\left\{-\frac{1}{h^2}\left[y^2-\frac{2y(x_i+x_j)}{2}+\frac{(x_i^2+x_j^2)}{2}\right]\right\}dy$$

$$= \frac{1}{2\pi h^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \exp\left\{ -\frac{1}{4h^2} (x_i - x_j)^2 \right\} \int_{-\infty}^{\infty} \exp\left\{ -\frac{1}{h^2} \left[y - \frac{(x_i + x_j)}{2} \right]^2 \right\} dy$$

Now $\int_{-\infty}^{\infty} \exp\left\{ -\frac{1}{h^2} \left[y - \frac{x_i + x_j}{2} \right]^2 \right\} dy = \frac{h\sqrt{2\pi}}{\sqrt{2}}$

Therefore

$$n^2 \int_{-\infty}^{\infty} \hat{f}(y) dy = \frac{1}{2\pi h^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \exp\left\{ -\frac{1}{4h^2} (x_i - x_j)^2 \right\} h\sqrt{\pi}$$

$$= \frac{1}{2h\sqrt{\pi}} \sum_{i=1}^{n} \sum_{j=1}^{n} \exp\left\{ -\frac{1}{4h^2} (x_i - x_j)^2 \right\}$$

Hence

$$\int_{-\infty}^{\infty} \hat{f}(y)^2 dy = \frac{1}{n^2} \frac{1}{2h\sqrt{\pi}} \sum_{i=1}^{n} \sum_{j=1}^{n} \exp\left\{ -\frac{1}{4h^2} (x_i - x_j)^2 \right\}$$

Hence, it follows that minimising

$$\frac{1}{n^2 2h\sqrt{\pi}} \sum_{i=1}^{n} \sum_{j=1}^{n} \exp\left\{ -\frac{1}{4h^2} (x_i - x_j)^2 \right\} - \frac{2}{n(n-1)} \sum_{i} \sum_{j \neq i} \frac{1}{h\sqrt{2\pi}} \exp\left\{ -\frac{1}{2h^2} (x_i - x_j)^2 \right\}$$

with respect to h is equivalent to minimising

$$\int_{-\infty}^{\infty} \hat{f}(y)^2 dy - \frac{2}{n} \sum_{i=1}^{n} \hat{f}_{-i}(X_i).$$

The smoothing parameter determined by this optimisation procedure is also only a guide since the global minimum achieved by minimising the ISE may produce some local features that are undesirable.

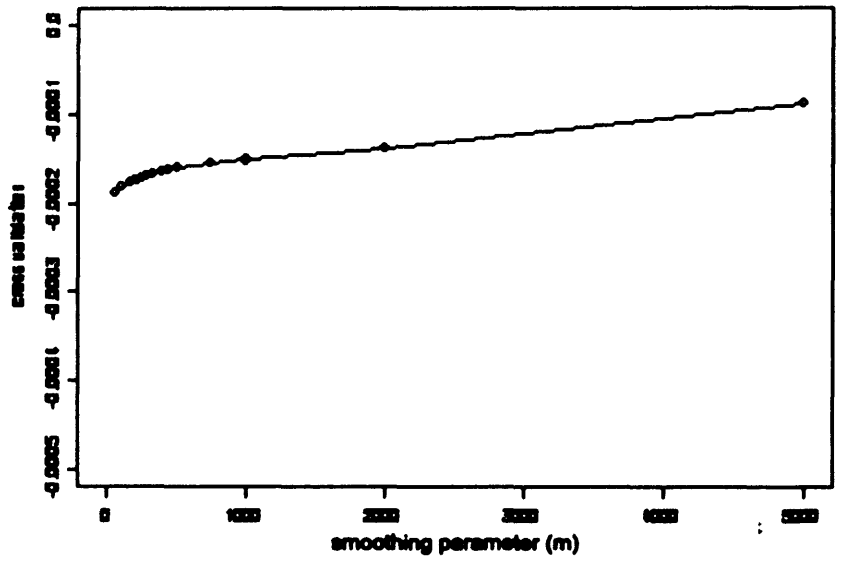
Typically, the cross validation method can be performed on data using Bowman's programs written for use in SPlus. However, the selected population for Nant Y Gwyddon has over 19000 individuals in it and subsequently Bowman's program for cross-validation cannot cope with the size of this dataset. In response to this a new program had to be written using the algebra above.

A2.2.12 Example of Selecting a Smoothing Parameter

The two values generated by the normal optimal smoothing and cross validation can be used as guidelines and then trial and error can be used to find a good smoothing parameter by looking at the kernel densities that are produced using different smoothing parameters around the optimal value calculated and selecting the smallest possible smoothing parameter that does not show too much detail. This is a subjective method but it is a good way of selecting a smoothing parameter that shows sufficient structural detail.

Using normal optimal smoothing for the distance data for Nant y Gwyddon displayed in Figure A2.1 to Figure A2.6, the value of h that minimises the MISE is 257 meters.

Figure A2.9: the smoothing parameter plotted against $\int_{-\infty}^{\infty} \hat{f}(y)^2 dy - \frac{2}{n}\sum_{i=1}^{n} \hat{f}_{-i}(X_i)$



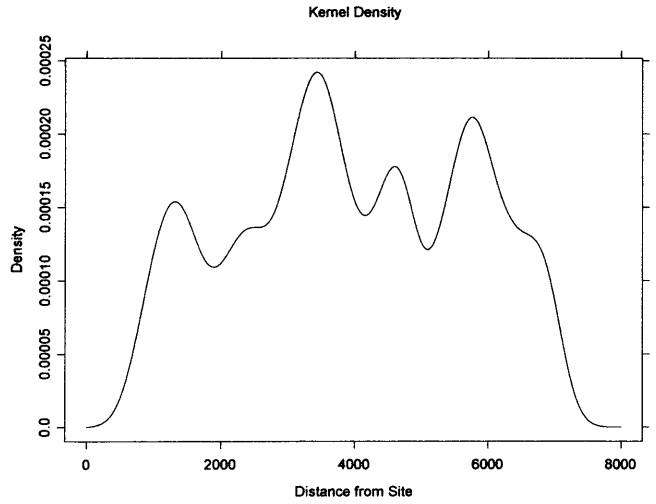
The cross validation method involves plotting the estimator for the ISE against different smoothing parameters, h. Figure A2.9 is the plot between selected values of h and the equivalent value of $\int_{-\infty}^{\infty} \hat{f}(y)^2 dy - \frac{2}{n}\sum_{i=1}^{n} \hat{f}_{-i}(X_i)$. There is no minimum in this curve therefore cross validation cannot be used to determine a specific recommended value for the smoothing parameter. However the curve in Figure A2.9 is very flat indicating that there are only small

differences in the values of $\int_{-\infty}^{\infty} \hat{f}(y)^2 dy - \frac{2}{n}\sum_{i=1}^{n} \hat{f}_{-i}(X_i)$ for different values of h. Therefore, from Figure A2.9 any of the values of h below 500m could be suitable.

The failure of using cross validation on the data considered here means that the value for the smoothing parameter resulting from normal optimal smoothing was taken as a starting point in the trial and error method suggested. Different values were subsequently taken to visually determine whether other values below 500m would be more suitable.

Figure A2.10 is the kernel density for the distance of births from the centre of the Nant y Gwyddon site between 1983 and 1997 where the smoothing parameter is 257m. The smoothing parameter here is too low because under smoothing has allowed too much variability to appear in the profile. The under smoothing has produced too much structural detail, some of which will not be genuine features of the distribution, and so create a complex profile unsuitable for comparison with other density estimators. Therefore higher smoothing parameters should be considered.

Figure A2.10: the kernel density estimator for the distance of births from the centre of Nant y Gwyddon in the circular area with radius 7km centred over the site (h = 257m)



If the value of the smoothing parameter is increased slightly to 500m, as shown in Figure A2.11, the detail is reduced. The smoothing parameter is too high, so over smoothing shifts density from the peaks into the tails. The features of the kernel density profile have disappeared from the plot because the density from the peaks has been displaced into the tails.

The result is over simplified and provides little information on the form of the actual distribution.

Figure A2.11: the kernel density estimator for the distance of births from the centre of Nant y Gwyddon in the circular area with radius 7km centred over the site (h = 500m)

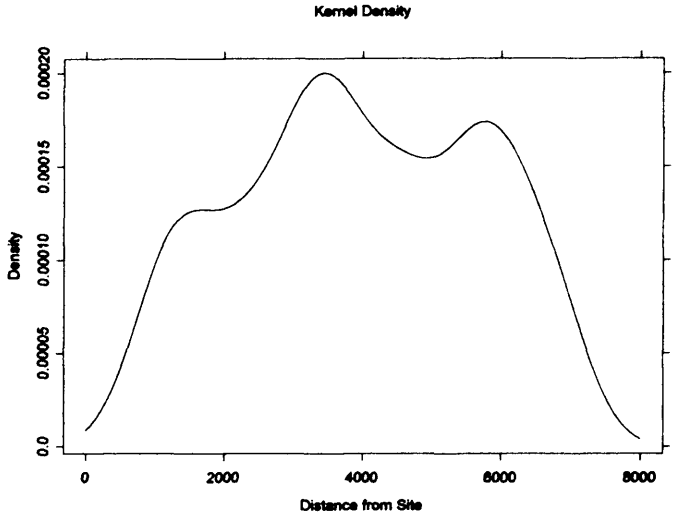
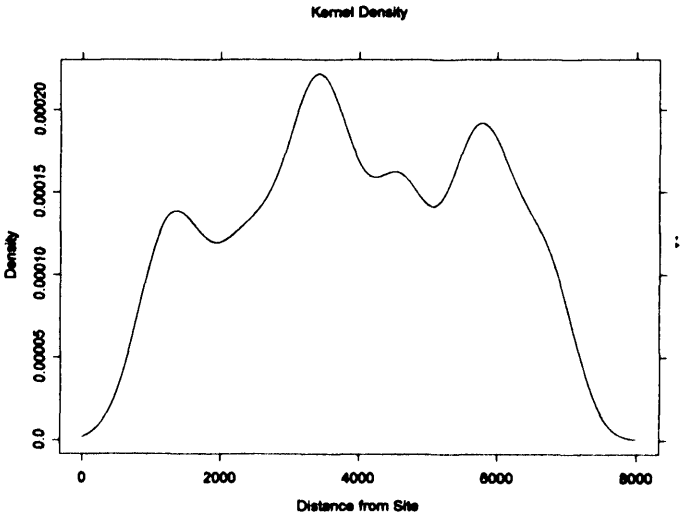


Figure A2.12 is an example where the smoothing parameter is between 257m and 500m. This is an improvement on Figure A2.11 where over smoothing had removed most of the features of the underlying distribution by shifting the density from the peaks into the tails. Also, it does not show the same amount of variability as Figure A2.10 which was overly complex.

Figure A2.12: the kernel density estimator for the distance of births from the centre of Nant y Gwyddon in the circular area with radius 7km centred over the site (h = 350m)



A2.3 Two-Dimensional Kernel Function

If there are two variables then a two dimensional kernel density can be constructed producing a surface. It is sufficient, in this case, to assume independence between the two variables because the form of the kernel density does not affect the form of the final estimate if there is enough data. Hence, the two-dimensional kernel function has the form

$$\hat{f}(x_1, x_2) = \frac{1}{nh_1h_2}\sum_{i=1}^{n} k\left(\frac{x_1 - X_{i1}}{h_1}\right)k\left(\frac{x_2 - X_{i2}}{h_2}\right).$$

This allows each variable to have its own smoothing parameter, h_1 and h_2, which is usually required if the two variables have different scales.

This can be simplified if the smoothing parameters are equal for the two variables so the kernel density has the form

$$\hat{f}(\mathbf{x}) = \frac{1}{nh^2}\sum_{i=1}^{n} k\left(\frac{x_1 - X_{i1}}{h}\right)k\left(\frac{x_2 - X_{i2}}{h}\right)$$

where h is the smoothing parameter.

A2.4 Significance Contours

From results presented in Appendix 2 the expectation and variance of the density estimator $\hat{f}(x)$ are defined as follows

$$E\left(\hat{f}(y)\right) = f(y) \quad \text{and} \quad \text{var}\left(\hat{f}(y)\right) = \frac{f(y)}{nh}\int_{-\infty}^{\infty} k^2(s)ds$$

for small h where $k(s)$ is the kernel function used. If $Y = h(X)$ is a transformation of X, then by Taylor's Theorem $Y = h(\mu) + (X - \mu)h'(\mu) + \ldots$ where $\mu = E[X]$. Hence, the variance of Y is approximately given by $\text{var}[Y] = h'(\mu)^2\,\text{var}[X]$.

It follows from this result that the variance of the square root of the probability density for the observed congenital malformations is

$$\text{var}\left[\sqrt{\hat{f}(x)}\right] = \left\{h'\left(E\left[\hat{f}(x)\right]\right)\right\}^2 \text{var}\left[\hat{f}(x)\right]$$

$$\approx \frac{1}{4E\left(\hat{f}(x)\right)}\left\{\frac{f(x)}{nh}\int_{-\infty}^{\infty}k^2(s)ds\right\} \text{ for small h}$$

$$\approx \frac{1}{4f(x)}\left\{\frac{f(x)}{nh}\int_{-\infty}^{\infty}k^2(s)ds\right\} \text{ for small h}$$

$$= \frac{1}{4nh}\int_{-\infty}^{\infty}k^2(s)ds$$

This is an interesting and important result since it does not depend on the form of the unknown density $f(x)$.

However, the variance of the square root of the probability density for the expected birth defects, denoted $\hat{g}(x)$, differs from the expressions quoted in the literature as the weights, w_i, which are used to adjust for confounders are involved in the algebra. The resultant variance is

$$\text{var}\left[\sqrt{\hat{g}(x)}\right] = \frac{1}{4n^2h\overline{w}^2}\int_{-\infty}^{\infty}k^2(s)ds\sum_{i=1}^{n}w_i$$

The correlation between the two densities $\hat{f}(x)$ and $\hat{g}(x)$ is negligible since $\hat{f}(x)$ has such small numbers of observations involved in its construction relative to $\hat{g}(x)$. Hence, it can be assumed that $\hat{f}(x)$ and $\hat{g}(x)$ are independent. Consequently, the variance of the difference between the square roots of two densities is

$$\text{var}\left\{\sqrt{\hat{f}(x)} - \sqrt{\hat{g}(x)}\right\} = \text{var}\left\{\sqrt{\hat{f}(x)}\right\} + \text{var}\left\{\sqrt{\hat{g}(x)}\right\}$$

$$= \frac{1}{4n_fh}\int_{-\infty}^{\infty}k^2(s)ds + \frac{1}{4n_g^2h\overline{w}^2}\int_{-\infty}^{\infty}k^2(s)ds\sum_{i=1}^{n}w_i^2$$

$$= \frac{n_g^2\overline{w}^2 + n_f\sum_{i=1}^{n}w_i^2}{4hn_fn_g^2\overline{w}^2}\int_{-\infty}^{\infty}k^2(s)ds$$

assuming that the same kernel function and smoothing parameter is used in both density estimators. This assumption is made and commented upon earlier.

In conclusion, the spatial significance between the estimated densities $\hat{f}$ and $\hat{g}$ at a point x is assessed by the function below

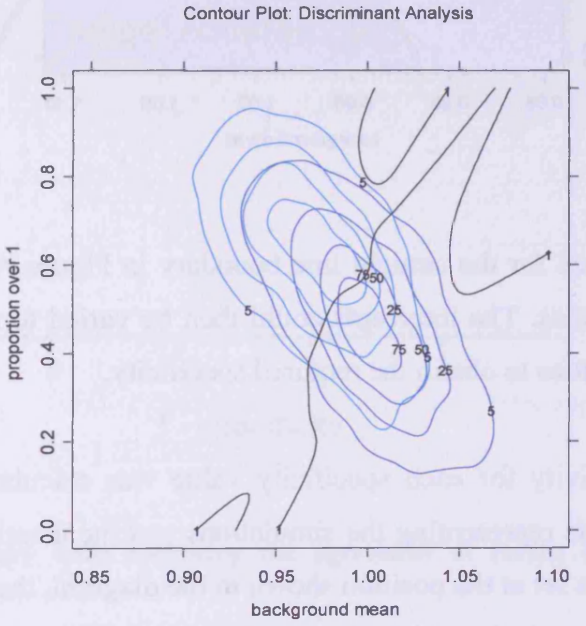
$$significance = \frac{\sqrt{\hat{f}(x)} - \sqrt{\hat{g}(x)}}{\sqrt{\frac{n_g^2 \overline{w}^2 + n_f \sum_{i=1}^{n} w_i^2}{4hn_f n_g^2 \overline{w}^2} \int_{-\infty}^{\infty} k^2(s) ds}}.$$

Appendix Three – Discriminant Analysis

A3.1 Linear Non-Parametric Discriminant Boundary

Figure A3.1 is the discriminant boundary for the background mean and the proportion for the case where the alternative hypothesis assumes a risk attributable to exposure that is weaker than it was initially ($\beta = 0.0125$). The resultant discriminant boundary not only has the primary line bisecting the region into the critical and acceptance region but there are three additional, smaller boundaries. These secondary contours are illogical resulting in misclassifications where even though the corresponding statistics are closer to that of one hypothesis the opposite hypothesis is selected using the boundary.

Figure A3.1: the contour plots for the kernel densities of the joint distributions of the background mean and the proportion along with the discriminant boundary with a weaker exposure risk

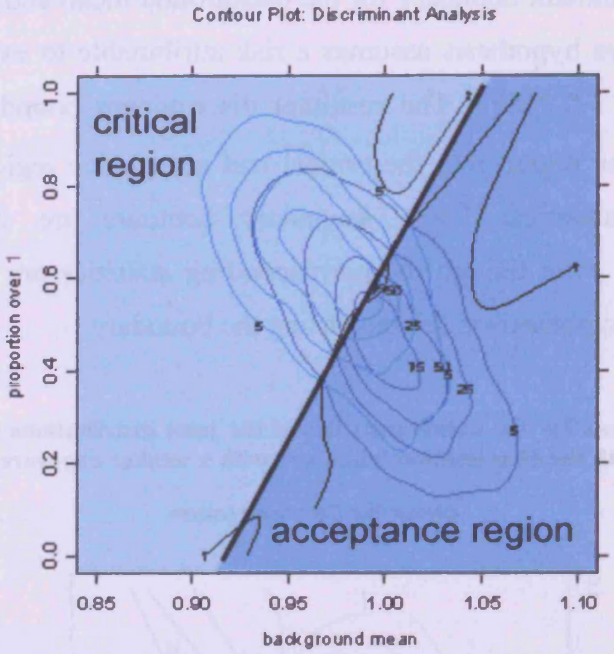


These unnecessary contours occur because of the instability in the regions of lower density. The impact of this erratic behaviour in the tails on the efficiency of the test can be checked by implementing an alternative non-parametric method using a linear discriminant boundary.

The boundary chosen is shown in Figure A3.2. A straight line was drawn along the general direction of the original discriminant boundary in the regions of higher density in Figure A3.1.

The straight line is easy to implement and bisects the region so that the statistics generated for a specific region can be used to make a clear decision as to whether or not there is an effect.

Figure A3.2: the new straight line boundary



The gradient was calculated for the straight line boundary in Figure A3.2. The gradient was kept at this value at all times. The intercepts could then be varied to move the discriminant boundaries along the densities to obtain the required specificity.

The corresponding sensitivity for each specificity value was calculated using the shortest distance between the points representing the simulations and the discriminant boundary. The discriminant boundary was set at the position shown in the diagram, the baseline position, and the shortest distances were calculated for all simulations under the null hypothesis and under the alternative hypothesis with this weaker exposure risk.

The shortest distance between the point (x_0, y_0) and the line $Ax + By + C = 0$ is given by

$$d = \frac{Ax_0 + By_0 + C}{\sqrt{A^2 + B^2}}.$$

If the specificity is 0.95 then the boundary is located at a distance D from the baseline position, where D is the 95th percentile of the 1000 distances calculated under the null. Therefore, the proportion of the distances for weak exposure that were above D is equivalent to the proportion of the weak exposure simulations that led to correctly rejecting the null hypothesis. Hence the proportion of the distances for weak exposure simulations above D is the sensitivity. The ROC curve, shown in Figure A3.3, is generated by varying the specificity and repeating this process.

Figure A3.3: the ROC curve comparing the two non-parametric discriminant boundaries

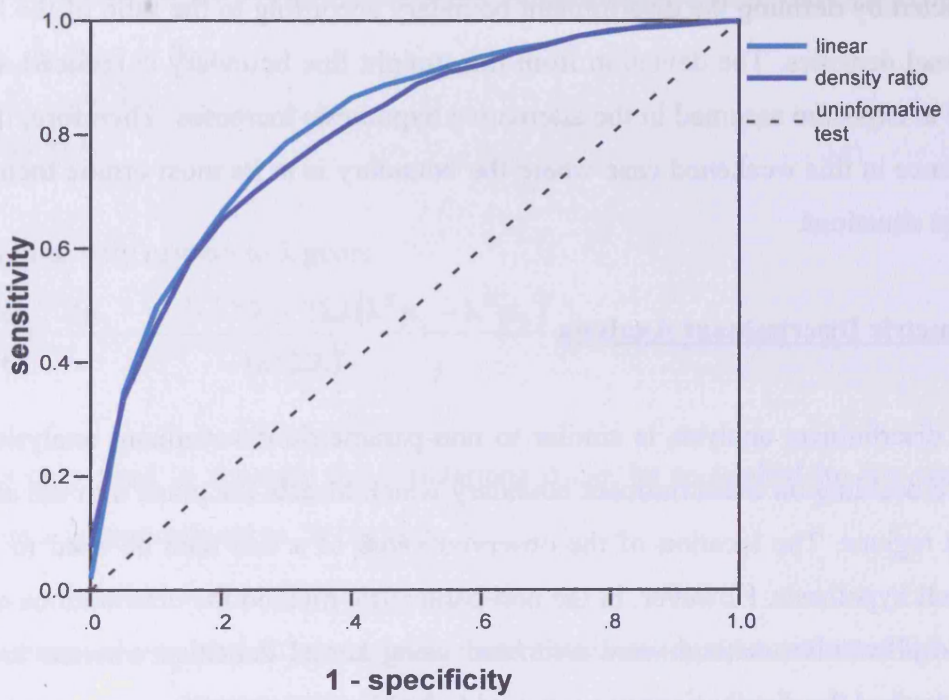


Table A3.1: a contingency table indicating the agreement of results for the two non-parametric discriminant boundaries

			density ratio		total
			acceptance region	critical region	
linear	acceptance region	count	642	11	653
		% of total	64.2	1.1	65.3
	critical region	count	4	343	347
		% of total	0.4	34.3	34.7
total		count	646	354	1000
		% of total	64.6	35.4	100

The two curves are very similar, particularly where the specificity is high, indicating that the instability of the discriminant boundary in regions of low density has little impact on the performance of the hypothesis test. This is because the erratic patterns in the discriminant boundaries only affect regions with low density so few simulations are affected. This is supported by Table A3.1 where the simulations under the alternative generally led to the same conclusions using both discriminant boundaries when the boundaries are positioned so that the specificity is 0.95.

It is therefore clear that when using non-parametric discriminant analysis the efficiency is not greatly affected by defining the discriminant boundary according to the ratio of the heights of the two kernel densities. The deviation from the straight line boundary is reduced as the risk attributable to exposure assumed in the alternative hypothesis increases. Therefore, if it makes little difference in this weakened case where the boundary is at its most erratic then it can be used in most situations.

A3.2 Parametric Discriminant Analysis

Parametric discriminant analysis is similar to non-parametric discriminant analysis because both involve deciding on a discriminant boundary which bisects the plane into the acceptance and critical regions. The location of the observed value of $\mathbf{x}$ can then be used to accept or reject the null hypothesis. However, in the non-parametric method the distributions of $\mathbf{x}$ under the null and alternative method were estimated using kernel densities whereas in the non-parametric method the distributions are assumed to be bivariate normal.

A3.2.1 Fisher's Linear Discriminant Analysis

Fisher's linear combination of the observations is denoted by $Y = \boldsymbol{\lambda}^T \mathbf{x}$. The coefficients of this linear combination are chosen such that the ratio of the difference of the means of the linear combinations in the two regions to its variance is maximised. This gives the maximum separation between the two hypotheses therefore making it less probable that the null hypothesis is incorrectly accepted or rejected.

The acceptance region is labelled as Π_1 and the critical region is denoted by Π_2. The mean of Y in Π_1 is $\bar{Y}_1 = \boldsymbol{\lambda}^T \boldsymbol{\mu}_1$ and in Π_2 is $\bar{Y}_2 = \boldsymbol{\lambda}^T \boldsymbol{\mu}_2$, where $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ are the means of $\mathbf{x}$ under the null hypothesis and under the alternative hypothesis, respectively.

