West, Joe (2005) Coeliac disease: studies of its frequency and consequence. PhD thesis, University of Nottingham.

Coeliac disease: studies of its frequency and consequence

Joe West

Thesis submitted to the University of Nottingham for the degree of Doctor of Philosophy, October 2004

Table of contents

Abstract

Background

The development of serological tests for the diagnosis of coeliac disease, including tests for endomysial and tissue transglutaminase antibodies, has made population screening for coeliac disease a realistic possibility. Several serological screening studies from European countries have shown that as many as 1% of the general population may have undetected coeliac disease. The implications of this diagnosis are unclear since the only data on the morbidity and physiological characteristics associated with previously undetected disease come from small, selected, case series. Most adult screening studies in the general population have identified only small numbers (i.e. less than 20 cases) of previously undetected cases and have therefore been unable to examine these issues through lack of statistical power.

Clinically diagnosed coeliac disease has traditionally been linked with a variety of adverse co-morbid conditions including osteoporosis, non-Hodgkin's lymphoma and an increased mortality in general. These conditions are thought to be partly a consequence of the altered nutritional status associated with the malabsorption that occurs with villous atrophy of the small bowel in coeliac disease. Although some of the adverse effects of, for example, vitamin and calcium deficiencies in coeliac disease have previously been explored whether there may be potentially beneficial effects of mild malabsorption have not.

There are two main aspects in this thesis. The first is to estimate the prevalence of undetected coeliac disease in England and explore the important physiologic correlates of this condition. The second is to examine the risk of fracture, vascular disease, malignancy and mortality in people with diagnosed coeliac disease compared to the general population.

Objectives

1. To estimate the seroprevalence of undetected coeliac disease in England.
2. To explore the relationship between undetected coeliac disease and various socio-demographic characteristics and physiological measures.
3. To quantify the impact of diagnosed coeliac disease (compared to the general population) on the risk of:
    a. Fracture
    b. Vascular disease (hypertension, high cholesterol, atrial fibrillation, myocardial infarction and stroke)
    c. Malignancy and mortality

Methods

To examine objectives 1 and 2 I utilised the Cambridge General Practice Health Study. This study identified individuals aged 45-76 registered with 12 general practices and invited them to complete a health survey, have a bone density measurement and submit a blood sample between 1990 and 1995. Serum samples from 7550 participants were tested for antiendomysial antibody (EMA). Seroprevalence of undetected coeliac disease was defined by EMA positivity. Differences between EMA positive and negative participants of various physiological measures and reported characteristics were estimated using multivariate logistic and linear regression and adjusted for age, gender, social class and smoking behaviour.

To examine objective 3 I performed a population based cohort study using the General Practice Research Database to quantify the risk of fracture, vascular disease, malignancy and mortality in people with coeliac disease compared to the general population. I identified 4732 people with coeliac disease and 23620 age and sex matched control subjects. I used Cox regression to estimate hazard ratios for fracture, myocardial infarction, stroke, malignancy and mortality, and conditional logistic regression to estimate the risk of diagnosed hypertension, hypercholesterolaemia and

atrial fibrillation, in people with coeliac disease compared to the general population.

Findings

The studies show that undetected coeliac disease is likely to affect about 1% of the population of England aged 45-76, a figure similar to several other countries. Those affected more commonly reported “good or excellent health”, however they do have an increased risk of osteoporosis and mild anaemia. In contrast they have a favourable cardiovascular risk profile including lower serum cholesterol and blood pressure.

In people with clinically diagnosed coeliac disease, compared to the general population, there were small increases in both the absolute and relative overall fracture incidence with a 2-fold increase in the risk of hip fracture. Adults with treated coeliac disease did have a favourable vascular disease risk factor profile but numbers having heart attacks or strokes were modest and rates of heart attack and stroke were not reduced. There were modest increases in the overall risks of malignancy and mortality in people with coeliac disease and most of this excess risk occurred in the first year of follow up after diagnosis, suggesting ascertainment bias. I found a marked reduction in the risk of breast and lung cancer in people with coeliac disease and the mechanism of this merits further attention as it may provide insight into the aetiology of these common malignancies.

Conclusions

I found that approximately 1% of general adult population of the UK has undetected coeliac disease. The findings suggest that although coeliac disease is associated with some adverse conditions; it may also have some beneficial health effects.

Contributions

The studies in this thesis made use of two existing datasets – The Cambridge General Practice Health Study and the General Practice Research Database.

For the study using the Cambridge Study I utilised the data collected originally at entry to the study. In addition I organised the logistics of transporting and testing the blood samples, managing the data collection, linking the blood sample data to the originally collected information and for the data I used checking its accuracy. For this process I had help from Rosemary Reader (Research Assistant, Cambridge University). The blood samples were tested in the laboratories of the Derbyshire Acute Hospitals NHS Trust. I then carried out the entire analysis myself with some additional advice from my supervisors and, for the statistical analysis, Sarah Lewis (Senior Lecturer in Medical Statistics, University of Nottingham).

For the GPRD studies I had the original ideas, wrote the grant that subsequently funded the studies and carried out the majority of the data management and the entirety of the statistical analysis myself. For the data management I had help and advice from Chris Smith (Research Assistant, University of Nottingham). For the study design and analysis I had advice and suggestions from my two supervisors and Dr Tim Card (Wellcome Research Training Fellow in Clinical Epidemiology).

Acknowledgements

I am very grateful for the supervision of Richard Logan and Richard Hubbard in the Division of Epidemiology and Public Health, University of Nottingham that helped me to complete this thesis.

Tim Card deserves particular thanks for his remarkable tolerance in dealing with my conversation (with him and others), cake obsession, dislike of code lists and general, all-round, Joe-ness. All of this work would have been of much poorer quality without his input. I hope to continue learning from and with him for many years hereafter.

Some of the many people who also made the work possible and kept it in perspective are listed below. Thank you.

Geoff Holmes, Peter Hill, Alison Lloyd, Rosemary Reader, Kay-Tee Khaw, Chris Smith, Masoud Solaymani-Dodaran, Sarah Lewis, Laila Tata, Gill Price, Dave Macafee, Dave Williams, Kathleen Eley, Lesley Bird, Sylvia Attah.

I would not have begun nor finished this work without the continuing support, inspiration and encouragement of my family, Denise Kendrick and Liz Doney. You are all brilliant.

Dedication

This thesis is dedicated to the memory of my dad who died whilst it was being finished. Below is some text taken from the tribute I made to him at his funeral.