In this method the covariance matrices are assumed to be equal in other words $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \boldsymbol{\Sigma}$. Hence, the variance of Y is $\boldsymbol{\lambda}^T \boldsymbol{\Sigma} \boldsymbol{\lambda}$ for both populations.

The aim is to choose $\boldsymbol{\lambda}$ to maximise the ratio given by the difference in the means for the two linear combinations for the populations divided by the variance. In other words, maximise

$$\phi = \frac{\boldsymbol{\lambda}^T \boldsymbol{\mu}_1 - \boldsymbol{\lambda}^T \boldsymbol{\mu}_2}{\boldsymbol{\lambda}^T \boldsymbol{\Sigma} \boldsymbol{\lambda}}$$

with respect to $\boldsymbol{\lambda}$.

Differentiating with respect to $\boldsymbol{\lambda}$ gives

$$\frac{\partial \phi}{\partial \boldsymbol{\lambda}} = \frac{2(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)\boldsymbol{\lambda}^T \boldsymbol{\Sigma} \boldsymbol{\lambda} - 2\boldsymbol{\Sigma} \boldsymbol{\lambda}(\boldsymbol{\lambda}^T \boldsymbol{\mu}_1 - \boldsymbol{\lambda}^T \boldsymbol{\mu}_2)}{(\boldsymbol{\lambda}^T \boldsymbol{\Sigma} \boldsymbol{\lambda})^2} = \mathbf{0}.$$

Since $\boldsymbol{\lambda}$ is only used to separate the populations it can be multiplied by any constant so it follows that $\boldsymbol{\lambda}$ is proportional to $\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)$.

The parameters $\boldsymbol{\mu}_1$, $\boldsymbol{\mu}_2$ and $\boldsymbol{\Sigma}$ are unknown and therefore must be estimated by $\bar{\mathbf{x}}_1$, $\bar{\mathbf{x}}_2$ and $\mathbf{S}$, respectively. Here, $\mathbf{S}$ is the pooled sample covariance matrix defined by

$$\mathbf{S} = \frac{(n_1 - 1)\mathbf{S}_1 + (n_2 - 1)\mathbf{S}_2}{n_1 + n_2 - 2}.$$

The assignment procedure is to assign an individual to Π_1 if $Y = (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)^T \mathbf{S}^{-1} \mathbf{x}$ is closer to $\bar{Y}_1 = (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)^T \mathbf{S}^{-1} \bar{\mathbf{x}}_1$ than to $\bar{Y}_2 = (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)^T \mathbf{S}^{-1} \bar{\mathbf{x}}_2$, otherwise assign to Π_2.

Y is closer to $\bar{Y}_1$ if $|Y - \bar{Y}_1| < |Y - \bar{Y}_2|$ or equivalently $|Y - \bar{Y}_1| - |Y - \bar{Y}_2| < 0$. The value of 0 on the right hand side of the latter equality is arbitrary and is equivalent to setting the discriminant boundary to where the height ratio is equal to 1 in the non-parametric

discriminant analysis. Therefore, changing this value moves the position of the discriminant boundary.

If the value on the right hand side of the inequality is set to the $(100 - \alpha)^{th}$ percentile of the 1000 values of $|Y - \bar{Y}_1| - |Y - \bar{Y}_2|$ for each of the simulations under the null hypothesis then the significance level is α% and the power is $(100 - \alpha)$%. The resultant power is estimated by the percentage of simulations under the alternative for which the value of $|Y - \bar{Y}_1| - |Y - \bar{Y}_2|$ is greater than the $(100 - \alpha)^{th}$ percentile.

The parametric method makes two assumptions about the two distributions of $\mathbf{x}$ under the null and under the alternative. The first is that the two distributions are bivariate normal if $\mathbf{x}$ has two dimensions. The second is that the two covariance matrices under both hypotheses are equal. Before the parametric discriminant analysis is performed these two assumptions must be tested.

A3.2.2 Testing the Assumptions

Test for bivariate normality

The test for joint normality of data uses the squared generalised distances denoted by

$$d_j^2 = (\mathbf{x}_j - \bar{\mathbf{x}})^T \mathbf{S}^{-1} (\mathbf{x}_j - \bar{\mathbf{x}}) \quad \text{for j} = 1, 2, \ldots, \text{n}$$

where $\mathbf{x}_j$ is the j^{th} sample observation.

If the parent population is bivariate normal with size greater than 25 or 30 then each of the squared distances d_1^2, $d_2^2, \ldots,$ d_n^2 should behave like a chi-squared random variable. Therefore, in a bivariate normal distribution the following things are true:

1. roughly half of the d_j^2 are less than or equal to $\chi_2^2\left(\frac{1}{2}\right)$; and
2. a plot of the ordered d_j^2 versus $\chi_2^2\left(\frac{j - 1/2}{n}\right)$ is a straight line.

Consequently, from the first point the set of bivariate normal outcomes for **x** should have a probability of $(\mathbf{x}-\bar{\mathbf{x}})^T \mathbf{S}^{-1}(\mathbf{x}-\bar{\mathbf{x}}) \leq \chi_2^2\left(\frac{1}{2}\right)$ equal to 0.5. Hence, 50% of the observations should lie in the elliptical region defined by $(\mathbf{x}-\bar{\mathbf{x}})^T \mathbf{S}^{-1}(\mathbf{x}-\bar{\mathbf{x}}) \leq \chi_2^2\left(\frac{1}{2}\right)$. The 50th percentile of a chi-squared distribution with 2 degrees of freedom is 1.386. Table A3.2 gives the percentage of outcomes under each hypothesis that lie in this region. All the percentages shown are very close to 50% with a maximum deviation of 1.7% from this figure.

Table A3.2: the percentage of outcomes for x that lie in the elliptical region

Hypothesis	% above 1.386
null	51.0
alternative	51.7

The second criterion was that the plot between the squared distances and the corresponding chi-squared percentiles forms a relatively straight line. This type of plot is called a chi-squared or gamma plot. In order to construct a chi-squared plot, the squared distances are ordered such that $d_{(1)}^2 \leq d_{(2)}^2 \leq \ldots \leq d_{(n)}^2$. Then the pairs $\left(d_{(j)}^2, \chi_2^2\left(\frac{j-1/2}{n}\right)\right)$ are plotted, where $\chi_2^2\left(\frac{j-1/2}{n}\right)$ is the $\frac{100(j-1/2)^{th}}{n}$ percentile of the chi-squared distribution with 2 degrees of freedom.

This plot should resemble a straight line. A systematic curve pattern suggests a lack of normality and one or more points too far to the right of the line indicate some outliers.

The chi-squared plots for background mean and proportion under both the null and alternative are given in Figure A3.4 and Figure A3.5, respectively. Both plots seem to be roughly straight however they do deviate slightly with higher values for the squared generalised distances.

The two criteria considered seem to indicate that the joint distributions for mean and proportion are bivariate Gaussian under both hypotheses. However, a bivariate Gaussian distribution should have elliptical contours of equal density height. The contours for the two distributions were generated for the non-parametric discriminant analysis shown in Figure

3.22b. These are not precisely elliptical and therefore the assumption of normality is questionable.

Figure A3.4: the chi-squared plot for the joint distribution of the mean and proportion under the null hypothesis

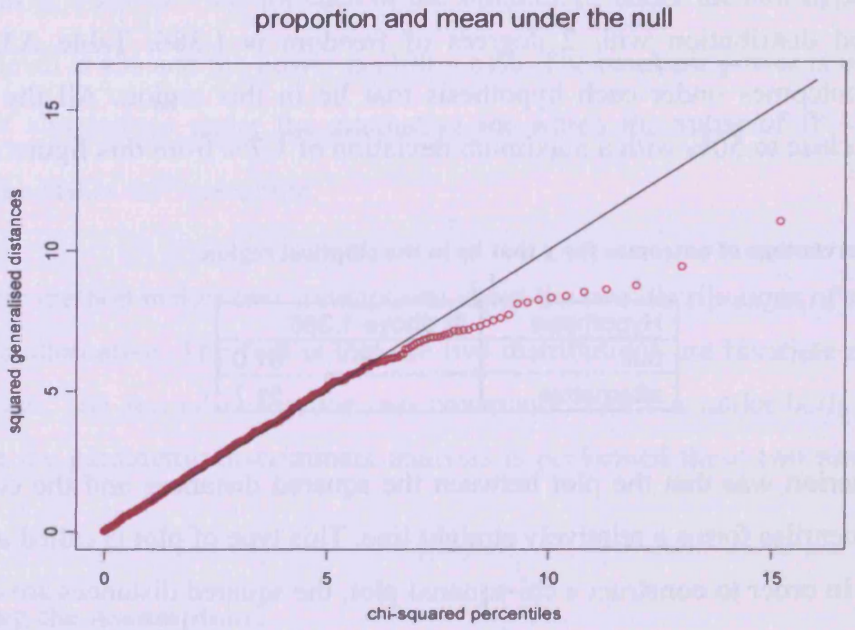
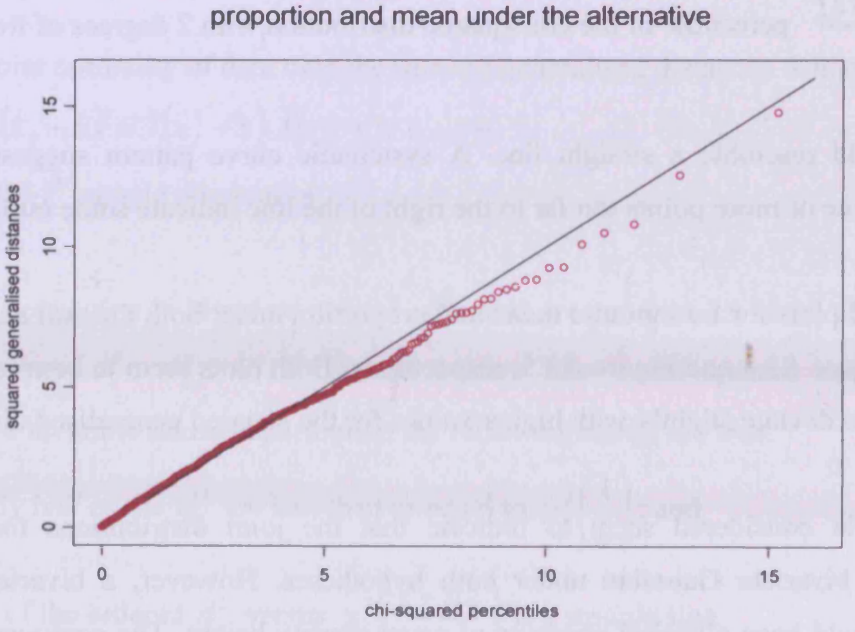


Figure A3.5: the chi-squared plot for the joint distribution of the mean and proportion under the alternative hypothesis



Testing for equality of covariance matrices

The equality of the covariance matrices under the null and alternative hypotheses can be tested using the Box's M test. Under the null hypothesis of equality, Box's M follows an F distribution. The test was carried out in SPSS for the covariance matrices for the mean and proportion. The p-value is displayed in Table A3.3 and is very low suggesting that there is significant evidence that the covariance matrices are different and thus the assumption has not been met.

Table A3.3: the p-values for the Box's M test

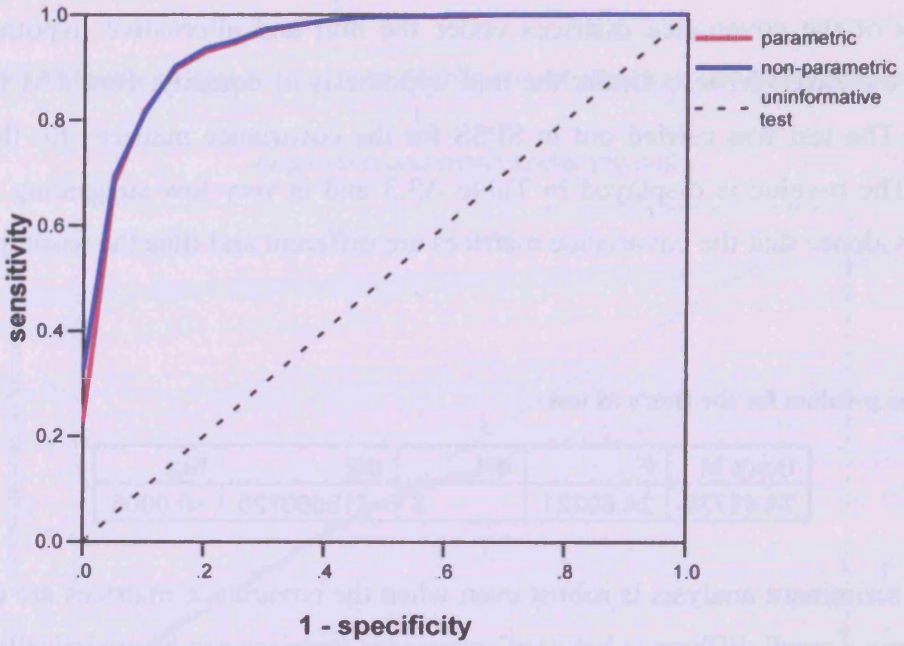
Box's M	F	df1	df2	Sig.
74.48739	24.80221	3	718560720	<0.0005

However, discriminant analysis is robust even when the covariance matrices are different and when n is large a small difference between covariance matrices can be statistically significant. If the covariance matrices are different then the discriminant boundary generated in the non-parametric discriminant analysis would be quadratic whereas in the regions where there is data these boundaries are relatively straight indicating that the covariance matrices may be equal under the null and the alternative hypotheses. Hence, it may be reasonable to make the assumption of equality of covariance matrices.

A3.2.3 Efficiency of the Parametric Approach

The parametric discriminant analysis was performed on the data, regardless of the dubious validity of the assumptions made and the ROC curve assessing its efficiency is displayed in Figure A3.6. The curve for the parametric approach is almost identical in shape to that for the non-parametric approach. However, when the specificity is above 0.95 the non-parametric method performs slightly better and is not built on possibly invalid assumptions like the parametric discriminant analysis. Therefore, in conclusion the non-parametric method in this situation is the favoured approach.

Figure A3.6: the ROC curve comparing the parametric and non-parametric approaches to discriminant analysis



Appendix Four: Congenital Malformations

Table A4.1: the logistic regression model results for all congenital malformations

Factor and level	Beta	S.E.	Wald	df	Sig.	Exp(Beta)
bthyear = 1983			737.98	14	0.000	
bthyear = 1984	0.049	0.061	0.66	1	0.416	1.051
bthyear = 1985	-0.088	0.062	2.02	1	0.156	0.915
bthyear = 1986	-0.125	0.062	4.04	1	0.044	0.883
bthyear = 1987	-0.191	0.062	9.34	1	0.002	0.826
bthyear = 1988	-0.121	0.061	3.91	1	0.048	0.886
bthyear = 1989	-0.182	0.062	8.60	1	0.003	0.834
bthyear = 1990	-0.516	0.067	59.04	1	0.000	0.597
bthyear = 1991	-0.545	0.068	64.09	1	0.000	0.580
bthyear = 1992	-0.911	0.076	142.03	1	0.000	0.402
bthyear = 1993	-0.832	0.075	122.55	1	0.000	0.435
bthyear = 1994	-0.945	0.079	143.74	1	0.000	0.389
bthyear = 1995	-0.847	0.077	121.12	1	0.000	0.429
bthyear = 1996	-1.278	0.089	205.27	1	0.000	0.279
bthyear = 1997	-0.572	0.071	65.63	1	0.000	0.565
sex = male	-0.296	0.026	127.22	1	0.000	0.743
matgrp = 12 to 20			12.61	3	0.006	
matgrp = 21 to 30	-0.016	0.041	0.14	1	0.708	0.985
matgrp = 31 to 40	0.045	0.047	0.89	1	0.344	1.046
matgrp = > 40	0.344	0.120	8.22	1	0.004	1.411
townquin = most affluent			10.36	4	0.035	
townquin = next affluent	0.023	0.042	0.31	1	0.580	1.023
townquin = median	0.095	0.042	5.18	1	0.023	1.099
townquin = next deprived	0.100	0.042	5.69	1	0.017	1.105
townquin = most deprived	0.103	0.043	5.82	1	0.016	1.109
hospital = Shrewsbury			481.93	19	0.000	
hospital = Glen Clwyd	-0.505	0.117	18.58	1	0.000	0.603
hospital = Wrexham	-0.097	0.114	0.72	1	0.395	0.908
hospital = West Cheshire	-0.061	0.137	0.20	1	0.657	0.941
hospital = County Hosp Here	0.114	0.184	0.38	1	0.536	1.121
hospital = Bronglais	-0.014	0.136	0.01	1	0.919	0.986
hospital = West Wales Gen	-0.185	0.124	2.23	1	0.135	0.831
hospital = Withybush	0.022	0.118	0.03	1	0.853	1.022
hospital = Nevill Hall	-0.987	0.131	56.38	1	0.000	0.373
hospital = Royal Gwent	-0.776	0.115	45.80	1	0.000	0.460
hospital = Ysbyty Gwynedd	-0.281	0.117	5.75	1	0.016	0.755
hospital = Princess of Wales	-0.026	0.115	0.05	1	0.823	0.975
hospital = Caerphilly Miners	-0.856	0.137	38.92	1	0.000	0.425
hospital = East Glamorgan	-0.033	0.114	0.09	1	0.770	0.967
hospital = Prince Charles	-0.795	0.127	39.06	1	0.000	0.452
hospital = Mid Wales	0.002	0.148	0.00	1	0.987	1.002
hospital = Uni Hosp Wales	-0.039	0.112	0.12	1	0.726	0.961
hospital = Llandough	-0.341	0.114	8.91	1	0.003	0.711
hospital = Singleton	-0.303	0.112	7.39	1	0.007	0.739
hospital = Neath	-0.510	0.124	16.86	1	0.000	0.601
Constant	-3.681	0.122	913.69	1	0.000	0.025

Table A4.2: the logistic regression model results for chromosomal anomalies

Factor and level	Beta	S.E.	Wald	df	Sig.	Exp(Beta)
bthyear = 1983			9.34	14	0.808	
bthyear = 1984	-0.031	0.63	0.00	1	0.961	0.970
bthyear = 1985	0.519	0.56	0.86	1	0.352	1.681
bthyear = 1986	0.480	0.56	0.74	1	0.390	1.615
bthyear = 1987	0.523	0.55	0.91	1	0.340	1.687
bthyear = 1988	0.484	0.55	0.78	1	0.377	1.622
bthyear = 1989	0.686	0.53	1.66	1	0.198	1.987
bthyear = 1990	0.473	0.55	0.74	1	0.388	1.605
bthyear = 1991	0.324	0.56	0.34	1	0.561	1.383
bthyear = 1992	0.090	0.59	0.02	1	0.878	1.094
bthyear = 1993	0.811	0.52	2.42	1	0.120	2.250
bthyear = 1994	0.657	0.53	1.52	1	0.218	1.929
bthyear = 1995	0.129	0.59	0.05	1	0.826	1.138
bthyear = 1996	0.091	0.59	0.02	1	0.876	1.096
bthyear = 1997	-0.049	0.61	0.01	1	0.935	0.952
sex = male	-0.198	0.18	1.27	1	0.259	0.820
matgrp = 12 to 20			103.47	3	0.000	
matgrp = 21 to 30	0.115	0.36	0.10	1	0.751	1.122
matgrp = 31 to 40	0.903	0.37	5.90	1	0.015	2.468
matgrp = > 40	3.029	0.43	49.69	1	0.000	20.682
townquin = most affluent			2.06	4	0.725	
townquin = next affluent	-0.168	0.26	0.42	1	0.518	0.846
townquin = median	-0.380	0.28	1.88	1	0.171	0.684
townquin = next deprived	-0.183	0.26	0.49	1	0.484	0.833
townquin = most deprived	-0.241	0.28	0.75	1	0.385	0.786
hospital = Shrewsbury			41.37	16	0.000	
hospital = Glen Clwyd	14.517	6613	0.00	1	0.998	2016246
hospital = Wrexham	14.636	6613	0.00	1	0.998	2271650
hospital = West Cheshire	14.565	6613	0.00	1	0.998	2115718
hospital = West Wales Gen	13.520	6613	0.00	1	0.998	743828
hospital = Withybush	14.251	6613	0.00	1	0.998	1545601
hospital = Nevill Hall	12.020	6613	0.00	1	0.999	166067
hospital = Royal Gwent	12.233	6613	0.00	1	0.999	205368
hospital = Ysbyty Gwynedd	13.659	6613	0.00	1	0.998	854893
hospital = Princess of Wales	14.295	6613	0.00	1	0.998	1615055
hospital = Caerphilly Miners	13.181	6613	0.00	1	0.998	530237
hospital = East Glamorgan	14.218	6613	0.00	1	0.998	1496051
hospital = Prince Charles	12.926	6613	0.00	1	0.998	410923
hospital = Uni Hosp Wales	14.116	6613	0.00	1	0.998	1350492
hospital = Llandough	13.708	6613	0.00	1	0.998	897744
hospital = Singleton	13.485	6613	0.00	1	0.998	718694
hospital = Neath	12.625	6613	0.00	1	0.998	304099
Constant	-21.825	6613	0.00	1	0.997	0.000

Table A4.3: the forward stepwise logistic regression model results for chromosomal anomalies

	Variable	Model Log Likelihood	Change in -2 Log Likelihood	df	Sig. of Change
Step 1	matgrp	-1107	72.5	3	<0.0005
Step 2	matgrp	-1075	68.3	3	<0.0005
	hospital	-1071	59.5	16	<0.0005

Factor and level	Beta	S.E.	Wald	df	Sig.	Exp(Beta)
matgrp = 12 to 20			106.361	3	0.000	
matgrp = 21 to 30	0.16	0.36	0.189	1	0.664	1.17
matgrp = 31 to 40	0.96	0.36	6.979	1	0.008	2.61
matgrp = > 40	3.07	0.42	52.519	1	0.000	21.52
hospital = Shrewsbury			44.525	16	0.000	
hospital = Glen Clwyd	14.9	6613	0.000	1	0.998	2926397
hospital = Wrexham	15.0	6613	0.000	1	0.998	3200691
hospital = West Cheshire	14.9	6613	0.000	1	0.998	3019320
hospital = West Wales Gen	13.9	6613	0.000	1	0.998	1048670
hospital = Withybush	14.6	6613	0.000	1	0.998	2154223
hospital = Nevill Hall	12.3	6613	0.000	1	0.999	222806
hospital = Royal Gwent	12.6	6613	0.000	1	0.998	284881
hospital = Ysbyty Gwynedd	13.9	6613	0.000	1	0.998	1143230
hospital = Princess of Wales	14.7	6613	0.000	1	0.998	2330957
hospital = Caerphilly Miners	13.5	6613	0.000	1	0.998	726508
hospital = East Glamorgan	14.5	6613	0.000	1	0.998	2044502
hospital = Prince Charles	13.2	6613	0.000	1	0.998	544601
hospital = Uni Hosp Wales	14.5	6613	0.000	1	0.998	1956996
hospital = Llandough	14.1	6613	0.000	1	0.998	1266273
hospital = Singleton	13.8	6613	0.000	1	0.998	980729
hospital = Neath	13.0	6613	0.000	1	0.998	424265
Constant	-22.1	6613	0.000	1	0.997	0.000