"…his desire to learn about people, personalities and places has to some extent been passed on. Only now he is not here do I realise how much I have learnt, was learning and will continue to learn from him. I have spent my whole life trying to learn new things and I hope I can honour his memory by continuing to do so. Even in the last 6 months of his life, when he was so desperately ill, he was teaching me. Not least during this time I feel I have learnt so much about myself both emotionally and professionally as a doctor, through his analysis of his illness, dignity in dealing with its effects and tenderness towards us all whilst we grieved with him. I miss him."

List of figures

List of tables

Abbreviations

General Practice Research Database	GPRD
Body mass index	BMI
Immunoglobulin A	IgA
Tissue transglutaminase	tTG
Antitissue transglutaminase antibody	tTGA
Antiendomysial antibody	EMA
Value Added Medical Products	VAMP
Oxford Medical Information Systems	OXMIS
Confidence interval	CI
Hazard Ratio	HR
Inflammatory bowel disease	IBD
Enzyme linked immunoabsorbent assay	ELISA
Major Histocompatibility Complex	MHC
Antigliadin antibodies	AGA
Human Leukocyte Antigen	HLA

1 Introduction

The aim of the first study in this thesis is to estimate the prevalence of previously undetected coeliac disease in England and explore the potential adverse and beneficial physiologic correlates and socio-demographic characteristics of those with evidence of the disease compared to those without. The aim of the subsequent studies is to examine the risk of fracture, vascular disease, malignancy and mortality in people with clinically diagnosed coeliac disease compared to the general population.

To understand the rationale for these studies this introduction contains a brief description of how coeliac disease is defined and diagnosed, and the clinical manifestation of the disease. This section will also describe what is already known and not known about its occurrence and the impact of both previously undetected and clinically diagnosed disease. Finally details of the objectives of the thesis are given.

1.1 What is coeliac disease?

1.1.1 Historical perspective

Samuel Gee is generally credited with the first accurate clinical description of coeliac disease when he gave a lecture 'On the coeliac affection' at The Hospital for Sick Children, Great Ormond Street, London, on the 5th of October 1887. This report was subsequently published in the St Bartholomew's Hospital Reports[1]. His summarised description is as follows:

"There is a kind of chronic indigestion which is met with persons of all ages, yet is especially apt to affect children between one and five years old. Signs of the disease are yielded by the faeces; being loose, not formed but not watery; more bulky than the food taken would seem to account for; pale in colour, as if devoid of bile; yeasty; frothy, an appearance probably due to fermentation; stinking, stench often very great, the food having undergone putrefaction rather than concoction….. The onset is usually gradual, so that time is hard to fix: sometimes the complaint sets in suddenly, like an

accidental diarrhoea; but even when this is so, the nature of the disease soon shows itself….. The course of the disease is always slow, whatever be its end; whether the patient live or die, he lingers ill for months or years…. But if the patient can be cured at all, it must be by means of diet."

During the 1920's attempts were made to "cure" coeliac disease with various restrictive diets for example the "banana diet" described by Hass[2]. However it was nearly forty years after Gee's death, in the late 1940's, that his suspicions were finally confirmed.

In the 1940's Williem Karel Dicke, a Dutch physician, noticed that during the second world war, children with coeliac disease in Holland thrived *when others were starving*, only to subsequently relapse when flour was airlifted into the Netherlands at the end of the war. Furthermore he demonstrated that malabsorption in coeliac disease was

"elicited or aggravated by certain types of flour, especially wheat and rye flours"

He came to this conclusion from the clinical observation that there was variation in the well being of children at different times during their stay in hospital[3]. Dicke correlated the mood alterations with variations in the stool weight and frequency. The diet of these children consisted of 'gruel'. This was a 'porridge like' substance which was common in the Dutch diet at the time. Dicke discovered that the constituents of the 'gruel' varied depending on the availability of wheat flour. He observed that paediatric patients given gruel with a rice or potato flour base, appeared to be far 'happier' than when eating the 'gruel' based on wheat.

During the 1940's and 1950's work was carried out to attempt to identify the pathological lesion in coeliac disease. Work looking at autopsy specimens was inconclusive due to post mortem autolysis but suggested small intestinal abnormalities[4, 5]. Himes and Adlersberg had noted similar changes in

jejunal biopsies in life as those observed in post-mortem specimens from four of their patients[6]. Paulley also noted the same changes in jejunal mucosal specimens taken at laparotomy[7]. By the mid 1950's it was well established that in coeliac disease there was a characteristic change in the jejunal mucosa of villous atrophy.

In 1953, J. H van de Kamer identified gliadin as the toxic factor in wheat[8] and that this protein was present in wheat, barley and rye. In 1961 Taylor discovered circulating antibodies against gliadin in the sera of patients with coeliac disease[9]. In combination with advances in genetics these two findings paved the way for the modern understanding of the pathogenesis of coeliac disease.

1.1.2 Pathogenesis of coeliac disease

The traditional view of the aetio-pathogenesis of coeliac disease is that it occurs in people who are genetically susceptible and are subsequently exposed to gluten. Gluten is a heterogeneous mixture of proteins termed gliadins and glutenins and exposure to gluten is considered to be a necessary factor in the development of coeliac disease. The mechanisms for the pathogenesis are discussed below.

1.1.2.1 Genetics

The familial aggregation of coeliac disease shows the importance of genetic predisposition to the disease. There have been various twin studies that have estimated monozygotic twin concordance for coeliac disease at about 70%[10]. Some of that concordance has been explained by Human Leukocyte Antigen (HLA) genetic studies.

Over 97% of people with coeliac disease express the *HLA-DQ2* or *HLA-DQ8* genes, coding for a class II Major Histocompatibility Complex (MHC) molecule comprising a DQA1*0501 chain with a DQB1*0201 or DRB1*0301 chain, or DQA1*0301 with a DRB1*0401 or DQB1*0302 chain[11, 12]. As the function of class II MHC molecules is to present short peptide-antigens (epitopes) to CD4+ T lymphocytes, this is one of the important pieces of

evidence in favour of the view that coeliac disease is an aberrant immune response to gluten.

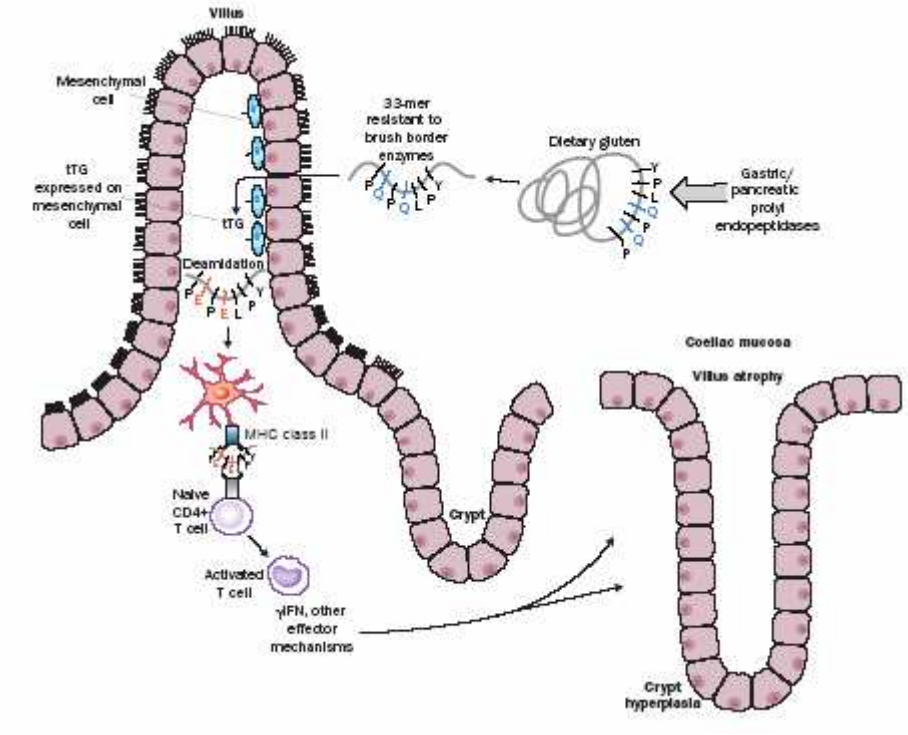
Further support for this mechanism is that α-gliadin-specific CD4+ T cells that produce interferon γ can be isolated from the intestinal mucosa of people with untreated coeliac disease. Such T cells are not present in normal intestine, or in people with coeliac disease on a gluten free diet[13]. In 2000 two independent studies identified immunodominant epitope peptides from the 57-75 region of α gliadin, which is considered the most toxic component of gliadin[14, 15]. These experiments, with the addition of a later report[16], indicated that a large proportion of CD4+ T cells from people with coeliac disease recognised three overlapping peptides rich in proline and glutamine (PFPQPQLPY, PQPQLPYPQ, and PYPQPQLPY) which derive from a region of α gliadin known to be recognised by antibodies found in people with coeliac disease.

1.1.2.2 Tissue transglutaminase

Tissue transglutaminase (tTG) is an ubiquitous enzyme found in all organs, including the intestine. It is known to be released upon cellular damage and to crosslink proteins in order to control tissue damage. This crosslinking occurs by forming a covalent bond between the sidechain of a glutamine in one protein with the aminogroup in the sidechain of a lysine in the other protein[17]. In 1997 Dieterich et al identified tTG as the autoantigen in coeliac disease[18] and, subsequently, other reports found that increased levels of tTGA appear to be very specific indicators of the presence of coeliac disease[19-21]. Each of the three peptide epitopes described above contains at least one glutamine residue which is a substrate for the deamidase activity of tTG. Deamidation of these residues is essential for significant T-cell stimulation, because deamidation exposes negatively charged aminoacids that are an essential part of the structural motif involved in binding to the HLA-DQ2 molecule[14-16]. Through these in vitro studies the concept that tTG plays a critical role in the pathogenesis of coeliac disease by generating the antigenic epitopes present in α gliadin has emerged.

Recent work by Shan et al has shown that digestion of recombinant α2 gliadin with gastric and pancreatic enzymes in vitro produces a highly stable 33-mer peptide that is rich in proline and glutamine and contains all three of the previously described epitopes[22]. This 33-mer peptide is resistant to in-vitro digestion with preparations of brush-border enzymes derived from the small intestine of rats or human beings. In addition this 33-mer peptide has a selective ability to survive digestion in vivo and may be present in significant amounts in the normal small intestine. Shan et al also demonstrated that this product has high specificity for deamidation by tTG and the resulting products are extremely stimulatory for all the *HLA-DQ2* restricted α gliadin specific T cells they examined. The hypothetical scheme proposed by this work is displayed in Figure 1-1.

Figure 1-1. Hypothetical scheme for interaction between intestinal processing proteins and the specific immune system in coeliac disease



Lancet 2003; **361**: 1290-1292

Although these findings are supportive of only a few peptides dependent on tTG dominating the T cell response in coeliac disease work from Koning's group is suggestive that epitopes that do not come from this 33-mer peptide can also stimulate T cells in people with coeliac disease[23]. This perhaps explains why although highly specific antibody tests for tTG are excellent predictors of coeliac disease there remains the occasional discrepancy with the finding of negative antibodies yet an abnormal mucosal lesion.

1.1.2.3 Enteropathy

Marsh classified the small intestinal lesion in people with coeliac disease in the early 1990's[24]. He classified various stages of the abnormalities of the small intestinal mucosa, comprising infiltration of the epithelium with lymphocytes, hyperplasia of crypts, progressive loss of surface epithelial cells and villous atrophy. In the Marsh I lesion (lymphocytic enteritis) the architecture of mucosa appears normal but the mucosal epithelium is invaded by lymphocytes. Marsh II (lymphocytic enteritis with crypt

hyperplasia), is characterised by intraepithelial lymphocytosis accompanied by hyperplasia of the crypts. Marsh III (flat lesion) consists of intraepithelial lymphocytosis, crypt hyperplasia and villous atrophy. Marsh also described a type IV lesion (irreversible hypoplastic/atrophic lesion) in which malignant (lymphomatous) transformation can develop. The observation of increased numbers of intraepithelial lymphocytes in the small bowel mucosa of people with coeliac disease has led to work on understanding their role in its pathogenesis.

1.1.2.4 Intraepithelial lymphocytes

That the increase in intraepithelial lymphocytes is important in the pathogenesis of coeliac disease was first recognised by Ferguson et al[25]. In particular it has now been recognised that the proportion of γ/δ intraepithelial T lymphocytes are increased in people with coeliac disease. These primitive lymphocytes recognise bacterial nonpeptide antigens and unprocessed stress-related proteins. Two important stress-induced proteins that are increased on intestinal epithelial cells by interferon gamma are MICA and MICB, which resemble major histocompatibility class I genes[26]. MICA and MICB gene expression is regulated by promotor heat-shock elements similar to heat-shock protein 70[26]. The receptor (NKG2D) for MICA on natural killer and γ/δ T cells has recently been identified[27]. Once activated γ/δ T cells secrete chemokines that attract and stimulate cells of the unspecific (innate) immune response (monocytes/macrophages, neutrophils, and eosinophils). However, they modulate the antigen-specific immune response by secreting IL-4, which dampens Th1 activity in favour of Th2 reactivity. Therefore γ/δ T cells appear to protect the intestinal mucosa from chronic exposure to damaging agents such as dietary gluten in gluten-intolerant individuals[17].