Table A4.4: the logistic regression model results for non-chromosomal anomalies

Factor and level	Beta	S.E.	Wald	df	Sig.	Exp(Beta)
bthyear = 1983			273.84	14	0.000	
bthyear = 1984	0.079	0.098	0.65	1	0.420	1.082
bthyear = 1985	-0.209	0.104	4.08	1	0.043	0.811
bthyear = 1986	-0.176	0.101	3.02	1	0.082	0.838
bthyear = 1987	-0.304	0.104	8.61	1	0.003	0.738
bthyear = 1988	-0.129	0.099	1.72	1	0.190	0.879
bthyear = 1989	-0.299	0.103	8.41	1	0.004	0.741
bthyear = 1990	-0.511	0.108	22.16	1	0.000	0.600
bthyear = 1991	-0.605	0.112	29.01	1	0.000	0.546
bthyear = 1992	-0.997	0.128	60.86	1	0.000	0.369
bthyear = 1993	-0.924	0.126	53.83	1	0.000	0.397
bthyear = 1994	-0.989	0.130	57.41	1	0.000	0.372
bthyear = 1995	-0.809	0.123	43.01	1	0.000	0.445
bthyear = 1996	-1.305	0.147	79.11	1	0.000	0.271
bthyear = 1997	-0.426	0.110	15.06	1	0.000	0.653
sex = male	-0.255	0.043	35.00	1	0.000	0.775
matgrp = 12 to 20			4.53	3	0.209	
matgrp = 21 to 30	-0.135	0.064	4.44	1	0.035	0.874
matgrp = 31 to 40	-0.128	0.075	2.93	1	0.087	0.879
matgrp = > 40	-0.148	0.244	0.37	1	0.544	0.862
townquin = most affluent			7.15	4	0.128	
townquin = next affluent	0.125	0.072	3.03	1	0.082	1.133
townquin = median	0.166	0.070	5.63	1	0.018	1.180
townquin = next deprived	0.156	0.069	5.11	1	0.024	1.169
townquin = most deprived	0.107	0.069	2.35	1	0.125	1.112
hospital = Shrewsbury			192.52	16	0.000	
hospital = Glen Clwyd	-1.324	1.028	1.66	1	0.198	0.266
hospital = Wrexham	-0.713	1.021	0.49	1	0.485	0.490
hospital = West Cheshire	-0.793	1.025	0.60	1	0.439	0.453
hospital = West Wales Gen	-0.926	1.029	0.81	1	0.368	0.396
hospital = Withybush	-0.584	1.028	0.32	1	0.570	0.558
hospital = Nevill Hall	-1.595	1.026	2.42	1	0.120	0.203
hospital = Royal Gwent	-1.442	1.022	1.99	1	0.158	0.236
hospital = Ysbyty Gwynedd	-0.770	1.025	0.56	1	0.452	0.463
hospital = Princess of Wales	-0.706	1.022	0.48	1	0.489	0.493
hospital = Caerphilly Miners	-1.581	1.031	2.35	1	0.125	0.206
hospital = East Glamorgan	-0.732	1.020	0.52	1	0.473	0.481
hospital = Prince Charles	-1.368	1.023	1.79	1	0.181	0.255
hospital = Uni Hosp Wales	-0.680	1.020	0.44	1	0.505	0.507
hospital = Llandough	-0.852	1.021	0.70	1	0.404	0.427
hospital = Singleton	-1.141	1.021	1.25	1	0.264	0.320
hospital = Neath	-1.201	1.024	1.38	1	0.241	0.301
Constant	-3.005	1.024	8.61	1	0.003	0.050

Table A4.5: the forward stepwise logistic regression model results for non-chromosomal anomalies

	Variable	Model Log Likelihood	Change in -2 Log Likelihood	df	Sig. of Change
Step 1	bthyear	-12375	299.7	14	<0.0005
Step 2	bthyear	-12276	302.5	14	<0.0005
	hospital	-12225	201.6	16	<0.0005
Step 3	sex	-12125	35.6	1	<0.0005
	bthyear	-12258	302.2	14	<0.0005
	hospital	-12208	201.4	16	<0.0005

Factor and level	Beta	S.E.	Wald	df	Sig.	Exp(Beta)
bthyear = 1983			276.5	14	0.000	
bthyear = 1984	0.080	0.098	0.7	1	0.414	1.083
bthyear = 1985	-0.208	0.104	4.0	1	0.044	0.812
bthyear = 1986	-0.176	0.101	3.0	1	0.083	0.839
bthyear = 1987	-0.303	0.104	8.6	1	0.003	0.738
bthyear = 1988	-0.127	0.099	1.7	1	0.196	0.880
bthyear = 1989	-0.300	0.103	8.5	1	0.004	0.741
bthyear = 1990	-0.511	0.108	22.2	1	0.000	0.600
bthyear = 1991	-0.606	0.112	29.1	1	0.000	0.545
bthyear = 1992	-0.998	0.128	61.0	1	0.000	0.369
bthyear = 1993	-0.926	0.126	54.1	1	0.000	0.396
bthyear = 1994	-0.991	0.130	57.8	1	0.000	0.371
bthyear = 1995	-0.811	0.123	43.3	1	0.000	0.444
bthyear = 1996	-1.304	0.147	79.1	1	0.000	0.271
bthyear = 1997	-0.426	0.110	15.2	1	0.000	0.653
sex = male	-0.256	0.043	35.2	1	0.000	0.774
hospital = Shrewsbury			187.6	16	0.000	
hospital = Glen Clwyd	-1.374	1.027	1.8	1	0.181	0.253
hospital = Wrexham	-0.758	1.020	0.6	1	0.458	0.469
hospital = West Cheshire	-0.842	1.024	0.7	1	0.411	0.431
hospital = West Wales Gen	-0.966	1.028	0.9	1	0.347	0.381
hospital = Withybush	-0.616	1.027	0.4	1	0.549	0.540
hospital = Nevill Hall	-1.609	1.025	2.5	1	0.116	0.200
hospital = Royal Gwent	-1.477	1.021	2.1	1	0.148	0.228
hospital = Ysbyty Gwynedd	-0.788	1.024	0.6	1	0.442	0.455
hospital = Princess of Wales	-0.758	1.020	0.6	1	0.458	0.469
hospital = Caerphilly Miners	-1.612	1.030	2.4	1	0.118	0.199
hospital = East Glamorgan	-0.756	1.020	0.6	1	0.458	0.469
hospital = Prince Charles	-1.378	1.022	1.8	1	0.178	0.252
hospital = Uni Hosp Wales	-0.735	1.019	0.5	1	0.471	0.480
hospital = Llandough	-0.896	1.020	0.8	1	0.380	0.408
hospital = Singleton	-1.168	1.020	1.3	1	0.252	0.311
hospital = Neath	-1.236	1.023	1.5	1	0.227	0.291
Constant	-2.973	1.020	8.5	1	0.004	0.051

Figure A4.1: the proportion of births in Wales with a chromosomal anomaly by year of birth 1983-1997

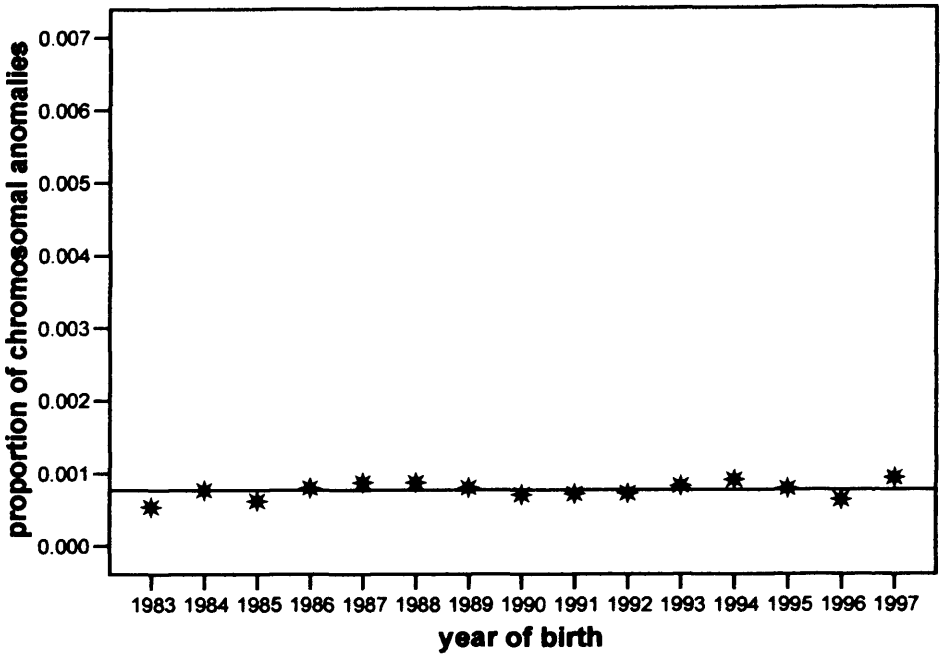


Table A4.6: the proportion of births in Wales with a chromosomal anomaly with 95% confidence intervals by year of birth 1983-1997

Year of Birth	Prop. of Births with a Chromosomal Anomaly (95% CI)	Total Number of Births
1983	0.000538 (0.000282, 0.000794)	31598
1984	0.000766 (0.000466, 0.001066)	32634
1985	0.000618 (0.000354, 0.000882)	33979
1986	0.000797 (0.000502, 0.001092)	35125
1987	0.000872 (0.000570, 0.001174)	36707
1988	0.000871 (0.000574, 0.001168)	37901
1989	0.000800 (0.000514, 0.001086)	37509
1990	0.000702 (0.000437, 0.000967)	38466
1991	0.000715 (0.000445, 0.000985)	37761
1992	0.000723 (0.000451, 0.000996)	37331
1993	0.000825 (0.000530, 0.001120)	36380
1994	0.000908 (0.000594, 0.001223)	35241
1995	0.000787 (0.000490, 0.001083)	34319
1996	0.000632 (0.000368, 0.000897)	34784
1997	0.000934 (0.000610, 0.001257)	34274

Figure A4.2: the proportion of births with a non-chromosomal anomaly by year of birth in Wales 1983-1997

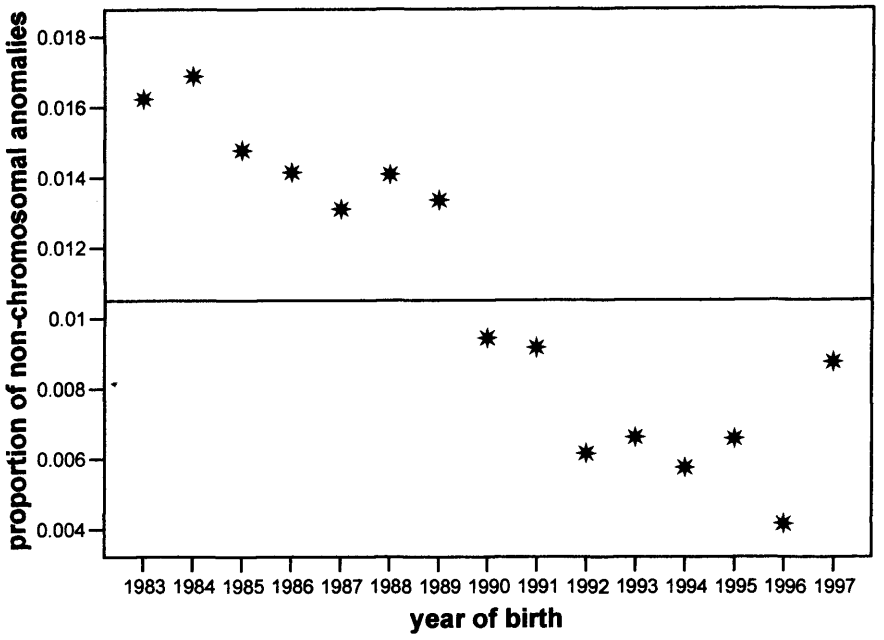


Table A4.7: the proportion of births in Wales with a non-chromosomal anomaly with 95% confidence intervals by year of birth 1983-1997

Year of Birth	Prop. of Births with a Non-Chromosomal Anomaly (95% CI)	Total Number of Births
1983	0.0162 (0.0148, 0.0176)	31598
1984	0.0169 (0.0155, 0.0183)	32634
1985	0.0148 (0.0135, 0.0161)	33979
1986	0.0141 (0.0129, 0.0154)	35125
1987	0.0131 (0.0119, 0.0143)	36707
1988	0.0141 (0.0129, 0.0153)	37901
1989	0.0134 (0.0122, 0.0145)	37509
1990	0.0094 (0.0085, 0.0104)	38466
1991	0.0092 (0.0082, 0.0101)	37761
1992	0.0062 (0.0054, 0.0070)	37331
1993	0.0066 (0.0058, 0.0075)	36380
1994	0.0058 (0.0050, 0.0066)	35241
1995	0.0066 (0.0057, 0.0074)	34319
1996	0.0042 (0.0035, 0.0048)	34784
1997	0.0088 (0.0078, 0.0097)	34274

Figure A4.3: the proportion of births in Wales with a chromosomal anomaly by maternal age 1983-1997

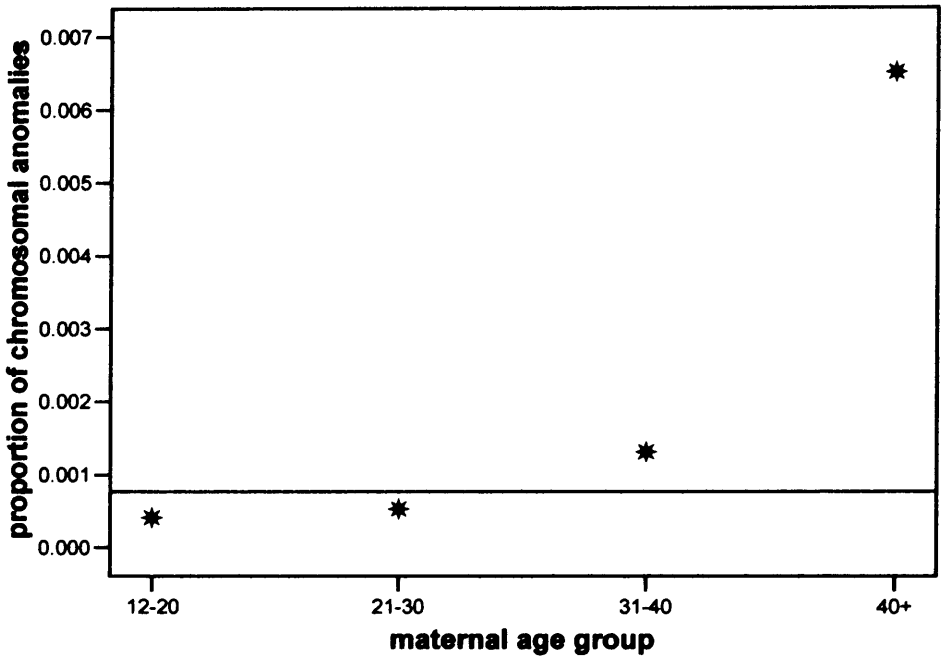


Table A4.8: the proportion of births in Wales with a chromosomal anomaly with 95% confidence intervals by maternal age 1983-1997

Maternal Age Group	Prop. of Births with a Chromosomal Anomaly (95% CI)	Total Number of Births
12-20	0.000412 (0.000254, 0.000571)	63078
21-30	0.000530 (0.000452, 0.000608)	333811
31-40	0.001312 (0.001116, 0.001507)	131897
>40	0.006510 (0.004329, 0.008691)	5223

Figure A4.4: the proportion of births in Wales with a non-chromosomal anomaly by maternal age 1983-1997

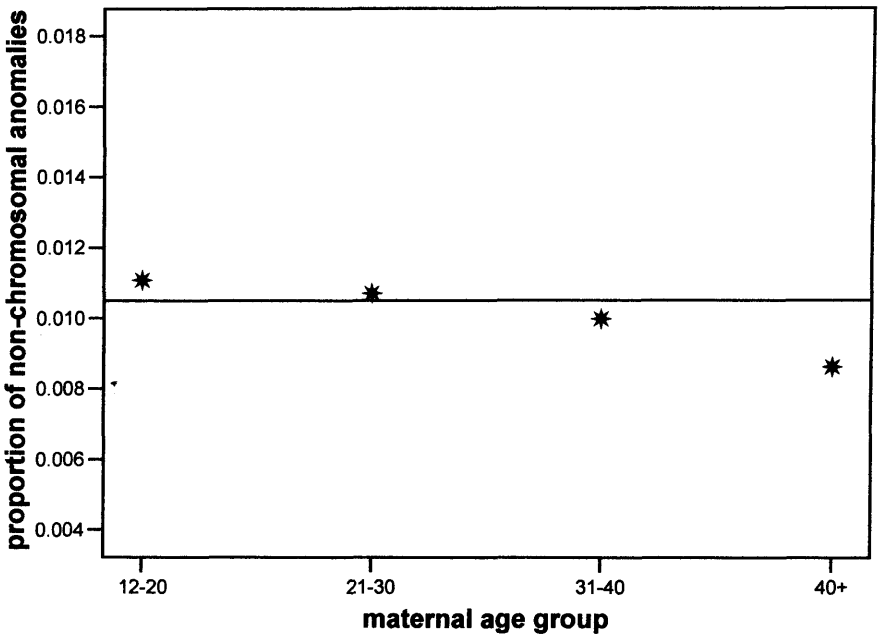


Table A4.9: the proportion of births in Wales with a non-chromosomal anomaly with 95% confidence intervals by maternal age 1983-1997

Maternal Age Group	Prop. of Births with a Non-Chromosomal Anomaly (95% CI)	Total Number of Births
12-20	0.0111 (0.0103, 0.0119)	63078
21-30	0.0107 (0.0104, 0.0111)	333811
31-40	0.0100 (0.0094, 0.0105)	131897
>40	0.0086 (0.0061, 0.0111)	5223

Figure A4.5: the proportion of births in Wales with a chromosomal anomaly by Townsend quintile 1983-1997

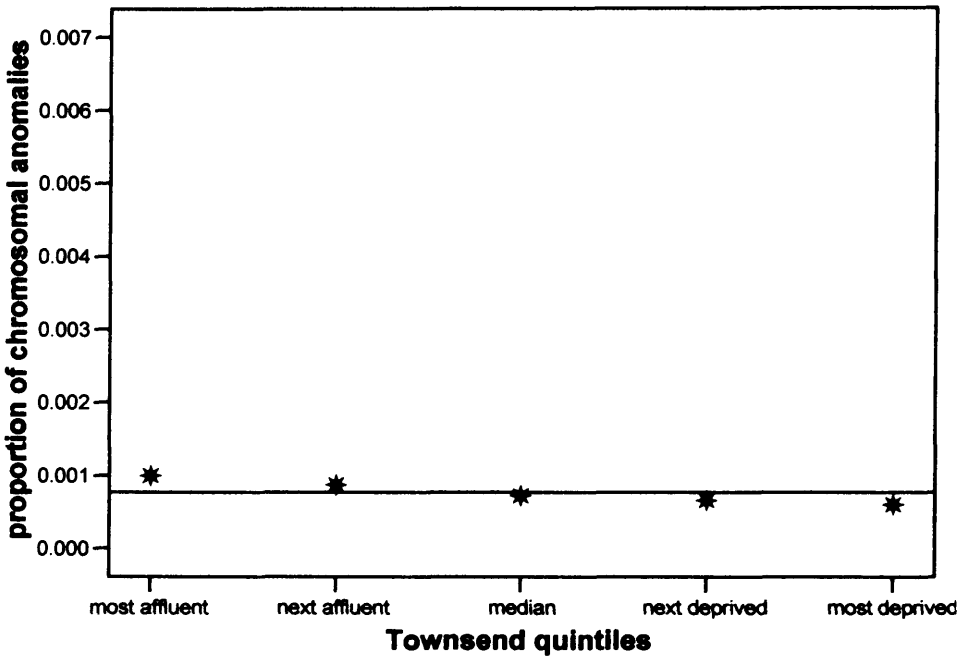


Table A4.10: the proportion of births in Wales with a chromosomal anomaly with 95% confidence intervals by Townsend quintile 1983-1997

Townsend Quintiles	Prop. of Births with a Chromosomal Anomaly (95% CI)	Total Number of Births
Most Affluent	0.000993 (0.000804, 0.001182)	106718
Next Affluent	0.000870 (0.000694, 0.001047)	106836
Median	0.000720 (0.000559, 0.000881)	106928
Next Deprived	0.000656 (0.000502, 0.000809)	106754
Most Deprived	0.000599 (0.000453, 0.000746)	106773

Figure A4.6: the proportion of births with a non-chromosomal anomaly by Townsend quintiles in Wales 1983-1997

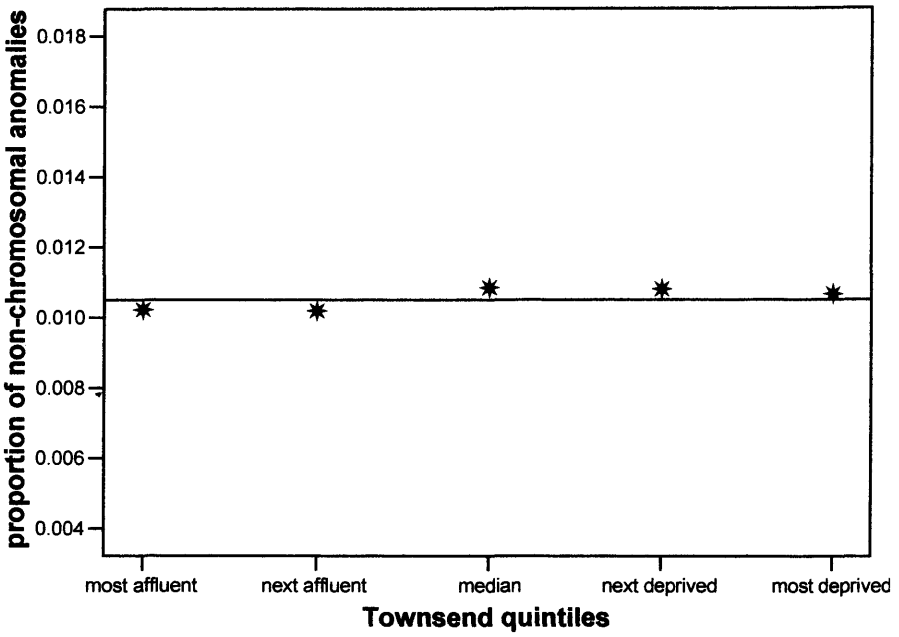


Table A4.11: the proportion of births with a non-chromosomal anomaly with 95% confidence intervals by Townsend quintiles 1983-1997

Townsend Quintiles	Prop. of Births with a Non-Chromosomal Anomaly (95% CI)	Total Number of Births
Most Affluent	0.0102 (0.0096, 0.0108)	106718
Next Affluent	0.0102 (0.0096, 0.0108)	106836
Median	0.0108 (0.0102, 0.0115)	106928
Next Deprived	0.0108 (0.0102, 0.0114)	106754
Most Deprived	0.0107 (0.0101, 0.0113)	106773

Figure A4.7: the proportion of births in Wales with a chromosomal anomaly by hospital of birth 1983-1997

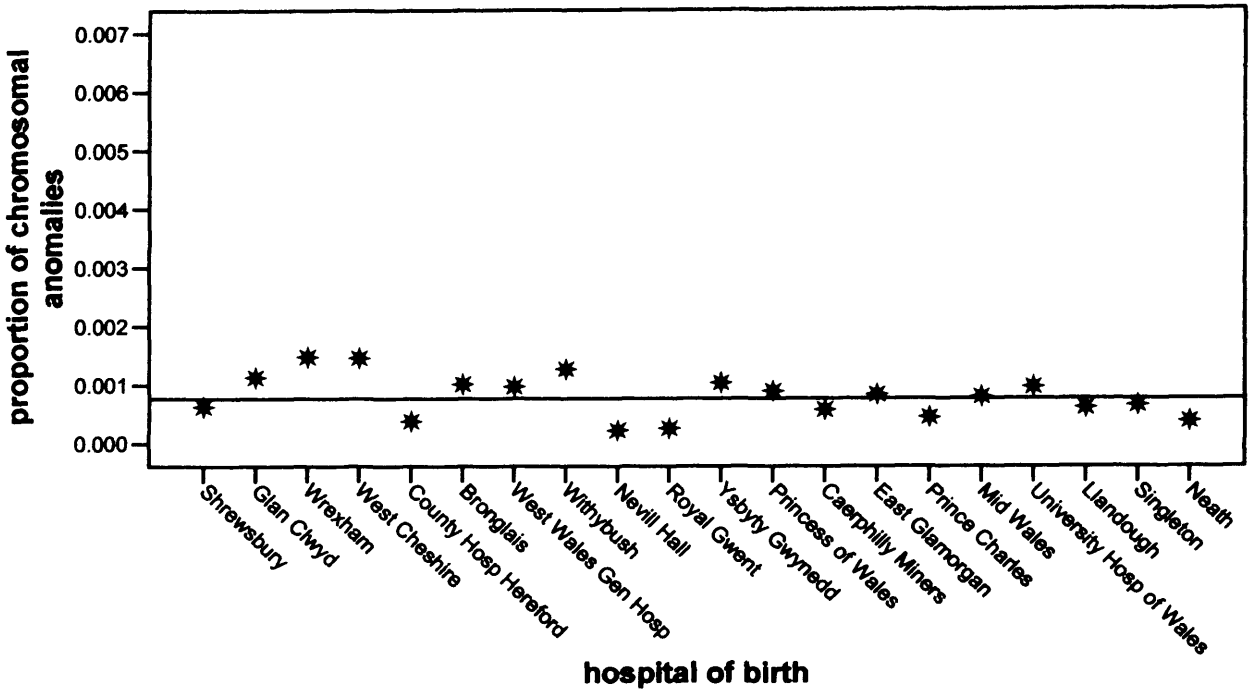


Table A4.12: the proportion of births in Wales with a chromosomal anomaly with 95% confidence intervals by hospital of birth 1983-1997