1.1.3 Aetiology of coeliac disease

As the concordance between monozygotic twins is at most 70% this underlines the importance of environmental factors, other than exposure to gluten, in the development of coeliac disease[10]. I have discussed below the areas that there is some information available.

1.1.3.1 Infant feeding practices

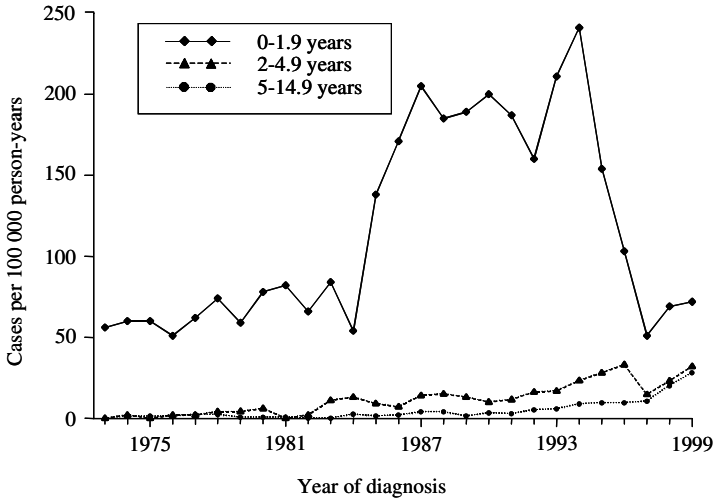
It was suggested as early as the 1950's that breast-fed infants have a later onset of coeliac disease[28]. In the 1970's there were reports from England, Ireland and Scotland that there appeared to be a decline in the incidence of childhood coeliac disease[29-32]. Changes in infant feeding practices that in Britain were promoted by new regulations were suggested as a possible explanation for the decline[33]. Those recommendations advocated breast-feeding for a minimum of two weeks and preferably for four to six months, use of infant formulas, avoidance of solids before the age of four months, and that cereals should not be added to the milk in bottle feeding. There was speculation that the age at onset in children was increasing yet during at the same time in Sweden there was no change in incidence with similar infant feeding recommendations[34].

Two large case-control studies carried out in Italy by Auricchio et al and by Greco et al examined infant feeding practices in relation to the risk of developing childhood coeliac disease[35, 36]. Auricchio et al examined feeding practices in 216 children with coeliac disease and compared them with those used for their siblings. Siblings eating gluten within the first 2 months of life had a slightly greater risk of developing coeliac disease than those who were started on gluten from age 3 months. Greco et al similarly looked at feeding practices in 201 children with coeliac disease and compared them with 1949 non-coeliac unrelated children. They found a two-fold increase in risk for developing coeliac disease when gluten had been introduced before the end of the second month of life although this was based on parents' recall so there might have been a reporting bias.

Recently Ivarrson et al investigated infant feeding practices and its impact on coeliac disease extensively with a series of population-based studies carried out in Sweden during the 1980's and 1990's. In an ecological analysis they examined national breast feeding practices and infant feeding with respect to the incidence of childhood coeliac disease[37]. This was in an era of high breast feeding and late gluten introduction in comparison with the UK. They

found a rise in incidence during the late 1980's and early 1990's, shown in Figure 1-2, that was preceded by an increase in the amount of gluten consumed. The national average daily consumption by children below the age of 2 years of flour from wheat, rye, barley and oats was doubled during the 1980's while from 1995 and onwards it declined by one-third.

Figure 1-2. Incidence of childhood coeliac disease in Sweden



Acta Paediatr 2000; **89**: 165-71

Complementing earlier studies this suggested a role for the timing of gluten introduction in the development of coeliac disease. In addition Ivarrson et al carried out the largest population based case-control study so far reported between 1992 and 1995 to examine early life risk factors for the development of coeliac disease[38]. They included 627 children with incident coeliac disease and 1254 age, sex and area of residence matched controls. They found that the risk for development of coeliac disease was reduced in children below the age of 2 if they were breast fed when gluten was introduced (adjusted Odds Ratio (OR) 0.59 95% CI 0.42-0.83) and this relationship was stronger for infants breast-fed beyond the introduction of gluten (OR 0.36 95% CI 0.26-0.51). They concluded that breast-feeding has an independent protective effect against the development of coeliac disease if ongoing when gluten-containing foods are introduced. The interpretation of these observations was that more gradual introduction of gluten containing foods, perhaps allowing “tolerance” to develop, influences the development of coeliac disease.

1.1.3.2 Cigarette smoking

Coeliac disease, like ulcerative colitis, appears to be associated with non-smoking although it is unclear whether this is a causal association[39-42].

Although most of these case-control studies show some inverse relation between current smoking and diagnosed coeliac disease the strength of the association has varied. This is probably due to the inconsistent reporting of smoking status among the control populations in comparison with the coeliac populations where the current smoking proportion was about 40%. In addition the studies have been small so some random variation must be expected. There have been no previous reports of "undiagnosed" coeliac disease and smoking status.

1.1.4 Clinical manifestation

The clinical manifestation of coeliac disease has its onset commonly in either early childhood, between 9 and 24 months, or in the third or fourth decade of life[43-47]. In contrast to children where the sex ratio is 1, in general twice as many women are diagnosed as men as adults. In severe disease a "classical" syndrome of gastrointestinal malabsorption can occur characterised by diarrhoea (due to steatorrhoea), weight loss and fatigue. However, the majority of people diagnosed with coeliac disease nowadays have a milder constellation of symptoms that include those already mentioned but may also include a variety of others such as abdominal discomfort or bloating, indigestion or non-gastrointestinal symptoms[43-47]. Since coeliac disease was first described the clinical manifestation appears to be changing, with increasing numbers being diagnosed as a result of the investigation of anaemia and/or non "classical" symptoms[48-52].

Logan et al described the clinical features in adult people with coeliac disease diagnosed in the Edinburgh and the Lothian areas of Scotland[50]. They compared features at presentation through the four quinquennia spanning 1960-1979 and found a lower age at diagnosis in women in the later years, 63% presenting with "typical malabsorption syndrome" in the 1960's compared with 21% in the 1975-79 and fewer haematological and biochemical abnormalities in the later periods. Other studies of comparable design have found similar patterns in adults diagnosed with coeliac disease[48, 49, 51, 52]. In addition Hin et al recently carried out a case finding study in primary care in the UK[53]. They tested people in primary

care who had irritable bowel syndrome, anaemia, malabsorption symptoms and people with fatigue. They found that people with fatigue and/or having a past or present diagnosis of microcytic anaemia had the highest prevalence of previously undetected coeliac disease.

Other haematological features of coeliac disease that have been reported repeatedly include macrocytic anaemia, hypoproteinaemia, folate deficiency, hypocalcaemia and abnormal liver function tests particularly hypertransaminasaemia[50, 54-60].