Hospital of Birth	Prop. of Births with a Chromosomal Anomaly (95% CI)	Total Number of Births
Shrewsbury	0.000630 (0.000013, 0.001247)	6349
Glen Clwyd	0.001130 (0.000793, 0.010133)	38037
Wrexham	0.001478 (0.012894, 0.015386)	34512
West Cheshire	0.001458 (0.000666, 0.002251)	8914
County Hosp Here	0.000382 (0.000000, 0.001129)	2621
Bronglais	0.001014 (0.000352, 0.001676)	8878
West Wales Gen	0.000971 (0.000510, 0.001432)	17511
Withybush	0.001266 (0.000789, 0.001744)	21322
Nevill Hall	0.000223 (0.000045, 0.000401)	26955
Royal Gwent	0.000257 (0.000131, 0.000383)	62274
Ysbyty Gwynedd	0.001035 (0.000676, 0.001393)	30929
Princess of Wales	0.000893 (0.000543, 0.001246)	27981
Caerphilly Miners	0.000576 (0.000236, 0.000917)	19089
East Glamorgan	0.000835 (0.000526, 0.001144)	33550
Prince Charles	0.000448 (0.000195, 0.000702)	26765
Mid Wales	0.000799 (0.000099, 0.001499)	6260
Uni Hosp Wales	0.000966 (0.000659, 0.001274)	39321
Llandough	0.000611 (0.000376, 0.000846)	42541
Singleton	0.000654 (0.000443, 0.000865)	56578
Neath	0.000381 (0.000132, 0.000630)	23622

Figure A4.8: the proportion of births with a non-chromosomal anomaly by hospital of birth 1983-1997

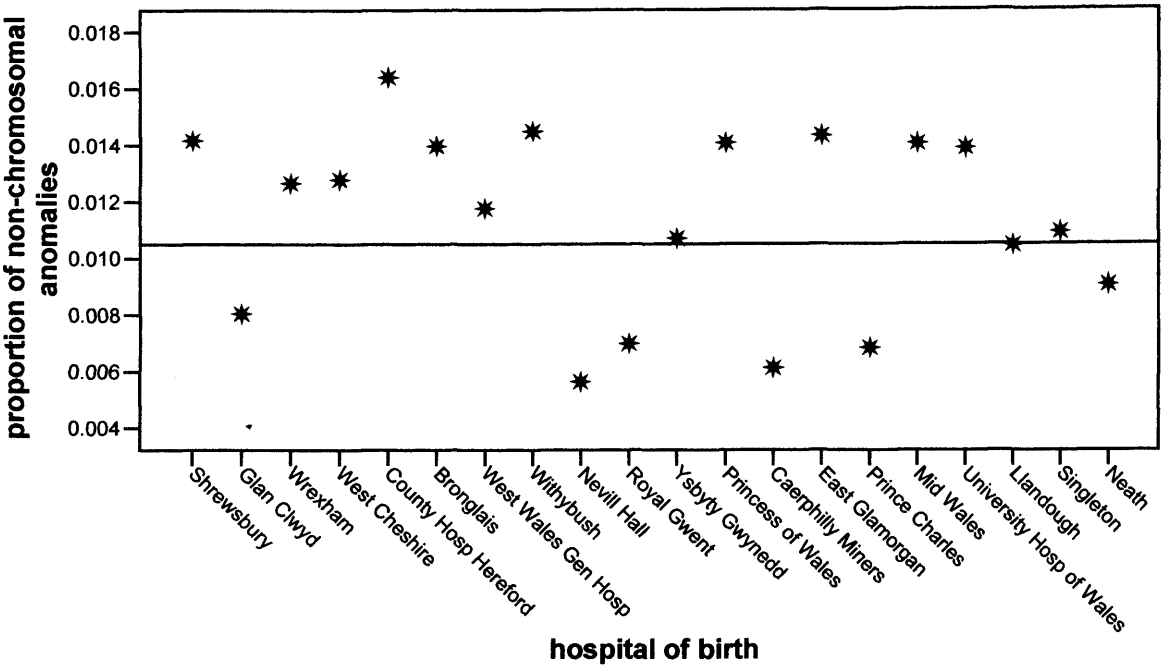


Table A4.13: the proportion of births with a non-chromosomal anomaly with 95% confidence intervals by hospital of birth 1983-1997

Hospital of Birth	Prop. of Births with a Non-Chromosomal Anomaly (95% CI)	Total Number of Births
Shrewsbury	0.0142 (0.0113, 0.0171)	6349
Glen Clwyd	0.0080 (0.0072, 0.0089)	38037
Wrexham	0.0127 (0.0115, 0.0138)	34512
West Cheshire	0.0128 (0.0105, 0.0151)	8914
County Hosp Here	0.0164 (0.0115, 0.0213)	2621
Bronglais	0.0140 (0.0115, 0.0164)	8878
West Wales Gen	0.0118 (0.0102, 0.0134)	17511
Withybush	0.0145 (0.0129, 0.0161)	21322
Nevill Hall	0.0056 (0.0048, 0.0065)	26955
Royal Gwent	0.0070 (0.0063, 0.0076)	62274
Ysbyty Gwynedd	0.0107 (0.0096, 0.0119)	30929
Princess of Wales	0.0141 (0.0127, 0.0155)	27981
Caerphilly Miners	0.0061 (0.0050, 0.0072)	19089
East Glamorgan	0.0143 (0.0131, 0.0156)	33550
Prince Charles	0.0068 (0.0059, 0.0078)	26765
Mid Wales	0.0141 (0.0111, 0.0170)	6260
Uni Hosp Wales	0.0139 (0.0127, 0.0150)	39321
Llandough	0.0105 (0.0095, 0.0114)	42541
Singleton	0.0109 (0.0101, 0.0118)	56578
Neath	0.0091 (0.0079, 0.0103)	23622

Table A4.14: the proportion of births in Wales with a chromosomal anomaly with 95% confidence intervals by gender 1983-1997

Gender	Prop. of Births with a Chromosomal Anomaly (95% CI)	Total Number of Births
Male	0.000815 (0.000708, 0.000922)	273587
Female	0.000718 (0.000615, 0.000821)	260422

Table A4.15: the proportion of births in Wales with a non-chromosomal anomaly with confidence intervals by gender 1983-1997

Gender	Prop. of Births with a Non-Chromosomal Anomaly (95% CI)	Total Number of Births
Male	0.0121 (0.0117, 0.0125)	273587
Female	0.0089 (0.0086, 0.0093)	260422

Figure A4.9: hotspots of excess density of chromosomal anomalies after adjustment for statistically significant confounders in the 20km square region centred over Nant y Gwyddon 1983-1997

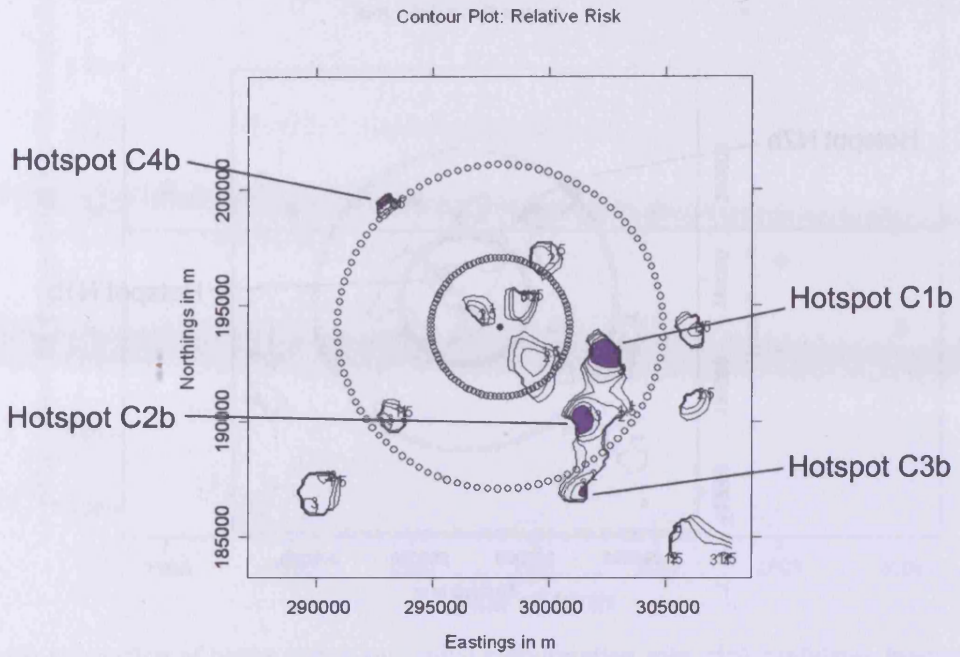


Table A4.16: results of the global significance test on hotspots of excess density of chromosomal anomalies after adjustment for statistically significant confounders in the 20km square region centred over Nant y Gwyddon 1983-1997

Hotspot		C1b	C2b	C3b	C4b
Number of births		443	243	92	584
Number of chromosomal anomalies		2	2	1	2
Observed mean adjusted relative risk		5.660	4.985	4.072	8.681
p-value	test 1	0.004	0.008	0.051	<0.001
	test 2	0.010	0.010	0.068	0.019
	test 3	0.054	0.010	0.071	0.020

Hotspots C2b and C4b are statistically significant.

Figure A4.10: hotspots of excess density of non-chromosomal anomalies after adjustment for statistically significant confounders in the 20km square region centred over Nant y Gwyddon 1983-1997

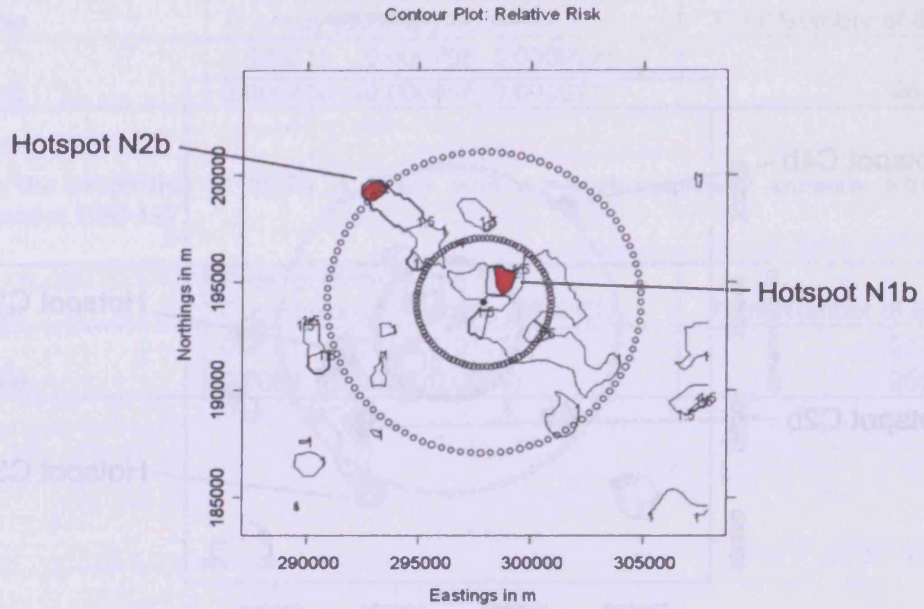


Table A4.17: results of the global significance test on hotspots of excess density of non-chromosomal anomalies after adjustment for statistically significant confounders in the 20km square region centred over Nant y Gwyddon 1983-1997

Hotspot		N1b	N2b
Number of births		543	259
Number of non-chromosomal anomalies		20	13
Observed mean adjusted relative risk		2.285	2.234
p-value	test 1	<0.001	<0.001
	test 2	<0.001	<0.001
	test 3	<0.001	<0.001

Hotspots N1b and N2b are statistically significant.

Figure A4.11: the proportion of births with a congenital malformation by year of birth 1998-2004

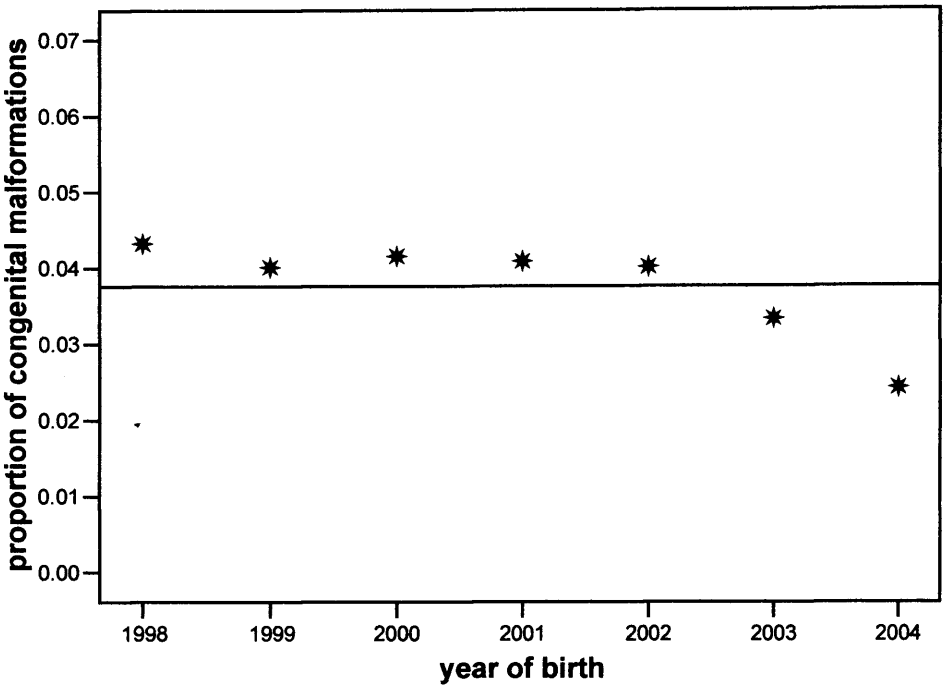


Table A4.18: proportion of births with a congenital malformation with 95% confidence intervals by year of birth 1998-2004

Year of Birth	Prop. of Births with a Congenital Anomaly (95% CI)	Total Number of Births
1998	0.0432 (0.0411, 0.0454)	33515
1999	0.0401 (0.0380, 0.0423)	32177
2000	0.0416 (0.0393, 0.0438)	31430
2001	0.0409 (0.0387, 0.0431)	30843
2002	0.0402 (0.0380, 0.0424)	30477
2003	0.0333 (0.0313, 0.0353)	31570
2004	0.0243 (0.0226, 0.0259)	32522

Table A4.19: the proportion of births in Wales with a congenital malformation with 95% confidence intervals by gender 1998-2004

Gender	Prop. of Births with a Congenital Anomaly (95% CI)	Total Number of Births
Female	0.0332 (0.0322, 0.0343)	108012
Male	0.0417 (0.0405, 0.0429)	114488
Indiscriminate	0.2353 (0.0851, 0.3855)	34

Figure A4.12: the proportion of births in Wales with a congenital malformation by maternal age 1998-2004

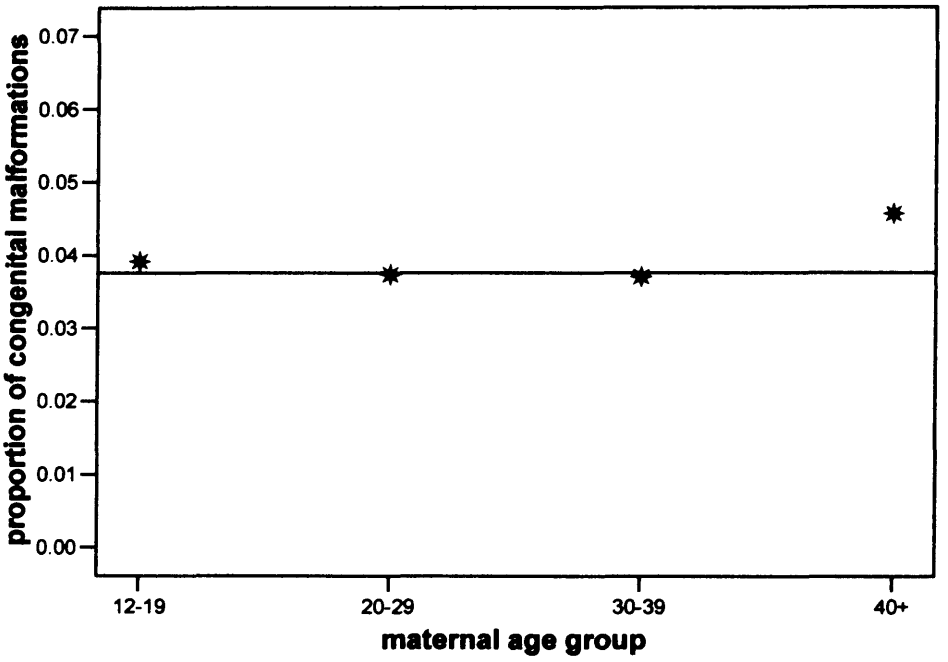


Table A4.20: the proportion of births in Wales with a congenital malformation with 95% confidence intervals by maternal age 1998-2004

Maternal Age	Prop. of Births with a Congenital Anomaly (95% CI)	Total Number of Births
12-19	0.0392 (0.0366, 0.0417)	22473
20-29	0.0374 (0.0362, 0.0384)	109459
30-39	0.0371 (0.0358, 0.0383)	85348
40+	0.0457 (0.0398, 0.0515)	4883

Figure A4.13: the proportion of births in Wales with a congenital anomaly by Townsend quintile 1998-2004

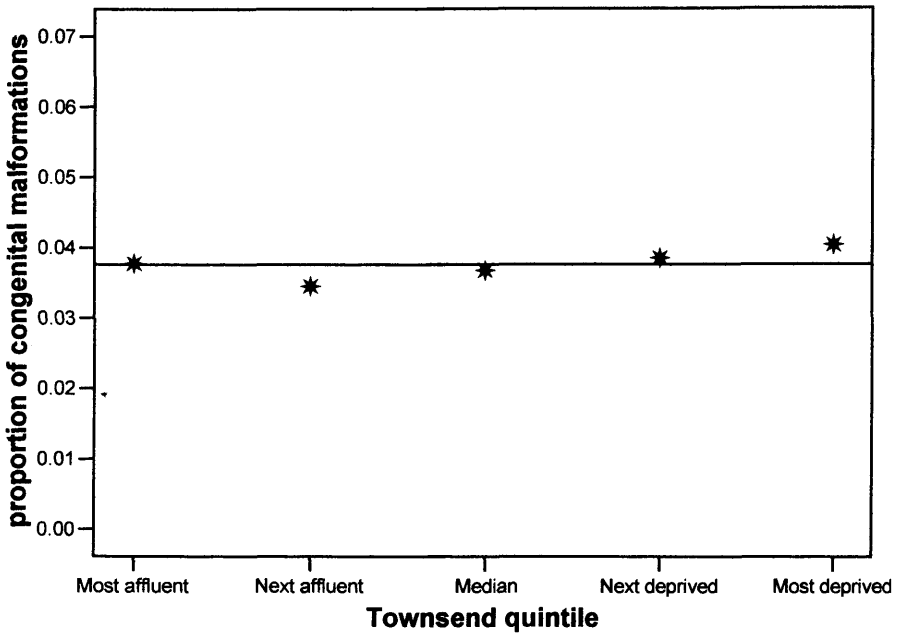


Table A4.21: the proportion of births in Wales with a congenital malformation with 95% confidence intervals by Townsend quintile 1998-2004

Townsend quintile	Prop. of Births with a Congenital Anomaly (95% CI)	Total Number of Births
Most affluent	0.0377 (0.0358, 0.0396)	38023
Next affluent	0.0345 (0.0327, 0.0363)	38562
Median	0.0367 (0.0350, 0.0385)	42535
Next deprived	0.0385 (0.0367, 0.0403)	45850
Most deprived	0.0404 (0.0387, 0.0420)	55147

Figure A4.14: the proportion of births in Wales with a congenital malformation by hospital of birth 1998-2004

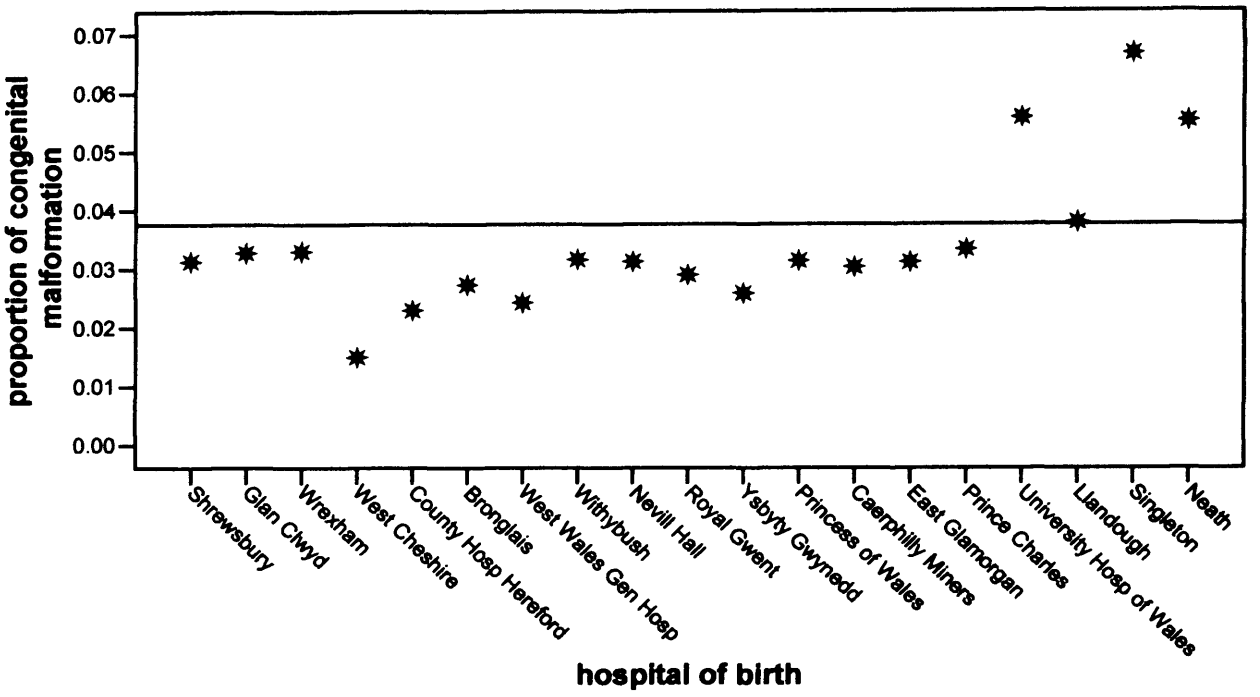


Table A4.22: the proportion of births in Wales with a congenital malformation with 95% confidence intervals by hospital of birth in Wales 1998-2004

Hospital of Birth	Prop. of Births with a Congenital Anomaly (95% CI)	Total Number of Births
Shrewsbury	0.0313 (0.0256, 0.0369)	3644
Glan Clwyd	0.0328 (0.0301, 0.0355)	16414
Wrexham	0.0330 (0.0302, 0.0358)	15538
West Cheshire	0.0150 (0.0110, 0.0190)	3598
County Hosp Here	0.0230 (0.0142, 0.0317)	1132
Bronglais	0.0273 (0.0225, 0.0321)	4433
West Wales Gen	0.0243 (0.0211, 0.0274)	9223
Withybush	0.0316 (0.0279, 0.0353)	8476
Nevill Hall	0.0313 (0.0282, 0.0344)	11993
Royal Gwent	0.0290 (0.0269, 0.0311)	25414
Ysbyty Gwynedd	0.0258 (0.0231, 0.0286)	13119
Princess of Wales	0.0314 (0.0282, 0.0346)	11504
Caerphilly Miners	0.0303 (0.0253, 0.0353)	4550
East Glamorgan	0.0312 (0.0283, 0.0340)	14315
Prince Charles	0.0333 (0.0299, 0.0367)	10536
Uni Hosp of Wales	0.0557 (0.0525, 0.0589)	19722
Llandough	0.0379 (0.0351, 0.0407)	17720
Singleton	0.0666 (0.0635, 0.0698)	23983
Neath	0.0551 (0.0499, 0.0604	7220

Appendix Five - Sex Ratios

Table A5.1: the number of births and sex ratio by region in Cardiff and the Vale of Glamorgan 1975-1979

Region		Number of Births	Proportion of Boys
Cardiff	Butetown	550	0.476
	Canton	791	0.518
	Cathays	1249	0.504
	Cyncoed	2734	0.512
	Ely	1412	0.505
	Fairwater	957	0.491
	Gabalfa	597	0.503
	Grangetown	610	0.505
	Lisvane and St. Mellons	229	0.528
	Llandaff	676	0.494
	Llanishen	773	0.519
	Plasnewydd	537	0.512
	Radyr and St. Fagans	317	0.486
	Rhiwbina	529	0.558
	Riverside	821	0.475
	Roath	995	0.471
	Rumney	1531	0.521
	Splott	1156	0.525
	Whitchurch and Tongwynlais	674	0.527
Vale of Glamorgan	Barry	2808	0.511
	Cowbridge	315	0.552
	Dinas Powys	489	0.476
	Llandow	116	0.440
	Llantwit Major	766	0.522
	Penarth	1345	0.517
	Peterson-Super-Ely	81	0.494
	Rhoose	209	0.450
	St. Athan	371	0.482
	Sully	53	0.472
	Wenvoe	97	0.412

Table A5.2: the number of births and sex ratio by region in Cardiff and the Vale of Glamorgan 1980-1984

Region		Number of Births	Proportion of Boys
Cardiff	Butetown	509	0.527
	Canton	1009	0.541
	Cathays	1160	0.531
	Cyncoed	2607	0.513
	Ely	2078	0.505
	Fairwater	1019	0.495
	Gabalfa	627	0.510
	Grangetown	893	0.531
	Lisvane and St. Mellons	785	0.504
	Llandaff	666	0.522
	Llanishen	959	0.506
	Plasnewydd	821	0.515
	Radyr and St. Fagans	254	0.535
	Rhiwbina	487	0.487
	Riverside	792	0.500
	Roath	686	0.503
	Rumney	1777	0.514
	Splott	1393	0.514
	Whitchurch and Tongwynlais	686	0.500
Vale of Glamorgan	Barry	3417	0.509
	Cowbridge	258	0.539
	Dinas Powys	601	0.537
	Llandow	86	0.605
	Llantwit Major	630	0.495
	Penarth	1297	0.503
	Peterson-Super-Ely	78	0.551
	Rhoose	288	0.458
	St. Athan	409	0.509
	Sully	121	0.471
	Wenvoe	83	0.434

Table A5.3: the number of births and sex ratio by region in Cardiff and the Vale of Glamorgan 1985-1989

Region		Number of Births	Proportion of Boys
Cardiff	Butetown	102	0.451
	Canton	386	0.479
	Cathays	434	0.447
	Cyncoed	1239	0.521
	Ely	736	0.515
	Fairwater	238	0.475
	Gabalfa	490	0.494
	Grangetown	380	0.495
	Lisvane and St. Mellons	760	0.525
	Llandaff	255	0.522
	Llanishen	885	0.521
	Plasnewydd	1062	0.508
	Radyr and St. Fagans	132	0.455
	Rhiwbina	261	0.517
	Riverside	255	0.490
	Roath	468	0.474
	Rumney	896	0.516
	Splott	895	0.504
	Whitchurch and Tongwynlais	472	0.517
Vale of Glamorgan	Barry	1016	0.469
	Cowbridge	97	0.515
	Dinas Powys	146	0.479
	Llandow	64	0.594
	Llantwit Major	214	0.551
	Penarth	396	0.515
	Peterson-Super-Ely	18	0.556
	Rhoose	92	0.565
	St. Athan	152	0.467
	Sully	65	0.446
	Wenvoe	34	0.559

Table A5.4: the number of births and sex ratio by region in Cardiff and the Vale of Glamorgan 1990-1994

Region		Number of Births	Proportion of Boys
Cardiff	Butetown	351	0.499
	Canton	1139	0.519
	Cathays	625	0.523
	Cyncoed	1932	0.518
	Ely	2524	0.517
	Fairwater	875	0.510
	Gabalfa	838	0.516
	Grangetown	1212	0.538
	Lisvane and St. Mellons	2098	0.523
	Llandaff	632	0.464
	Llanishen	1620	0.525
	Plasnewydd	1062	0.500
	Radyr and St. Fagans	249	0.518
	Rhiwbina	462	0.513
	Riverside	908	0.525
	Roath	619	0.506
	Rumney	1468	0.516
	Splott	1801	0.529
	Whitchurch and Tongwynlais	938	0.479
Vale of Glamorgan	Barry	3723	0.529
	Cowbridge	243	0.531
	Dinas Powys	427	0.527
	Llandow	124	0.508
	Llantwit Major	575	0.466
	Penarth	1344	0.516
	Peterson-Super-Ely	46	0.522
	Rhoose	256	0.547
	St. Athan	472	0.557
	Sully	216	0.532
	Wenvoe	117	0.410

Table A5.5: the number of births and sex ratio by region in Cardiff and the Vale of Glamorgan 1995-1999

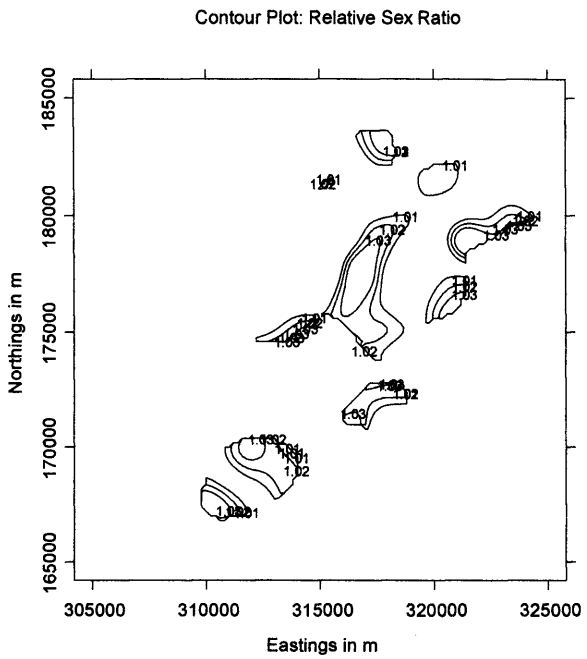
Region		Number of Births	Proportion of Boys
Cardiff	Butetown	357	0.487
	Canton	1027	0.523
	Cathays	445	0.530
	Cyncoed	1724	0.497
	Ely	2309	0.519
	Fairwater	786	0.518
	Gabalfa	837	0.524
	Grangetown	1254	0.518
	Lisvane and St. Mellons	2069	0.511
	Llandaff	543	0.457
	Llanishen	1581	0.489
	Plasnewydd	891	0.515
	Radyr and St. Fagans	234	0.513
	Rhiwbina	444	0.482
	Riverside	704	0.504
	Roath	626	0.513
	Rumney	1306	0.498
	Splott	1691	0.504
	Whitchurch and Tongwynlais	849	0.495
Vale of Glamorgan	Barry	3345	0.502
	Cowbridge	229	0.533
	Dinas Powys	390	0.503
	Llandow	108	0.491
	Llantwit Major	556	0.527
	Penarth	1315	0.513
	Peterson-Super-Ely	44	0.523
	Rhoose	315	0.508
	St. Athan	408	0.527
	Sully	143	0.455
	Wenvoe	112	0.473

Table A5.6: the number of births and sex ratio by region in Cardiff and the Vale of Glamorgan 2000-2004

Region		Number of Births	Proportion of Boys
Cardiff	Butetown	277	0.502
	Canton	613	0.527
	Cathays	210	0.519
	Cyncoed	1005	0.524
	Ely	1452	0.506
	Fairwater	538	0.513
	Gabalfa	544	0.520
	Grangetown	883	0.506
	Lisvane and St. Mellons	1456	0.519
	Llandaff	272	0.518
	Llanishen	1065	0.509
	Plasnewydd	523	0.528
	Radyr and St. Fagans	264	0.504
	Rhiwbina	310	0.535
	Riverside	606	0.508
	Roath	337	0.534
	Rumney	310	0.474
	Splott	1069	0.503
	Whitchurch and Tongwynlais	472	0.553
Vale of Glamorgan	Barry	2144	0.514
	Cowbridge	123	0.577
	Dinas Powys	332	0.512
	Llandow	29	0.552
	Llantwit Major	341	0.519
	Penarth	805	0.522
	Peterson-Super-Ely	60	0.517
	Rhoose	233	0.485
	St. Athan	175	0.480
	Sully	142	0.479
	Wenvoe	65	0.615

Figure A5.1: relative density contours for excess male births in Cardiff and the Vale of Glamorgan 1990-1994 calculated with the density of a) all births and b) female births as the denominator

a)



b)

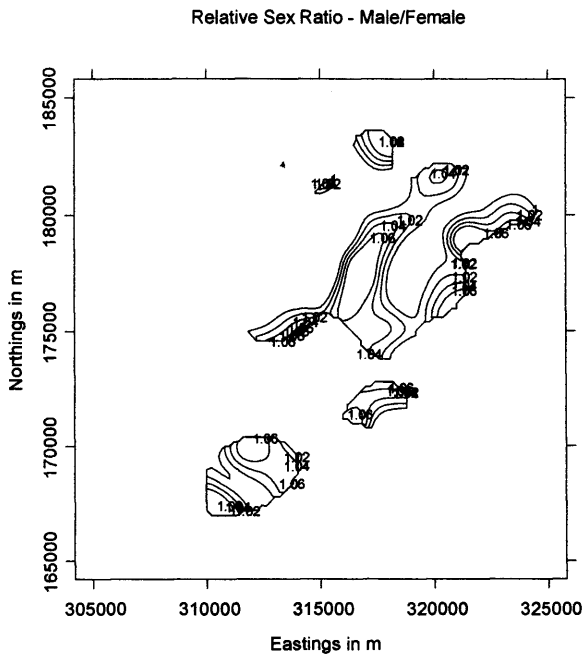


Figure A5.2: the sex ratio by maternal age in Wales 1983-1997

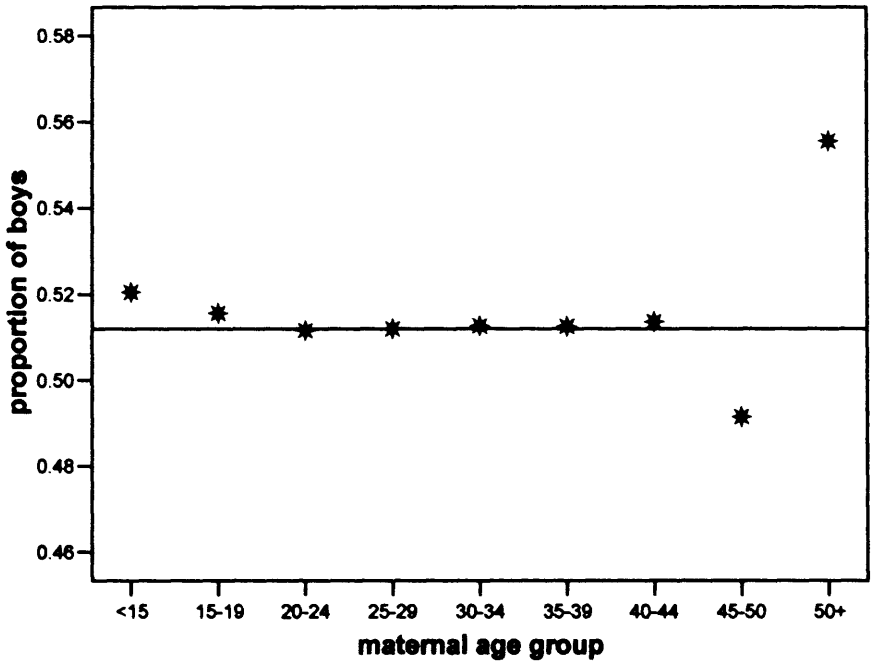


Figure A5.3: the sex ratio by maternal age in the USA 1940-2002

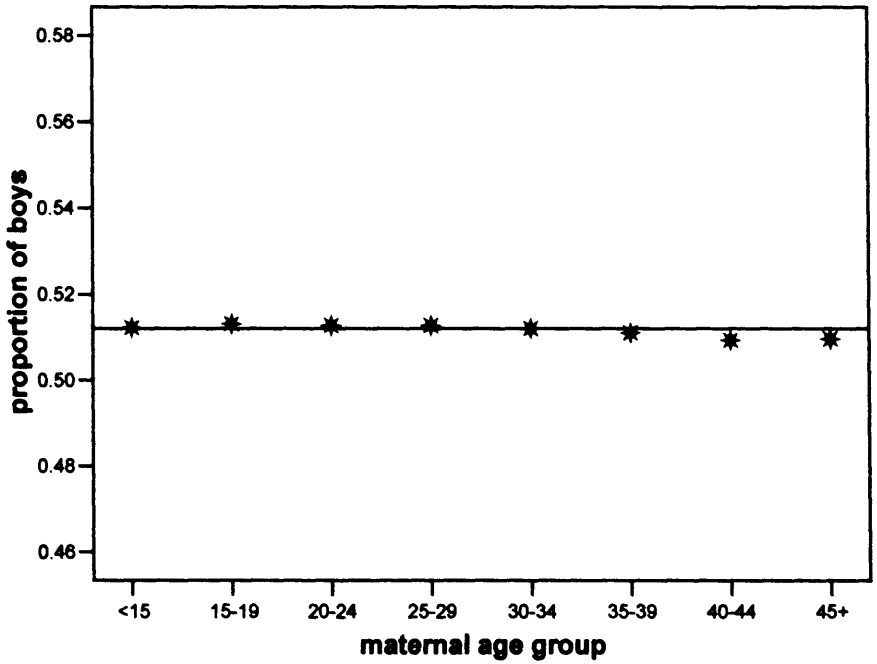


Figure A5.4: relative density contours of births to mothers over 30 years of age in Cardiff and the Vale of Glamorgan a) 1990-1994; b) 1995-1999; c) 2000-2004

a)

b)

c)

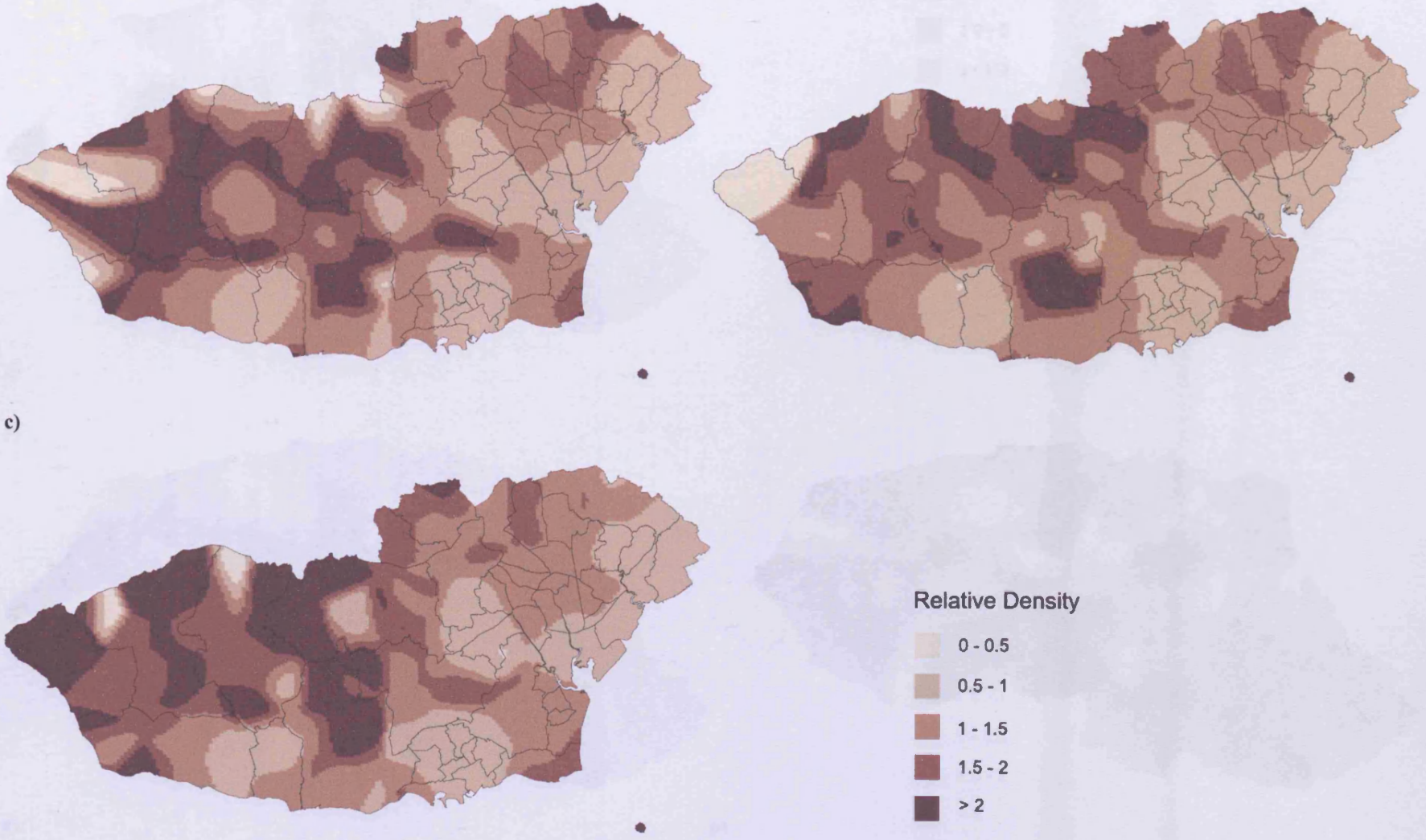
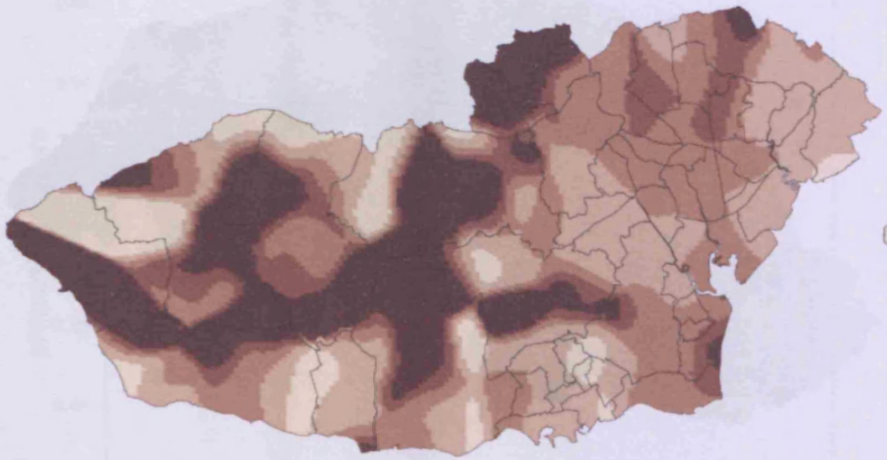
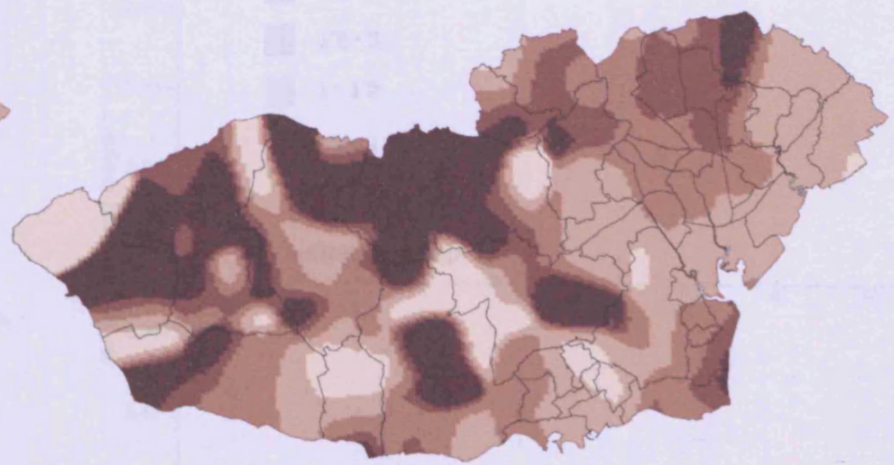


Figure A5.5: relative density contours of births to mothers over 35 years of age in Cardiff and the Vale of Glamorgan a) 1990-1994; b) 1995-1999; c) 2000-2004

a)



b)



c)

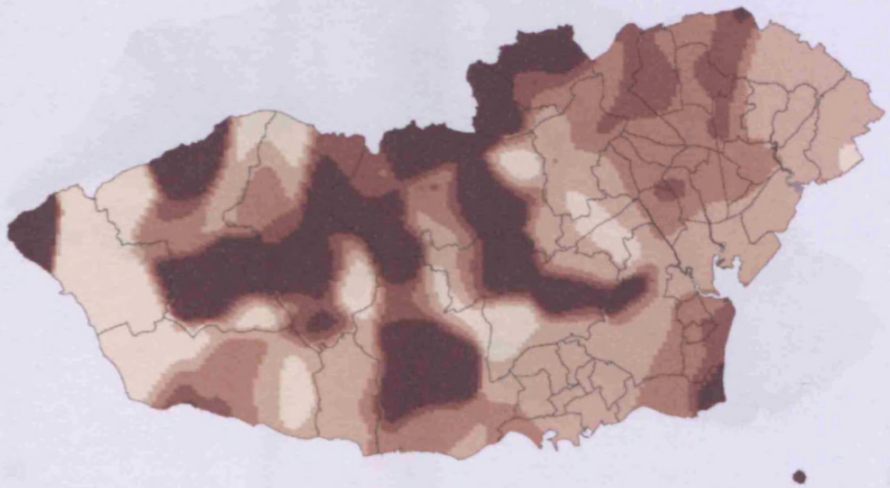
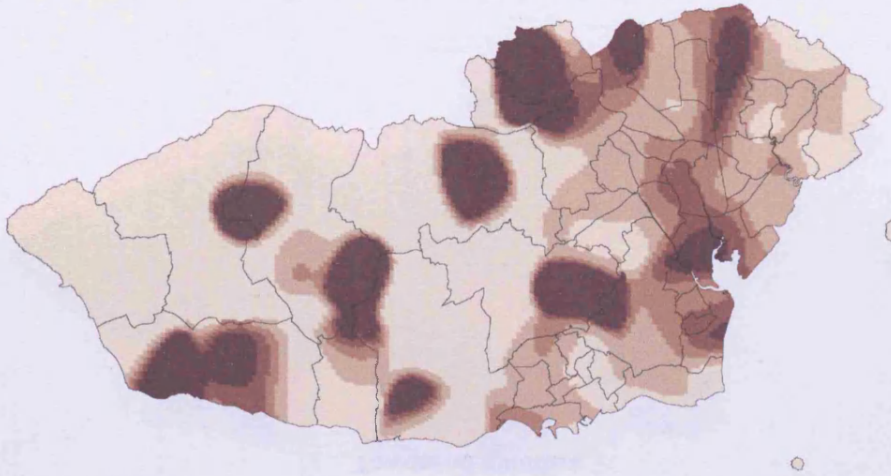
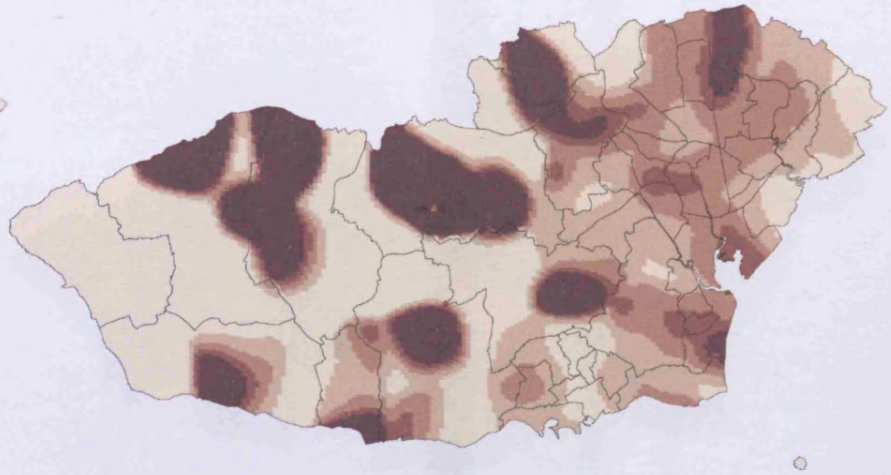


Figure A5.6: relative density contours of births to mothers over 40 years of age in Cardiff and the Vale of Glamorgan a) 1990-1994; b) 1995-1999; c) 2000-2004