The data supporting the thought that the clinical manifestation in children with coeliac disease is also changing with a greater proportion are presenting with a milder constellation of symptoms similar to the change seen in adults is less consistent. In a study of the incidence and clinical presentation of childhood coeliac disease in the Netherlands George et al reported a decrease in the proportions of children with clinical growth failure in height and weight at diagnosis between 1975 and 1994. In contrast Greco et al described features of newly diagnosed cases in their study the Naples area, Italy between 1973 and 1986[61], and found little change across time. Most other studies have been too small to make sensible comparisons across time but the consensus appears to be that childhood coeliac disease diagnosed more recently may be presenting with a less severe clinical illness[62-65].

One explanation for any changes in presentation could be that the natural history of the disease is changing, perhaps in response to changing environmental stimuli like infant feeding practices in children or cigarette smoking in adults. A more likely explanation is that the ability to make the diagnosis has improved in both quality and accessibility throughout the last 20 years with the development of both accurate serological markers of the disease and increasing use of endoscopic biopsy techniques rather than the traditional Crosby capsule biopsy. Therefore a broader spectrum of people are being investigated for coeliac disease and consequently being diagnosed.

1.1.5 Serological markers

In the 1980's Chorzelski et al described the production of anti-endomysial antibodies in people with dermatitis herpetiformis and coeliac disease[66]. Endomysium is a connective tissue protein found in the collagenous matrix of human and monkey oesophagus. Antibodies to endomysium can be measured in the serum with the use of indirect immunoflorescence[66]. The autoantigen recognised by endomysial antibody is tTG which is clearly important in the pathogenesis of coeliac disease as has been described in section 1.1.2.2. The Immunoglobulin A antiendomysial antibody (EMA) test can use either monkey oesophagus or human umbilical cord as substrate and its diagnostic utility has been shown to be very good, with specificity estimated at 99% and sensitivity over 90%[67]. In some laboratories specificity is as high as 99.8% (personal communication, Peter Hill). However the test is labour intensive and qualitative so requires money, time and expertise to perform. This has led to the development of enzyme linked immunoabsorbent assay (ELISA)-based tests for the measurement of IgA tissue transglutaminase antibody levels that are of comparable sensitivity and specificity to the EMA test[21, 68-70]. Measuring tTG antibody levels is quicker, easier and quantitative, so has clear advantages over the EMA test[44]. Both EMA and tTG have superseded the use of antigliadin antibodies (AGA) which having been identified as a useful serological marker in coeliac disease[71] have subsequently been shown to have inferior diagnostic accuracy[72] with a sensitivity as low as 76%.

Coeliac disease has an association with IgA deficiency so it is possible to incorrectly label a person as not having coeliac disease in such cases particularly when using the IgA dependent immunoflorescence EMA test[73]. Recently the improvements in techniques for measuring tTGA have allowed measurement of low serum IgA using recombinant human tTGA ELISAs[74].

1.1.6 Definition of coeliac disease in this thesis

The current diagnostic criteria for coeliac disease are largely based on the revised guidelines published by the European Society for Paediatric Gastroenterology, Hepatology and Nutrition in 1990 that recommended

reliance on the finding of a structurally abnormal small intestinal intestinal mucosa as described in section 1.1.2.3, followed by a clear clinical or histological remission on a gluten free diet[75]. Serological markers are used to either add additional evidence in favour of the diagnosis and/or to initially identify people who may have the condition.

The definition of coeliac disease in this thesis varies according to epidemiological need. The underlying assumption I have used is that coeliac disease is a life-long condition consequent upon genetic predisposition and exposure to gluten in the diet. In the study of prevalence of undetected coeliac disease I rely solely on serological evidence as a marker of "undetected coeliac disease", whereas in the studies of health impact of clinically diagnosed coeliac disease I rely on the premise that those people who have such a diagnosis recorded by their general practitioner were given it correctly. The merits and limitations of both definitions are discussed in detail with respect to the findings of the studies.

1.1.7 Occurrence

1.1.7.1 Prevalence

1.1.7.1.1 Undetected coeliac disease

Several serological screening studies from Europe, South America, Australasia and the United States of America have shown that approximately 0.5-1% of these populations may have undetected coeliac disease[76-84]. There is however some variation in the prevalence estimates in screening studies ranging from about 1 in 20 to 1 in 700. To discuss the methodology and findings I have split the studies into three groups: adult blood donor studies (Table 1-1), general population studies in children (Table 1-2) and general population studies in adults (Table 1-3). From the available data I have calculated the proportion of cases identified and the 95% Confidence Intervals (95% CI) using the binomial distribution. To assess whether the estimates vary by chance I carried out χ^2 tests on the data in each of the tables.

Adult blood donor studies

In all of the studies in Table 1-1 specimens for serological analysis were acquired from blood transfusion services in the country of origin. In three of the five their screening strategy used antiendomysial antibodies (EMA) as their test of prevalence[85-87], whereas in the remaining two studies antigiadin antibodies (AGA) were used first with EMA used in those who were positive.

Four of the five studies reported the age and sex distribution of their sample population. The mean age of participants varied from 31 to 41 years and in these studies the proportion of male participants ranged from 52% to 88%. Apart from the American study all the rest confirmed coeliac disease by assessing the small intestinal mucosa for histological abnormalities consistent with the disease. As can be seen in Table 1-1 the estimates of prevalence show little variation. When tested for variation using a χ^2 test there was no greater variation than expected by chance (Pearson χ^2=2.1 p=0.7).

I believe that this is unsurprising as the study methodologies were similar and the study populations are likely to be not too heterogeneous being all blood donors. These estimates show in general lower prevalence than the general population estimates shown in Table 1-2 and Table 1-3. This is also unsurprising as most transfusion services have some form of haemoglobin related entry criteria which are likely to have excluded some people with coeliac disease on the grounds of anaemia[50].

Table 1-1. Estimates of prevalence of coeliac disease from adult blood donor screening studies

Area	Mean or median* age	Proportion of males	Cases	Number screened	Prevalence	Proportion identified	95% CI of proportion	Year published	Reference
Sweden	41	65%	4	1970	1 in 492	0.20	0.06-0.52	1999	[88]
Italy	35	75%	10	4000	1 in 400	0.25	0.12-0.46	1999	[85]
Holland	Not reported	Not reported	3	1000	1 in 330	0.30	0.06-0.87	1999	[86]
Brazil	33*	87%	3	2045	1 in 681	0.15	0.03-0.43	2000	[89]
USA	39	52%	8	2000	1 in 250	0.40	0.17-0.79	2001	[87]

General population based studies in children

In the five studies in Table 1-2 general population based samples of children were recruited for screening for coeliac disease. In the Italian school district of Pesaro-Urbino Catassi et al recruited 66% of eligible school children aged 11-15 years between 1992 and 1993[90]. They used a negative AGA test to rule out coeliac disease and for all those positive tested them for EMA and, where they agreed, small bowel biopsy. Mean age was 12.8 years and 49% were male. In Sardinia, Italy Meloni et al carried out a similar study of 1607 school children during 1993-1994[91]. They did not report the overall number eligible for recruitment but the age range of participants was 6-14 years and 53% were male. In this study they used AGA to rule out coeliac disease but in addition tested 53 AGA negative participants for EMA as a control. Of these 53, one was positive but subsequently refused small intestinal biopsy. Those positive for AGA were offered small intestinal biopsy. Overall they found a prevalence of 1 in 93 children with previously undetected coeliac disease.