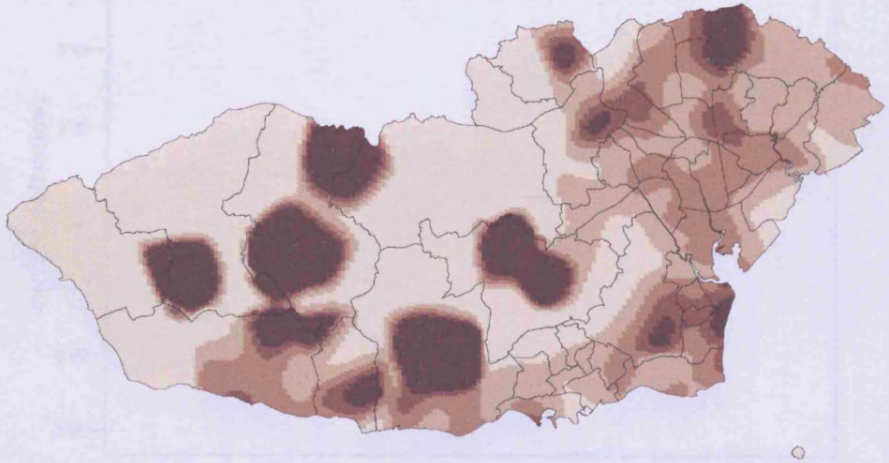
a)



b)



c)



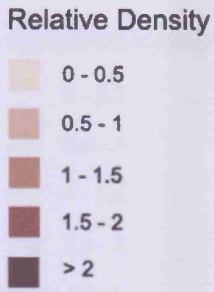


Figure A5.7: the sex ratio by Townsend quintile in Wales 1983-1997

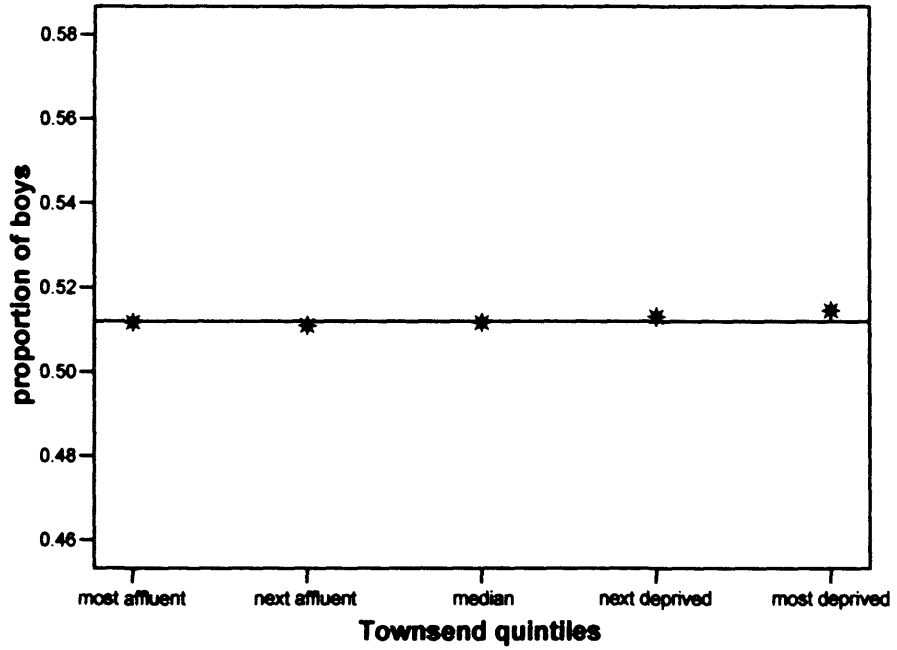


Figure A5.8: relative density contours of births to the most affluent families in Cardiff and the Vale of Glamorgan a) 1990-1994; b) 1995-1999; c) 2000-2004

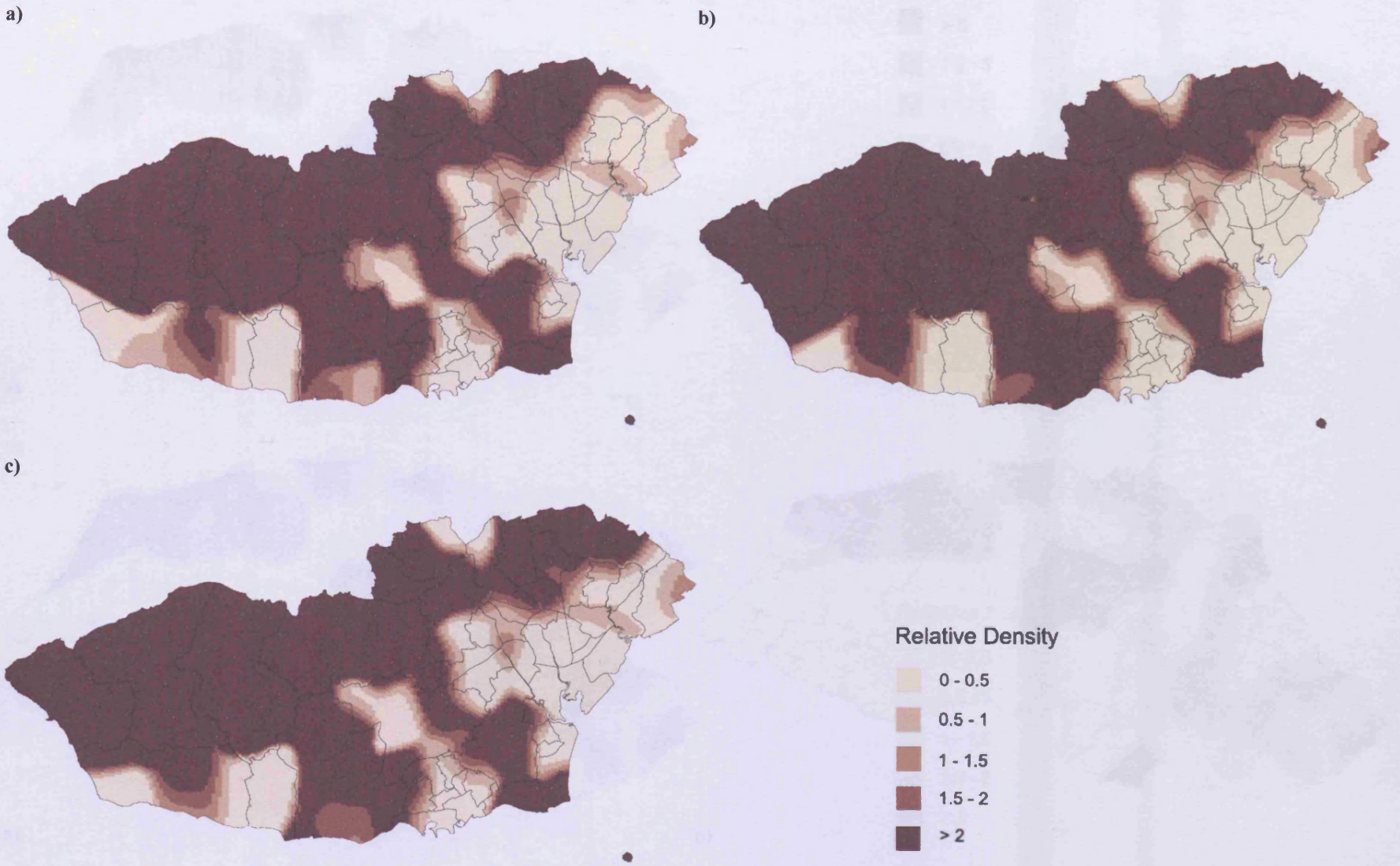
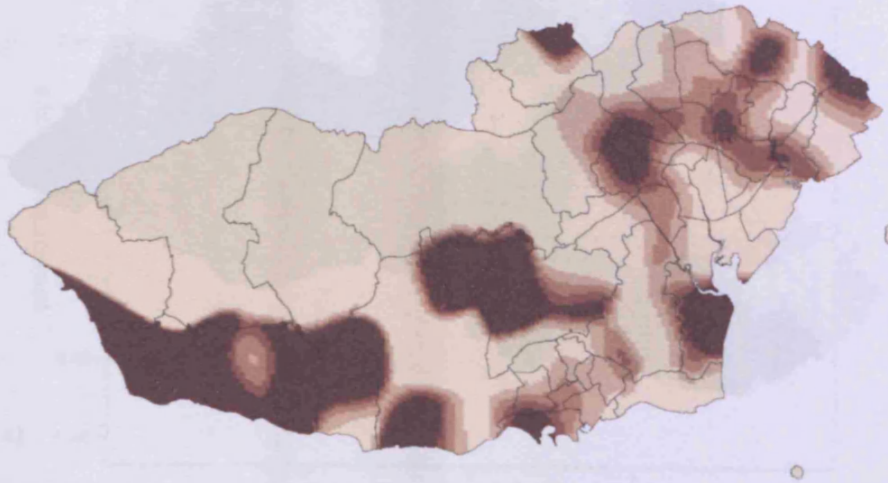
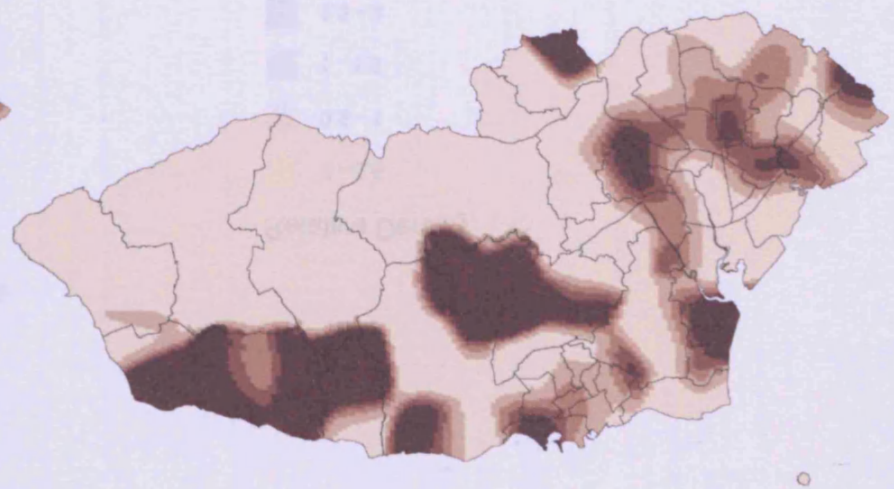


Figure A5.9: relative density contours of births to the next affluent families in Cardiff and the Vale of Glamorgan a) 1990-1994; b) 1995-1999; c) 2000-2004

a)



b)



c)

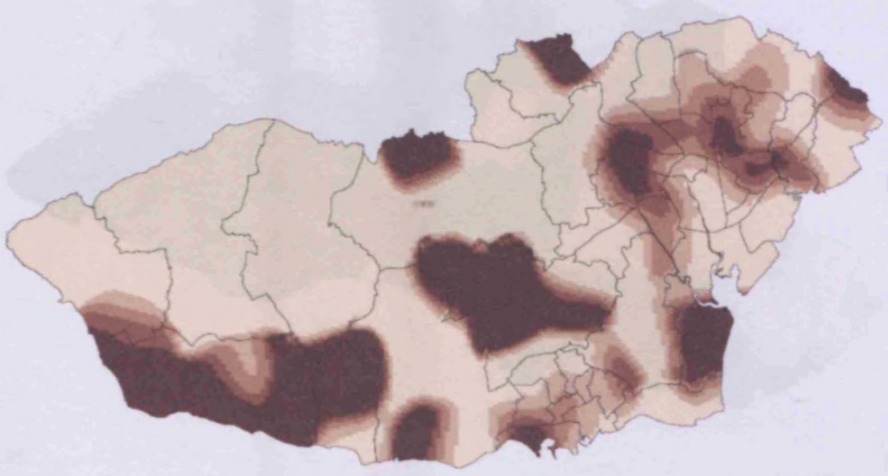


Figure A5.10: relative density contours of births to moderately deprived families in Cardiff and the Vale of Glamorgan a) 1990-1994; b) 1995-1999; c) 2000-2004

a)

b)

c)

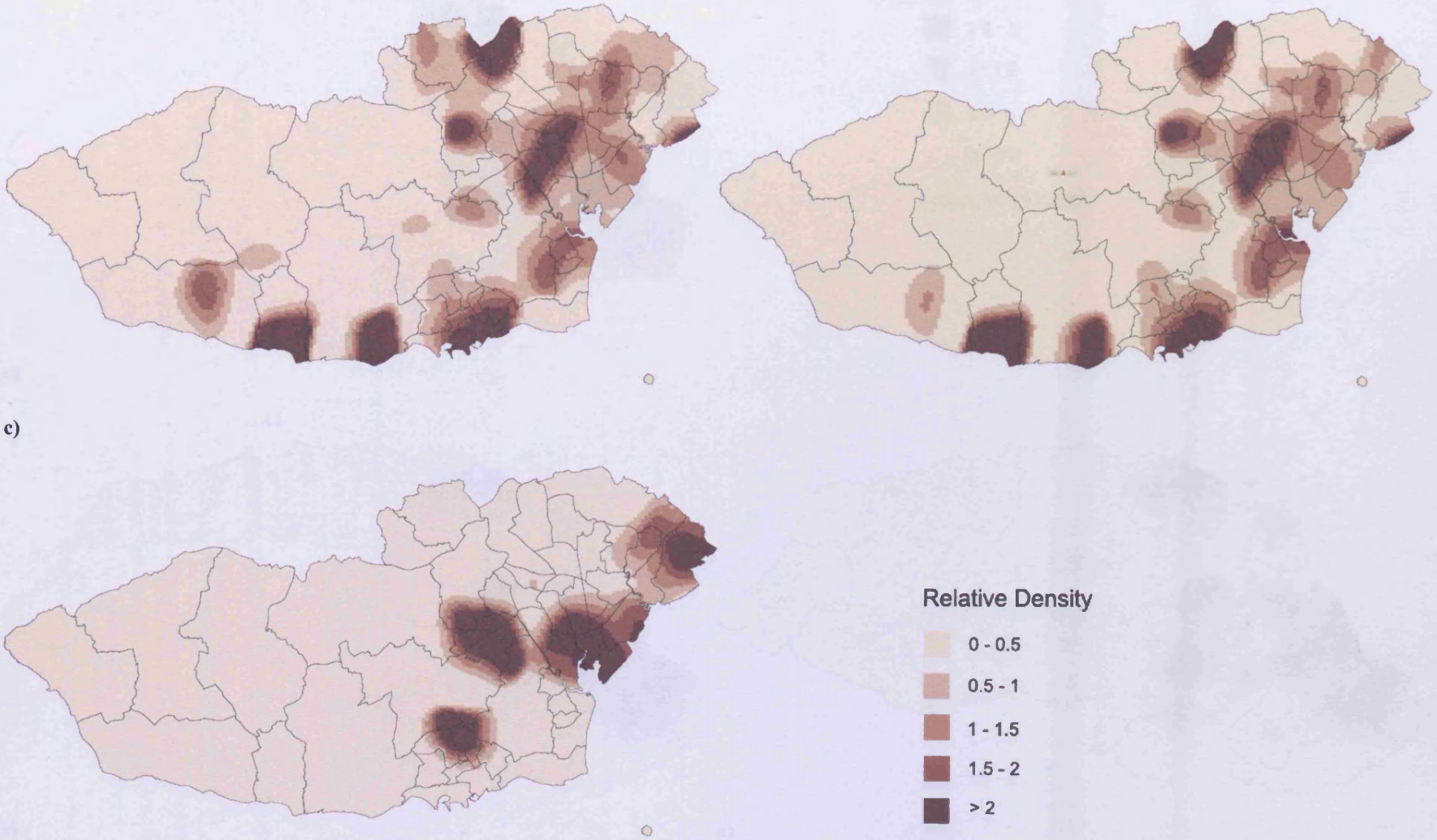


Figure A5.11: relative density contours of births to the next deprived families in Cardiff and the Vale of Glamorgan a) 1990-1994; b) 1995-1999; c) 2000-2004

a)

b)

c)

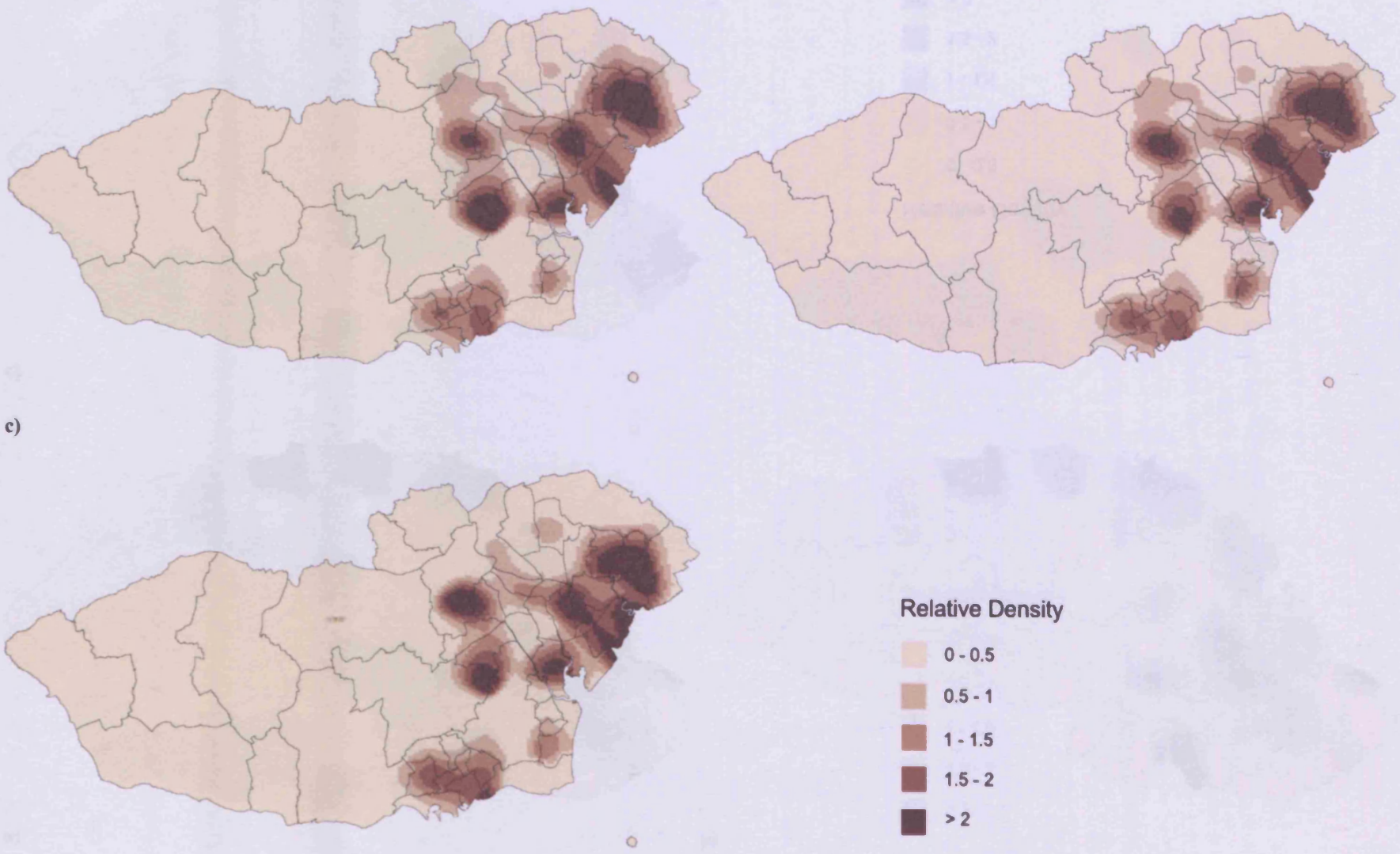


Figure A5.12: relative density contours of births to the most deprived families in Cardiff and the Vale of Glamorgan a) 1990-1994; b) 1995-1999; c) 2000-2004

a)

b)

c)

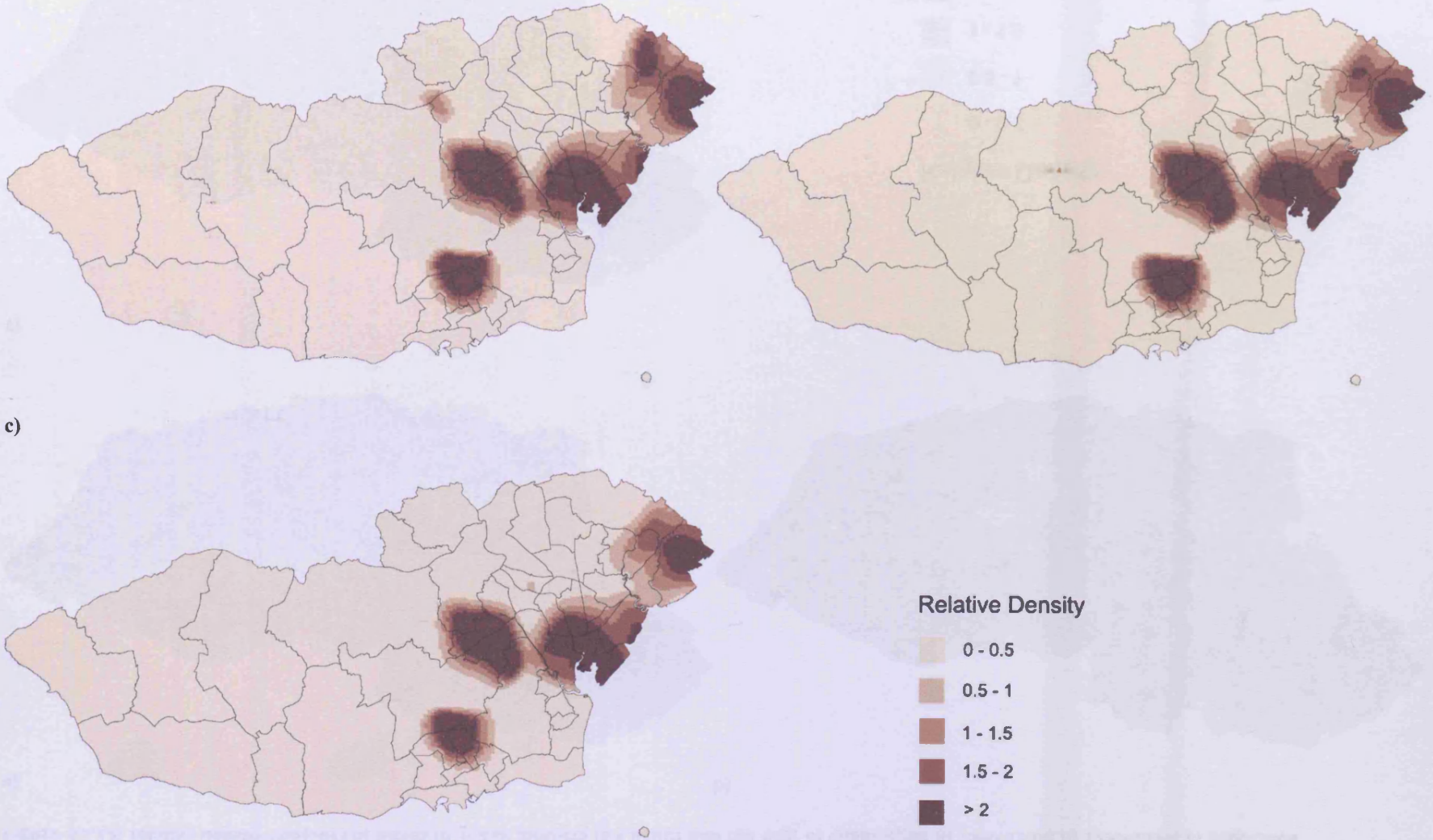


Figure A5.13: relative density contours of births to White mothers in Cardiff and the Vale of Glamorgan a) 1990-1994; b) 1995-1999; c) 2000-2004

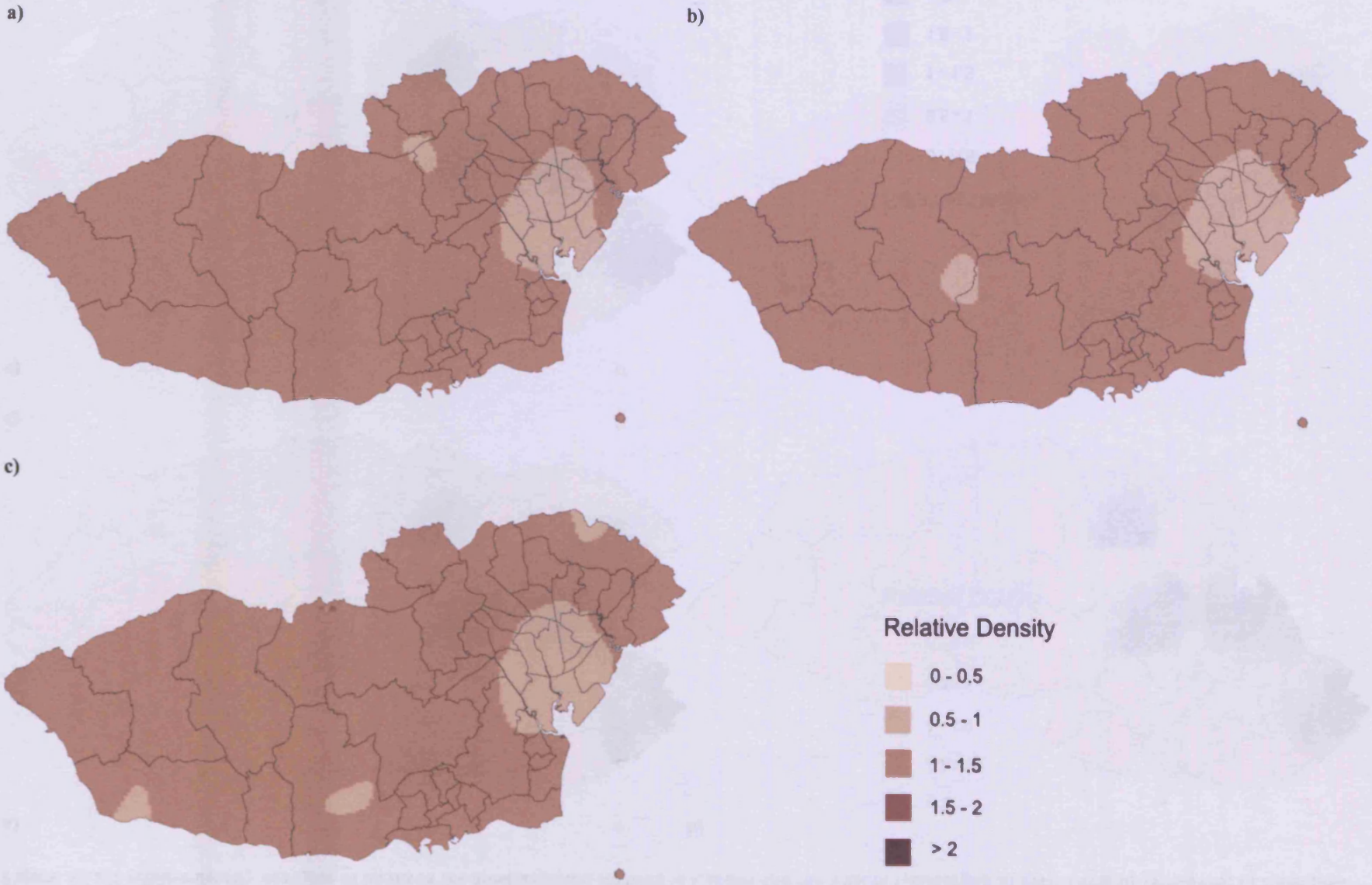


Figure A5.14: relative density contours of births to Black mothers in Cardiff and the Vale of Glamorgan a) 1990-1994; b) 1995-1999; c) 2000-2004

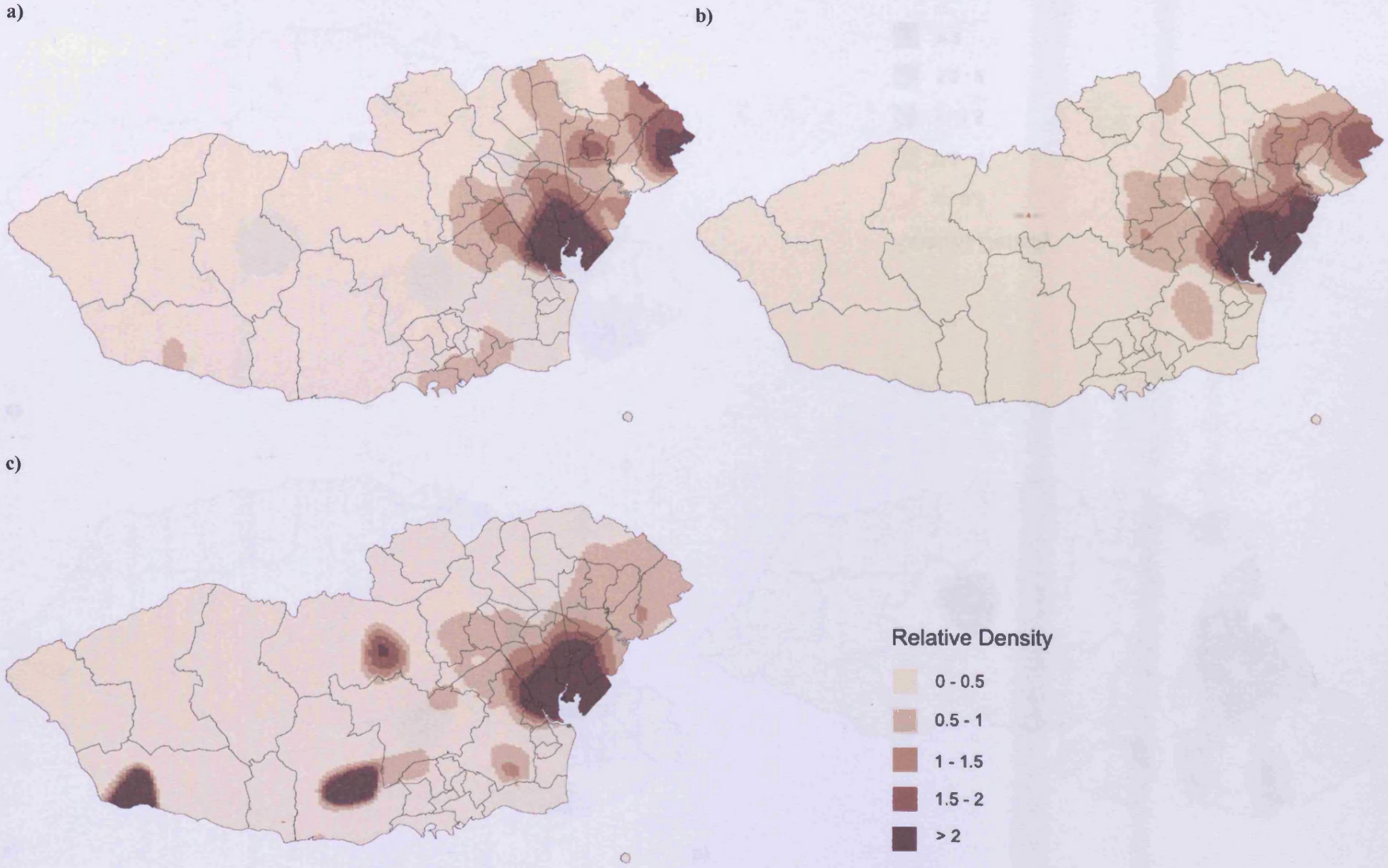


Figure A5.15: relative density contours of births to Indian mothers in Cardiff and the Vale of Glamorgan a) 1990-1994; b) 1995-1999; c) 2000-2004

a)

b)

c)

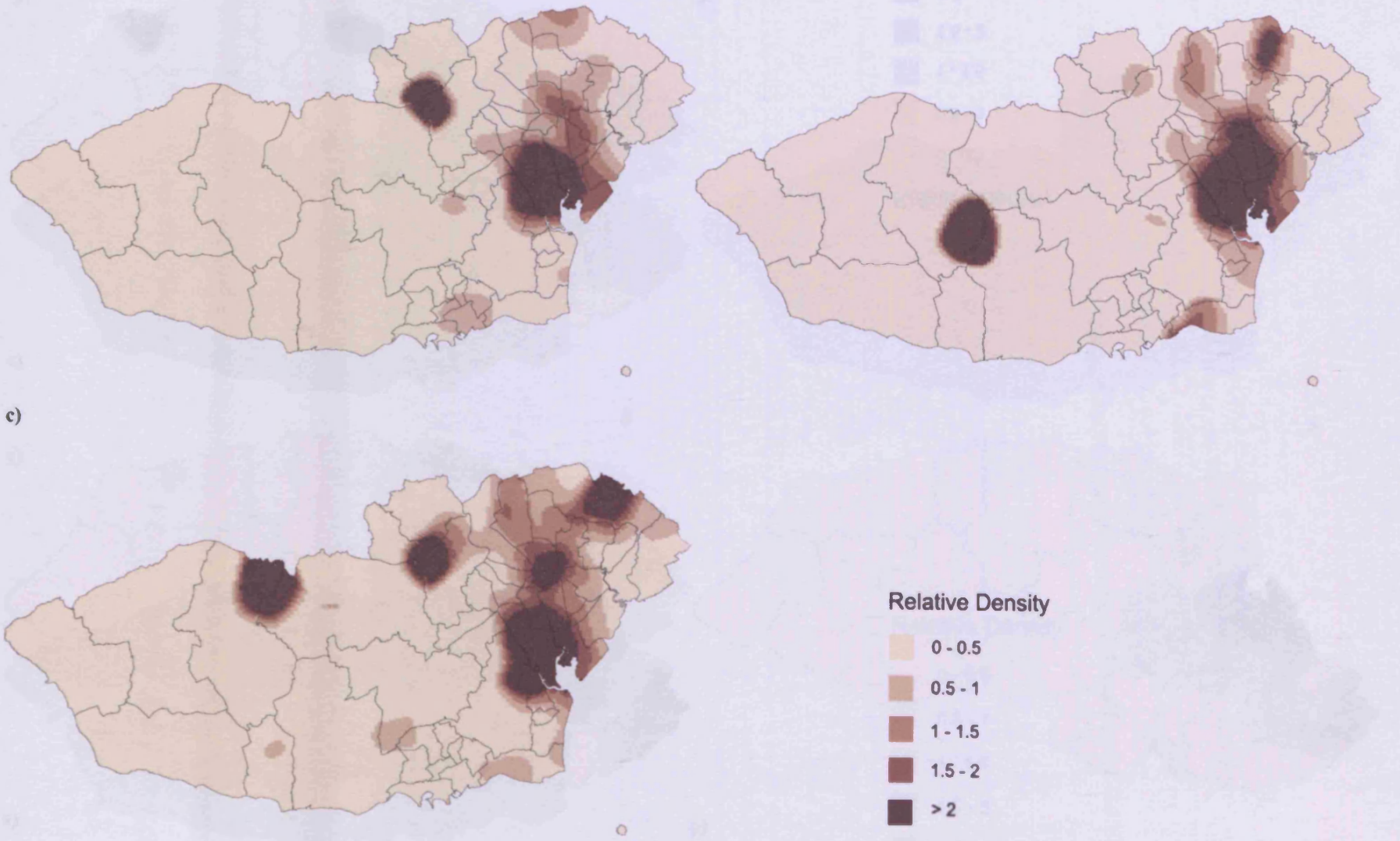


Figure A5.16: relative density contours of births to Pakistani mothers in Cardiff and the Vale of Glamorgan a) 1990-1994; b) 1995-1999; c) 2000-2004

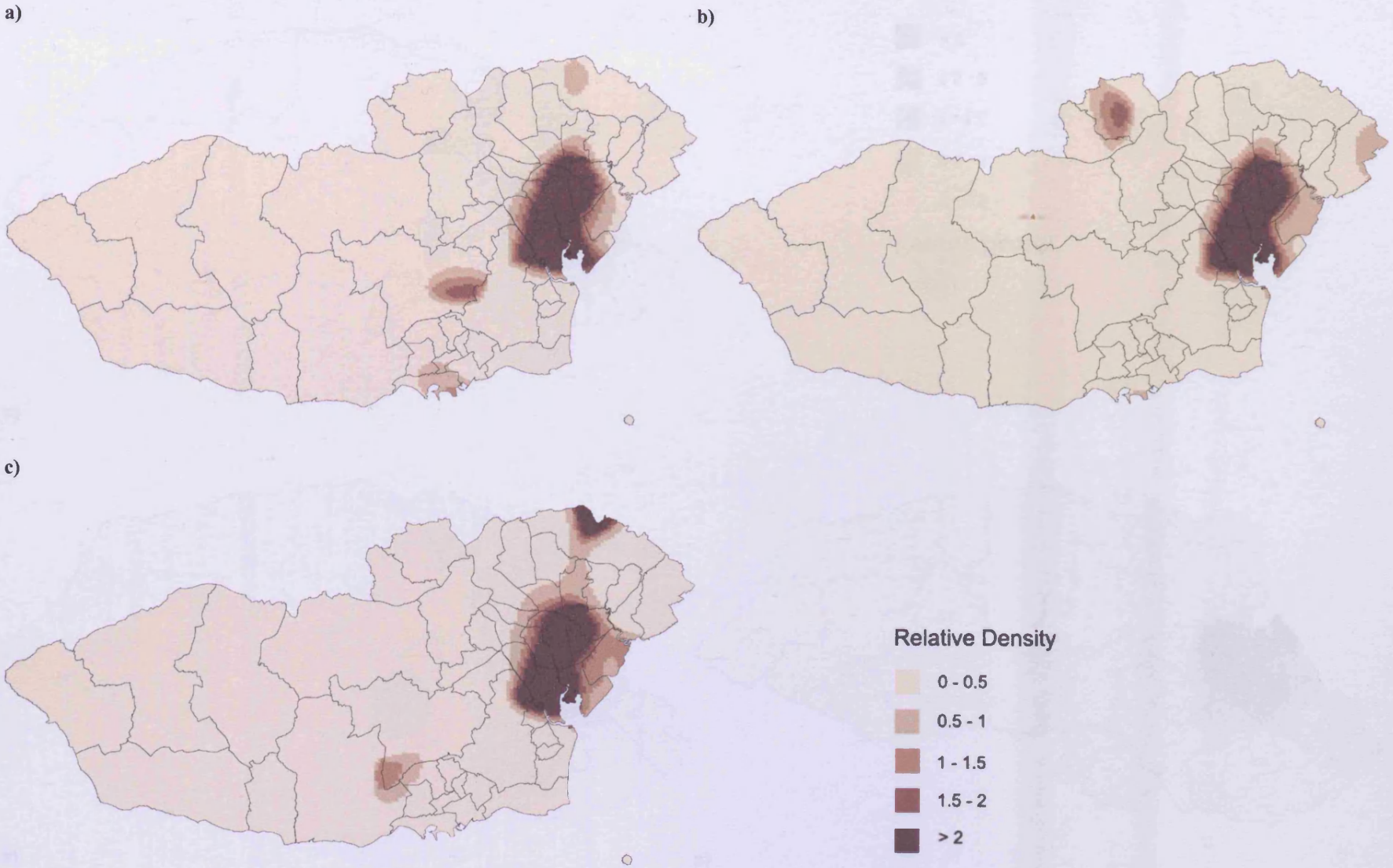


Figure A5.17: relative density contours of births to Bangladeshi mothers in Cardiff and the Vale of Glamorgan a) 1990-1994; b) 1995-1999; c) 2000-2004

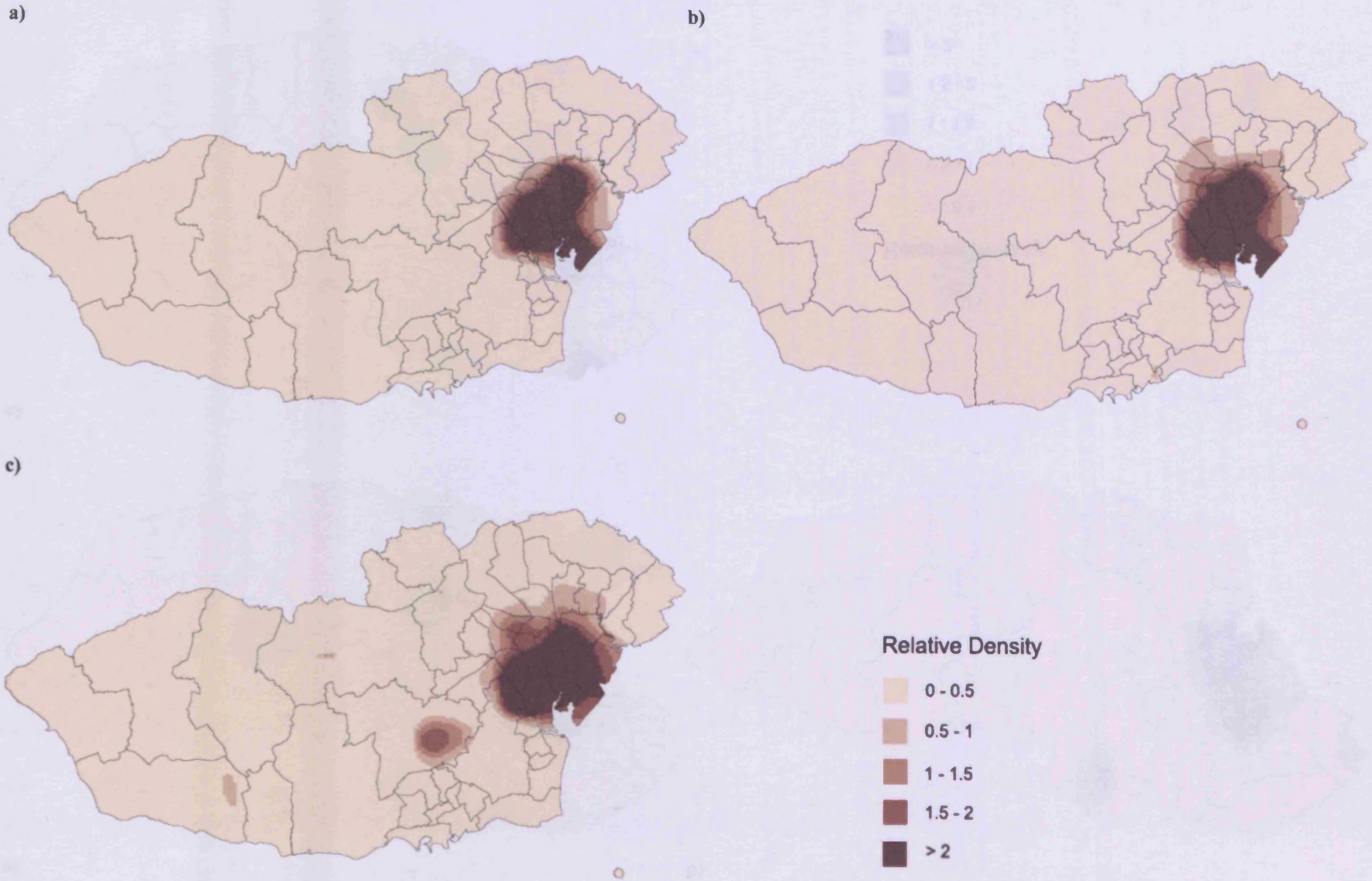
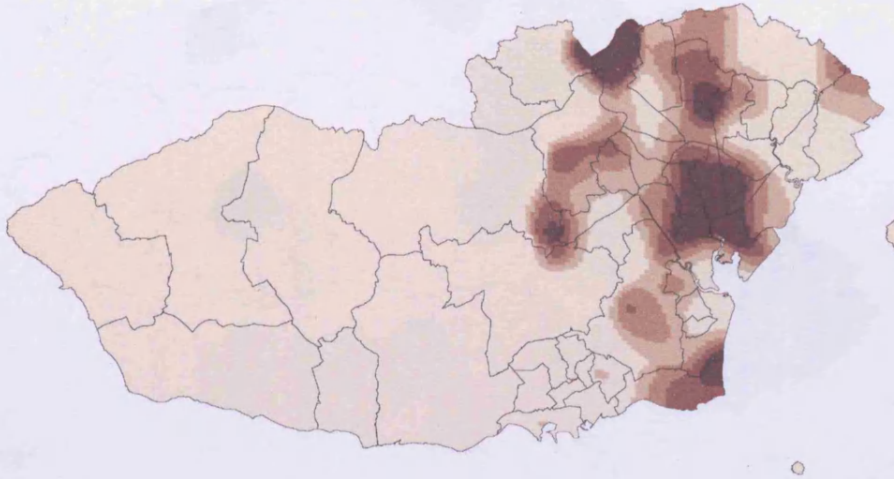


Figure A5.18: relative density contours of births to Chinese mothers in Cardiff and the Vale of Glamorgan a) 1990-1994; b) 1995-1999; c) 2000-2004

a)



b)

c)

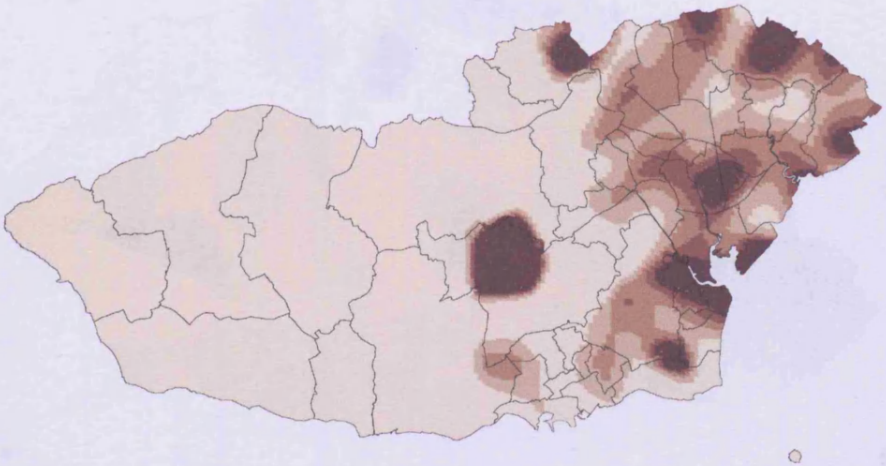


Figure A5.19: relative density contours of births to underweight mothers in Cardiff and the Vale of Glamorgan a) 1990-1994; b) 1995-1999; c) 2000-2004

a)

b)

c)

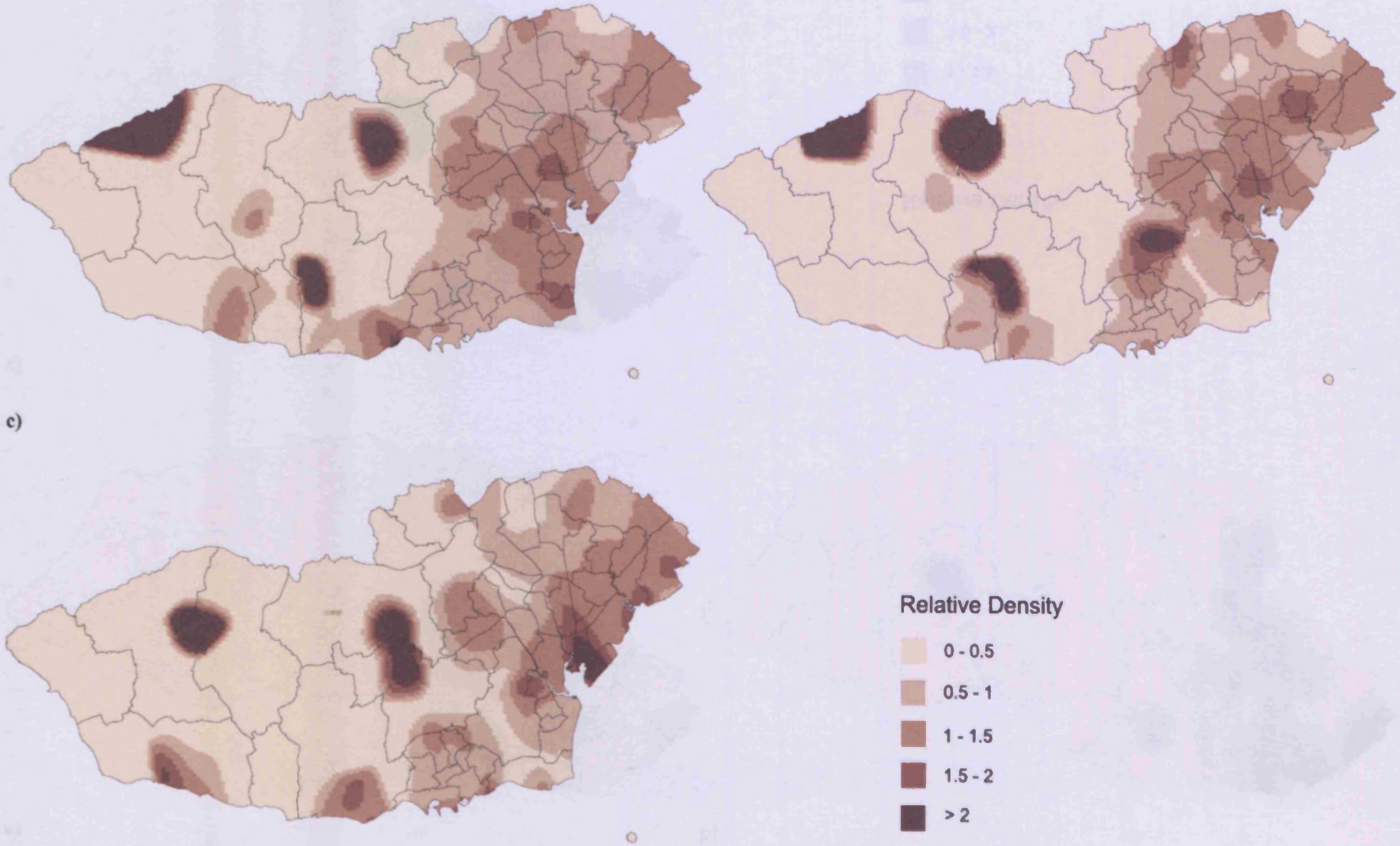


Figure A5.20: relative density contours of births to obese mothers in Cardiff and the Vale of Glamorgan a) 1990-1994; b) 1995-1999; c) 2000-2004

a)

b)

c)