This latter finding has been replicated in two further large studies, one from the UK and one from Finland. The UK study used blood stored anonymously for the Avon Longitudinal Study of Parents and Children which is a population based birth cohort study established in 1990[92]. This study tested 5470 children aged 7.5 years for tTGA and then EMA with 54 found to be positive for the latter test. Maki et al tested 3654 students aged between 7 and 16 years in Finland. They tested all subjects for both tTGA and EMA, 50% of who were male, and offered small intestinal biopsy to those who were positive. Both these studies found a prevalence of about 1 in 100.

The study of Saharawi children carried out by Catassi et al in 1998 surprisingly showed the highest prevalence of all screening studies[93]. In this study 989 children (mean age 7.4 years, males 47%) of Saharawi descent who were now living in an Algeria province as refugees were tested for EMA and a sample of positive participants had their small intestinal

mucosa biopsied. They found that 5% of the children were positive for EMA. Reasons for this high prevalence, which is markedly different to all the other general population estimates, are unclear. The authors speculated that coeliac disease might confer some "protection" against intestinal infections or parasites.

If all studies in Table 1-2 are included in the χ^2 analysis then there is variation greater than expected by chance (Pearson χ^2=160.4 $p<0.01$) yet if the three most recent European studies are compared the findings did not differ.

Table 1-2. Estimates of prevalence of coeliac disease from general population based screening studies in children

Area	Mean age or age range*	Proportion of males	Cases	Number screened	Prevalence	Proportion identified	95% CI of proportion	Year published	Reference
Italy	13	50%	11	3351	1 in 305	0.33	0.16-0.59	1994	[90]
Italy	6-14*	53%	17	1607	1 in 93	1.06	0.62-1.69	1999	[91]
Sahara	7.4	53%	56	989	1 in 18	5.66	4.31-7.29	1999	[93]
Finland	7-16*	Not reported	37	3654	1 in 99	1.01	0.71-1.39	2003	[94]
Bristol, UK	7.5	Not reported	54	5470	1 in 100	0.99	0.74-1.29	2004	[92]

General population based studies in adults

Table 1-3 shows adult general population based screening studies for coeliac disease. Three of these studies used populations recruited for the World Health Organisation MONICA project (Monitoring of trends and determinants in Cardiovascular disease)[76, 77, 82]. Samples of each Country's general population were randomly selected from population registers stratified by age and sex. They were similar in design although one of the studies has been reported only as an abstract to date so the details available are not comprehensive. All three studies used a combination of AGA and EMA to identify people with previously undetected coeliac disease but not all had had small intestinal biopsy at the time of publication.

All the other studies except the Argentinean study used EMA as the main test for identifying people with undetected coeliac disease. In general the study populations were selected randomly with age and sex stratification yet the summary measures of the age distributions shown in Table 1-3 are suggestive of some variation. When I analysed the data in Table 1-3 using a χ^2 test there was greater variation in the estimates than expected by chance (Pearson χ^2 =16.4 $p<0.05$).

None of the studies provide age and sex standardised estimates of prevalence so it is difficult to assess whether differences in the estimates are due to differences in the demographics of the populations studied. There did not seem to be a consistent pattern either geographical or methodologically in the studies to explain this variation but putative reasons for the observed differences could be related to infant feeding practices, cigarette smoking, or other as yet unidentified important environmental factors.

Table 1-3. Estimates of prevalence of coeliac disease from population based screening studies in adults

Area	Mean age and/or age range*	Proportion of males	Cases	Number screened	Prevalence	Proportion identified	95% CI of proportion	Year published	Reference
Ireland	15-65	Not reported	15	1823	1 in 122	0.82	0.46-1.35	1997	[77]
Italy	44, 20-89	47%	4	2237	1 in 555	0.18	0.05-0.46	1997	[95]
Sweden	50, 25-74*	50%	10	1894	1 in 188	0.53	0.25-0.97	1999	[76]
Spain	45, 2-89*	45%	3	1170	1 in 389	0.26	0.05-0.75	2000	[79]
France	35-64*	Not reported	3	1163	1 in 388	0.26	0.05-0.75	2000	[82]
Italy	12-65*	48%	17	3483	1 in 204	0.49	0.28-0.78	2001	[80]
Argentina	Median 29, 16-79*	50%	12	2000	1 in 167	0.60	0.31-1.05	2001	[84]
Australia	20-79*	50%	7	3011	1 in 430	0.23	0.09-0.48	2001	[83]

1.1.7.1.2 Clinically diagnosed disease

Estimates of the prevalence of clinically diagnosed coeliac disease ranging from about 0.05% to 0.27% are available from several other studies[32, 96-98]. From data in Derby, UK the estimated prevalence of clinically diagnosed coeliac disease at the end of 1999 was 1 in 714 or 0.14% (unpublished data). These estimates probably vary due to differences in case ascertainment in the different areas reflecting some local interest in coeliac disease.