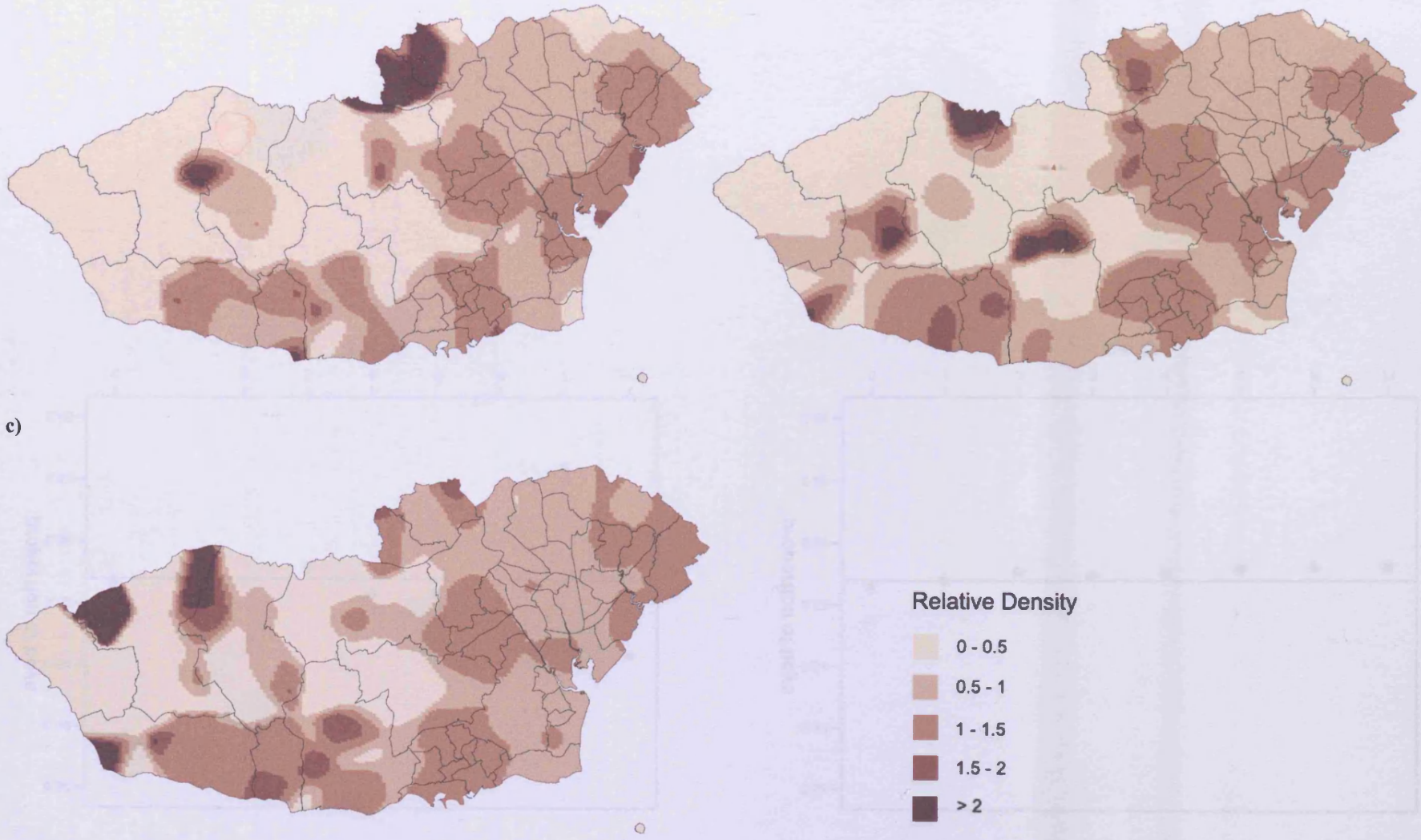


Figure A5.21: the sex ratio by parity in Wales 1983-1997

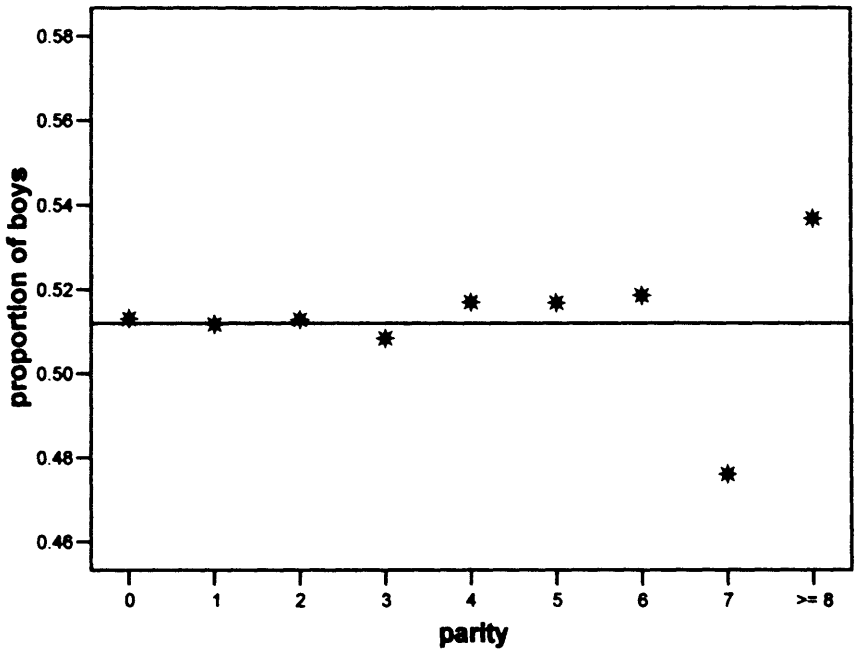


Figure A5.22: the sex ratio by parity in the USA 1943-2002

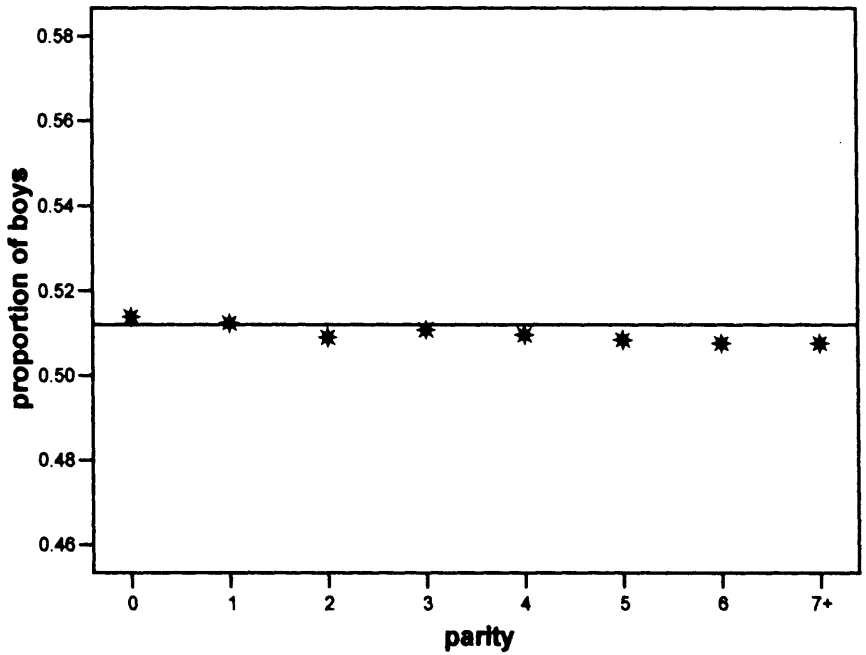


Figure A5.23: relative density contours of births with a parity of 5 or more in Cardiff and the Vale of Glamorgan a) 1990-1994; b) 1995-1999; c) 2000-2004

a)

b)

c)

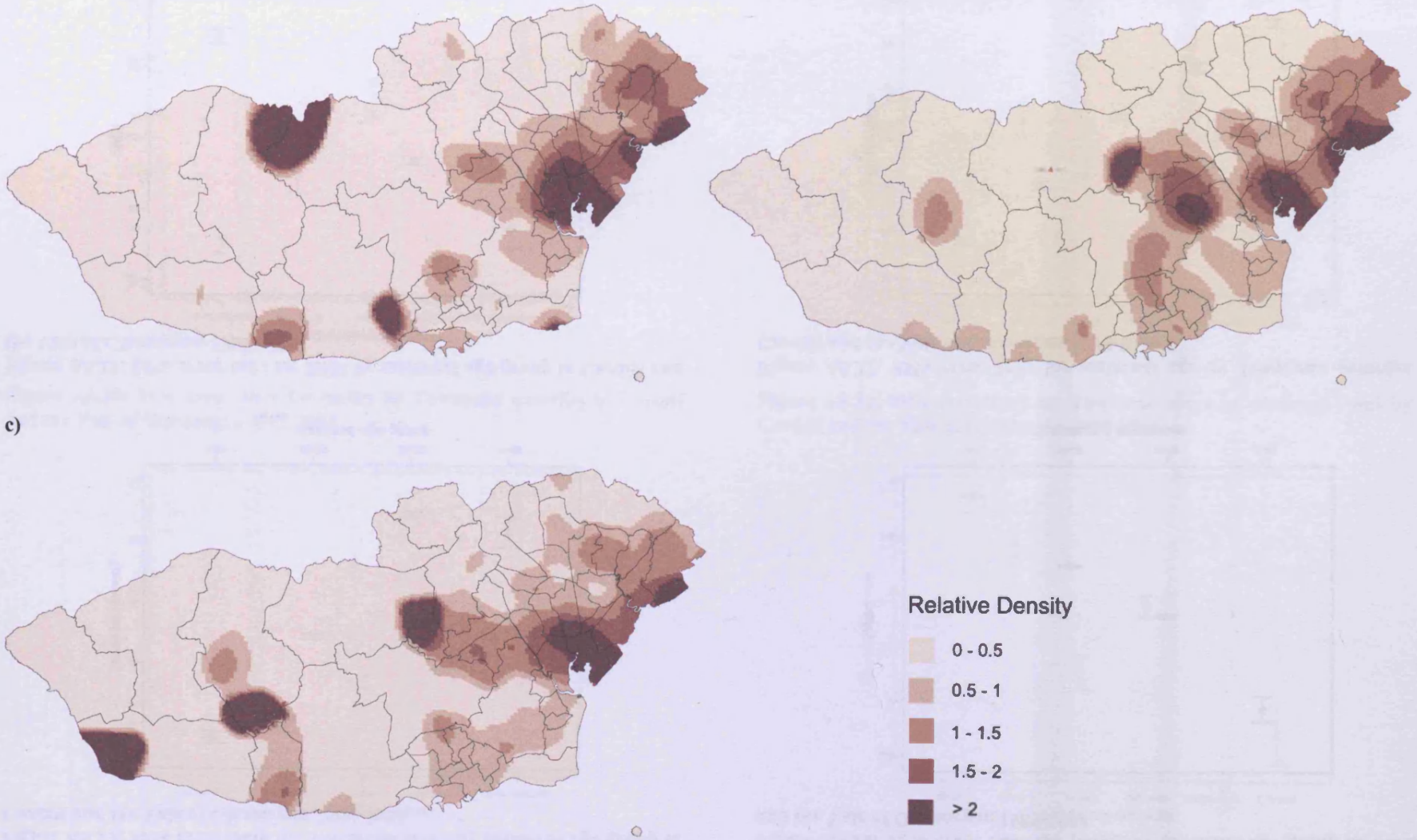


Figure A5.24: 95% error bars for Townsend score by maternal age group in Cardiff and the Vale of Glamorgan 1989-2004

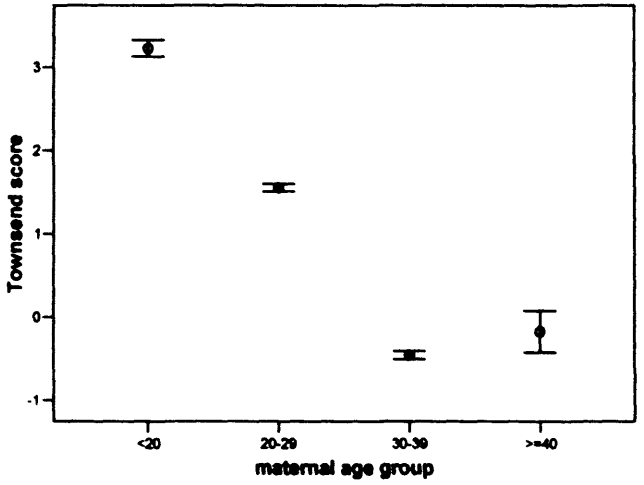


Figure A5.25: 95% error bars for BMI by maternal age group in Cardiff and the Vale of Glamorgan 1989-2004

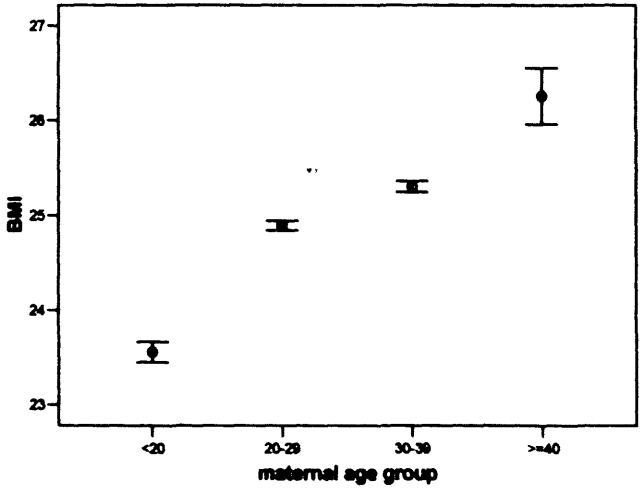


Figure A5.26: 95% error bars for parity by maternal age group in Cardiff and the Vale of Glamorgan 1989-2004

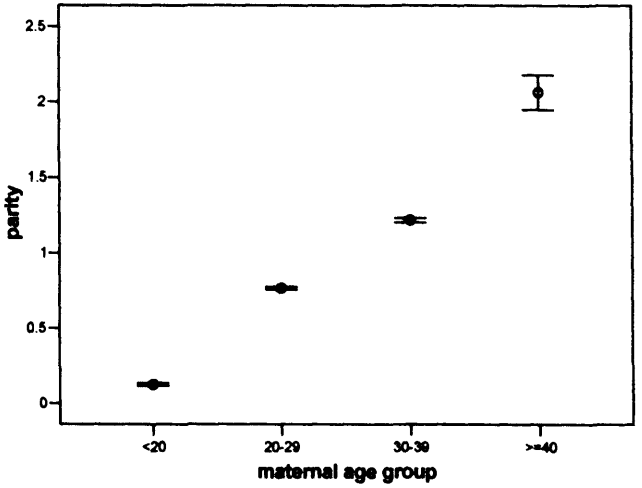


Figure A5.27: 95% error bars for maternal age by Townsend quintiles in Cardiff and the Vale of Glamorgan 1989-2004

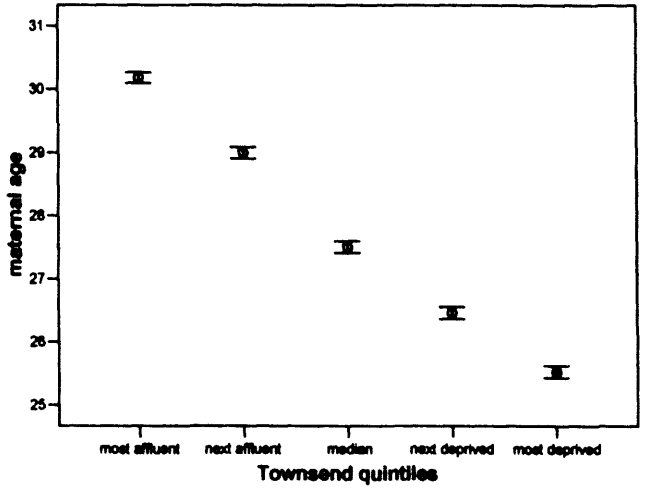


Figure A5.28: 95% error bars for BMI by Townsend quintiles in Cardiff and the Vale of Glamorgan 1989-2004

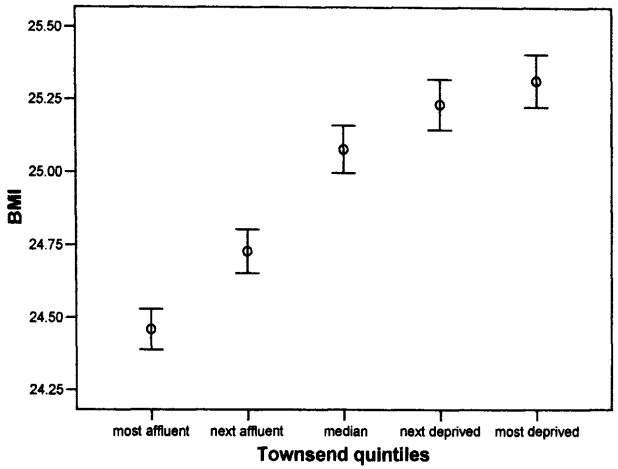


Figure A5.29: 95% error bars for parity by Townsend quintiles in Cardiff and the Vale of Glamorgan 1989-2004

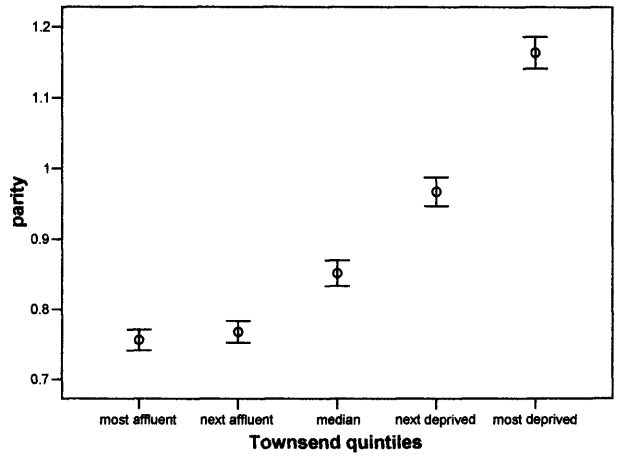


Figure A5.30: 95% error bars for maternal age by maternal ethnicity in Cardiff and the Vale of Glamorgan 1989-2004

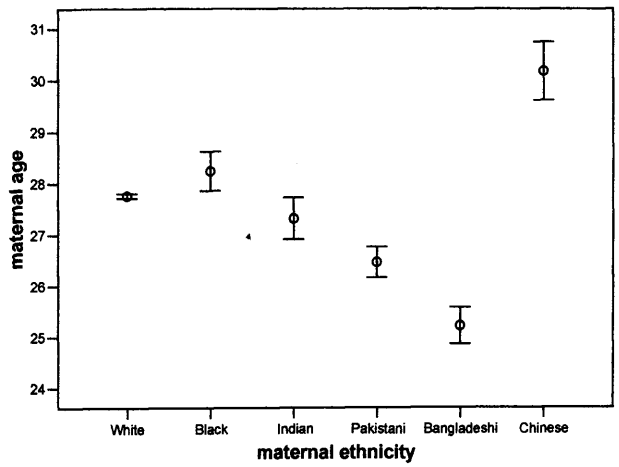


Figure A5.31: 95% error bars for Townsend score by maternal ethnicity in Cardiff and the Vale of Glamorgan 1989-2004

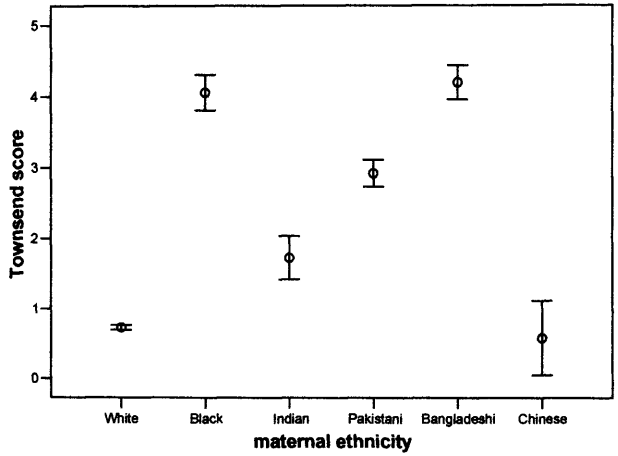


Figure A5.32: 95% error bars for BMI by maternal ethnicity in Cardiff and the Vale of Glamorgan 1989-2004

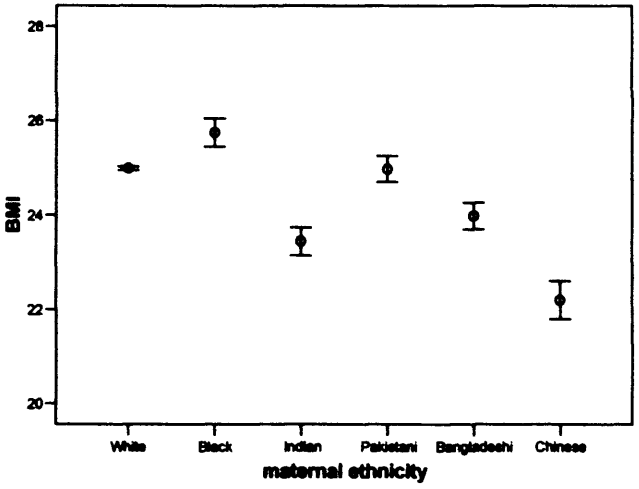


Figure A5.33: 95% error bars for parity by maternal ethnicity in Cardiff and the Vale of Glamorgan 1989-2004

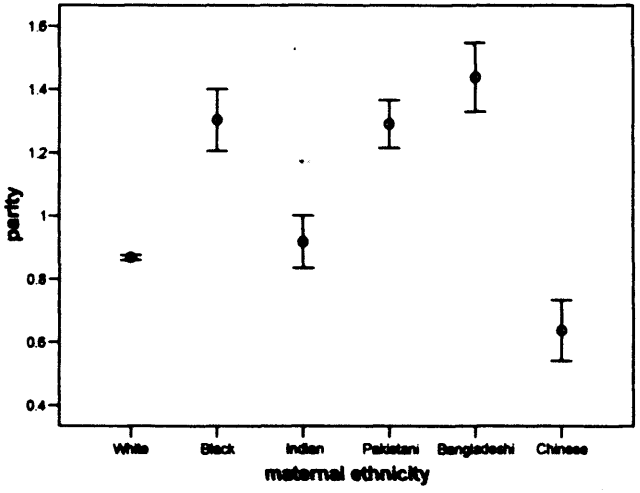


Figure A5.34: 95% error bars for maternal age by parity in Cardiff and the Vale of Glamorgan 1989-2004

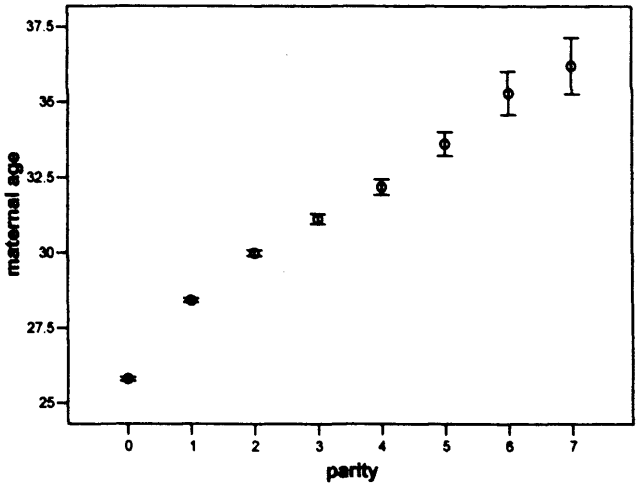


Figure A5.35: 95% error bars for Townsend score by parity in Cardiff and the Vale of Glamorgan 1989-2004