1.1.7.2 Incidence

1.1.7.2.1 Childhood coeliac disease

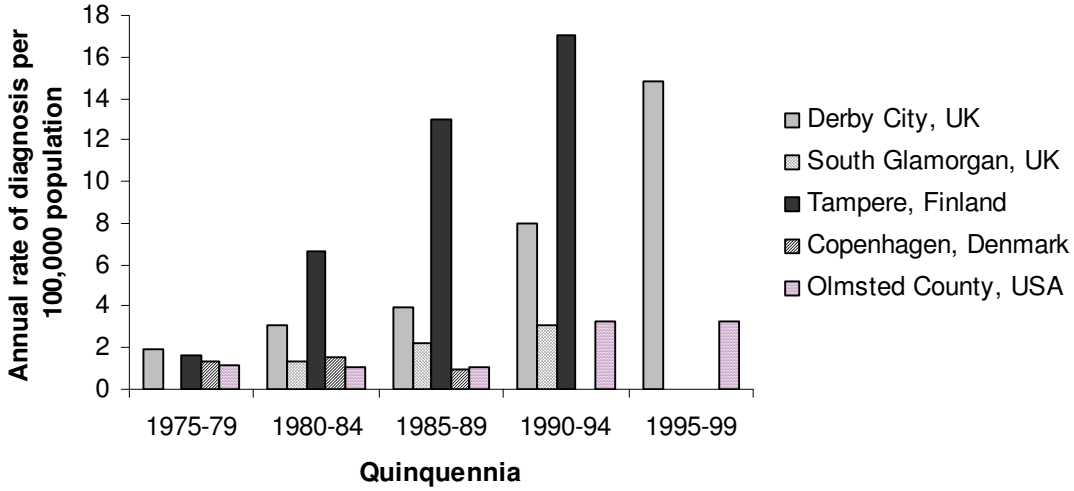
Studies of the incidence of childhood coeliac disease in the UK have shown a general decline in incidence during the mid 1970's[29-31]. Other countries, apart from Sweden, have also showed some evidence of a decline[99, 100]. Most of these studies calculated cumulative incidence as the number of cases identified by year of birth with the denominator being the total births in the same year. Other studies from that era used hospital admission data, self- or parent-reported coeliac disease, and membership of national coeliac societies all of which are liable to variation by ascertainment[101, 102]. The explanation for this decline may be related to infant feeding practices as discussed in section 1.1.3.1. Another explanation may be that the age of diagnosis of coeliac disease rose during this period however there is little direct evidence of this in the available data.

More recently some studies have suggested that the incidence of childhood coeliac disease may be rising during the late 1980's and 1990's. The data from Sweden reported by Ivarsson et al related the rise in childhood coeliac disease seen to infant feeding practices[37]. In a recent population based study from the Netherlands however there was no such change in infant feeding practices during the period studied yet an increase in incidence[63]. In the South Glamorgan, UK incidence rates in children appear stable but the population covered in this study was not large enough to allow valid comparisons over time[48].

1.1.7.2.2 Adult coeliac disease

The rate of diagnosis of adult coeliac disease has risen dramatically in most areas of the world where there is data available to monitor such trends[48, 96-98]. Although rate of diagnosis does not completely represent incidence, as diagnostic mechanisms have changed with the advent of serological markers, certainly coeliac disease is being more commonly recognised. The estimated annual rate of diagnoses from various areas is shown in Figure 1-3. An interest in coeliac disease research in some centres probably combined with an active “case-finding” strategy may explain the variation apparent in these figures. In combination with the diagnosed disease prevalence estimates they though do indicate the substantial gap between the number of people with clinically diagnosed and undetected disease. The ratio of “undetected” disease to symptomatic disease is probably therefore around 8:1.

Figure 1-3. Estimated annual rate of diagnosis (by quinquennia) of coeliac disease from various geographical areas



1.2 What is the impact of previously undetected coeliac disease?

The implications of recognising undetected coeliac disease at a general population level are unclear since the few reported data on the morbidity and physiological characteristics associated with previously undetected disease are from small, selected, case series. No studies so far have been able to look at a wide variety of socio-demographic and physiological factors with respect to undiagnosed coeliac disease as most adult screening studies in the general population have identified only small numbers of previously undiagnosed cases and have therefore been unable to examine any associations in comparison with the general population[103-105].

Of the studies that have looked at the differences between clinically diagnosed disease and “screening” or previously undetected disease, most have focussed on bone mineral density and anthropometric measurements. The findings, although not consistent, suggest that people with undetected coeliac disease have a slight tendency towards low bone density and measurements in keeping with mildly subnormal nutritional status[103-105]. There have been no studies large enough to look at relationships between undetected coeliac disease and important co-morbidity such as cardiovascular disease or stroke nor have they been able to examine mortality or malignancy in such a group.

1.3 What is the impact of clinically diagnosed coeliac disease?

1.3.1 Osteoporosis and fracture

The clearly documented association between osteoporosis and coeliac disease may be due to a combination of malabsorption, secondary hyperparathyroidism and abnormal calcium homeostasis[106-108]. As a consequence of osteoporosis there may be an increased risk of fracture in people with coeliac disease. Due to this perceived increase in fracture risk, some groups have recommended screening and surveillance of people with coeliac disease for decreased bone mineral density in order to implement treatment with bisphosphonates or hormone replacement therapy[43, 109, 110]. However, precise estimates of the excess fracture risk experienced by people with coeliac disease in comparison with the general population are not available, mainly because previous studies have been limited by their size, selected nature and inability to adjust for potential confounders[111-114]. Vasquez et al compared the fracture experience of 165 patients (median age 40 years) with that of controls with functional gastrointestinal disorders and found a three-fold increase in overall fracture risk (Odds ratio OR 3.5 95% CI 1.8-7.2) based on 25% of patients reporting fractures and only 8% of controls[112]. A subsequent study from the same group found that the increase in fracture experience was confined to people with coeliac disease presenting with “classical malabsorption” (OR 5.2 95% CI 2.8-9.8) compared to age-sex matched controls[115]. Fickling et al reported a “relative risk” of fracture of 7 based on a survey of 75 patients with a mean age of 52 years and age and sex matched controls selected from patients who had attended for bone densitometry[114]. While 21% of their patients with coeliac disease reported a past fracture, only two (3%) of the controls did.

In contrast both Vestergaard et al, in a database study of 1021 hospital diagnosed subjects with coeliac disease, and Thomason et al, using a mailed questionnaire survey of 244 cases, found no increase in the overall

fracture risk compared to the general population but with wide confidence limits[111, 113].

1.3.2 Vascular disease

Whether coeliac disease might afford protection from certain diseases, due partly to chronic malabsorption, needs consideration particularly since any protection might be reduced by treatment with a gluten free diet. This possibility was first raised by Whorwell et al who found a 40% reduction in ischaemic heart disease mortality in coeliac disease[116]. The possibility exists that low-grade chronic malabsorption, although leading to some vitamin and mineral deficiencies, may confer benefit through fat malabsorption or possibly altered salt homeostasis leading to, for example, lower serum cholesterol or lower blood pressure. Only one previous study has looked at serum cholesterol in diagnosed coeliac disease and concluded that “cholesterol malabsorption” led to the relative hypocholesteroaemia they found[117].

1.3.3 Malignancy and mortality

The early studies of the risk of malignancy and mortality in people with coeliac disease suggested a 2-fold increase in mortality rate, and greatly increased risks of lymphoproliferative malignancies. Most studies were small or not population-based, and their findings probably do not reflect the risks in contemporary coeliac disease[118-122]. More recent data from Sweden based on cases from their hospital inpatient register have suggested more modest increases in the risks but still found that people with coeliac disease were at excess risk of certain malignancies and death[123, 124]. Although large and population-based these studies were dependent on hospital admission of the index case for ascertainment. It is possible, therefore, that this may have led to an overestimate of the risks. The majority of other studies have found overall increased risks for malignancy or mortality of 2-fold or greater[118, 119, 121, 122]. These studies have been in cohorts of people with coeliac disease diagnosed and followed up some time ago, or from specialist referral centres. Generally they have been limited by not being population based or too small to provide robust estimates.

In contrast to the overall increased risks of mortality and malignancy two studies have suggested a decrease in the risk of breast cancer in people with coeliac disease, the reasons for which are not clear[121, 124]. A lack of breast cancer in people with coeliac disease was observed in both the Lothian and Swedish cohorts yet the former study was too small to be sure of the association and Askling et al were concerned that, among many comparisons, it may have been a chance finding.

1.4 Justification for the studies in this thesis

The development of accurate serological testing for coeliac disease has led some to debate the limitations and merits of population screening and/or screening “at risk groups” for the disease[125-128]. Indeed, coeliac disease fulfils several of the requirements of a condition suitable for population screening, according to the World Health Organisation criteria namely it appears to be relatively common, a suitable screening test is available and an effective treatment exists. However there are still areas where there is an absence of good information, in particular, about the natural history of the disease and whether by early detection (and treatment) of coeliac disease there can be an improvement in health at either a population or individual level. The possibility of doing harm by identifying previously unknown coeliac disease, and changing its physiological effects through treatment, is rarely discussed.

By more clearly understanding the physiological and socio-demographic associations of previously undetected disease and how common it is I might be able to add to this debate. The morbidity and mortality associated with clinically diagnosed coeliac disease are also not well quantified. Through understanding the impact of clinically diagnosed coeliac disease on health I may be able to provide information for people with the disease, and in addition possibly make inferences about the impact of previously undetected disease.

1.5 Objectives

The overall aim of this thesis is to achieve the following objectives. The section number in which each objective is addressed is given below.

1. To estimate the seroprevalence of undetected coeliac disease in England (section 2)
2. To explore the relationship between undetected coeliac disease and various socio-demographic characteristics and physiological measures (section 2)
3. To quantify the impact of diagnosed coeliac disease compared to the general population on the risk of:
    a. Fracture (section 4)
    b. Vascular disease – hypertension, high cholesterol diagnosis, atrial fibrillation, myocardial infarction and stroke (section 5)
    c. Malignancy (section 6)
    d. Mortality (section 6)

The overall methods for objectives 1 and 2 are described in section 2 and for objective 3 are described in section 3.

2 Seroprevalence, correlates and characteristics of previously undetected coeliac disease in England

2.1 Introduction

This section will describe a general population-based study of the prevalence of undetected coeliac disease. The section also describes the physiological and socio-demographic associations with undetected coeliac disease.

2.2 Methods

2.2.1 Participants

The Cambridge General Practice Health Study identified individuals aged 45-76 registered with 12 general practices (a list of registered names were held by the Cambridgeshire Family Practitioner Committee) and invited them for a health survey and bone density measurement between 1990 and 1995.[129-131] All those consenting completed a health and lifestyle questionnaire that included questions on occupation, past medical history, cigarette smoking habit and whether they rated their general health as "excellent, good, moderate or poor". They were also asked about known illnesses using the question: "Have you ever been diagnosed by a doctor as suffering from any of the following conditions? (please tick yes or no)" followed by a list of specific conditions. Self reported cardiovascular disease was defined as a positive answer for any one or more of the conditions: heart attack, angina, high blood pressure and stroke. Participants then attended for a physical examination including measurement of blood pressure and bone mineral density (BMD), the latter measured at the total hip, and total spine by dual energy X-ray absorptiometry, using the Hologic QDR-1000 densitometer (Hologic, Waltham, MA). Blood samples were taken by venepuncture and serum lipids were measured in fresh samples in the Hinchingbrooke Hospital biochemistry laboratory. All participants have been flagged at the Office for National Statistics for mortality and have been followed up to the end of May 2001. Death certificates were coded according to the 9th Revision of the International Classification of Diseases.[132] I was unable to gain access to participants for the purposes of duodenal biopsy.

2.2.2 Ethical approval

The study was approved by the Cambridge District Local Research Ethics Committee.

2.2.3 Serology

Serum samples have been stored at -15 degrees Celcius or below and had been thawed a maximum of three times previously. The sera were investigated for the presence of IgA class antiendomysial antibody (EMA) using indirect immunofluorescence on commercial monkey oesophagus sections (The Binding Site, Birmingham, UK), using a 1 in 10 serum dilution. A positive control was included with every batch of 40 samples. Samples positive for EMA were further tested for human anti-tissue transglutaminase antibody (tTGA), using a commercially available quantitative ELISA kit (Celikey®, Pharmacia diagnostics AB, Freiburg, Germany). Results of <3 U/mL were considered to be negative based on the experience in the laboratory at the Derbyshire Royal Infirmary. Total serum IgA was measured in all sera and considered results of <0.05g/L to indicate selective IgA deficiency.

2.2.4 Definitions

2.2.4.1 Coeliac disease

I defined undetected coeliac disease as those participants without self reported coeliac disease that had a positve EMA test. Definitions for the presence or absence of coeliac disease in participants follow:

- Participants with probable coeliac disease (treated): those who reported taking a gluten free diet and having a medical condition coded as malabsorption (including coeliac disease) and who were EMA negative

- Participants with probable coeliac disease (untreated): those who reported having a medical condition coded as malabsorption (including coeliac disease) but did not report taking a gluten free diet, and who were EMA positive

- Participants with possible coeliac disease: those who reported taking a gluten free diet but did not report having a medical condition coded as malabsorption (including coeliac disease) and who were EMA negative

- Participants not previously diagnosed with coeliac disease: those that did not report being on a gluten free diet, or having any other medical condition coded as malabsorption (including coeliac disease). This group was further subdivided into those who were EMA negative (no evidence of coeliac disease) and those who were EMA positive (undetected coeliac disease). The latter group was further subdivided on the basis of tTGA results

2.2.4.2 Socio-economic status

Socio-economic status was coded from the highest reported occupation of the participant or their spouse using the registrar general's classification and grouped as: professional and managerial, lower professional – group 1, non-manual skilled, manual skilled – group 2, partly skilled, unskilled, armed forces, inadequately described, housewife/homemaker, retired – group 3.

2.2.4.3 Osteoporosis

Participants were defined as having osteoporosis if their BMD measurement was 2.5 standard deviations (SD) or more below the young adult mean.

2.2.4.4 Anaemia

Participants were categorised into those who were anaemic and those who were not (anaemia = haemoglobin in men<13 g/dL and in women<11.5 g/dL)

2.3 Statistical analysis

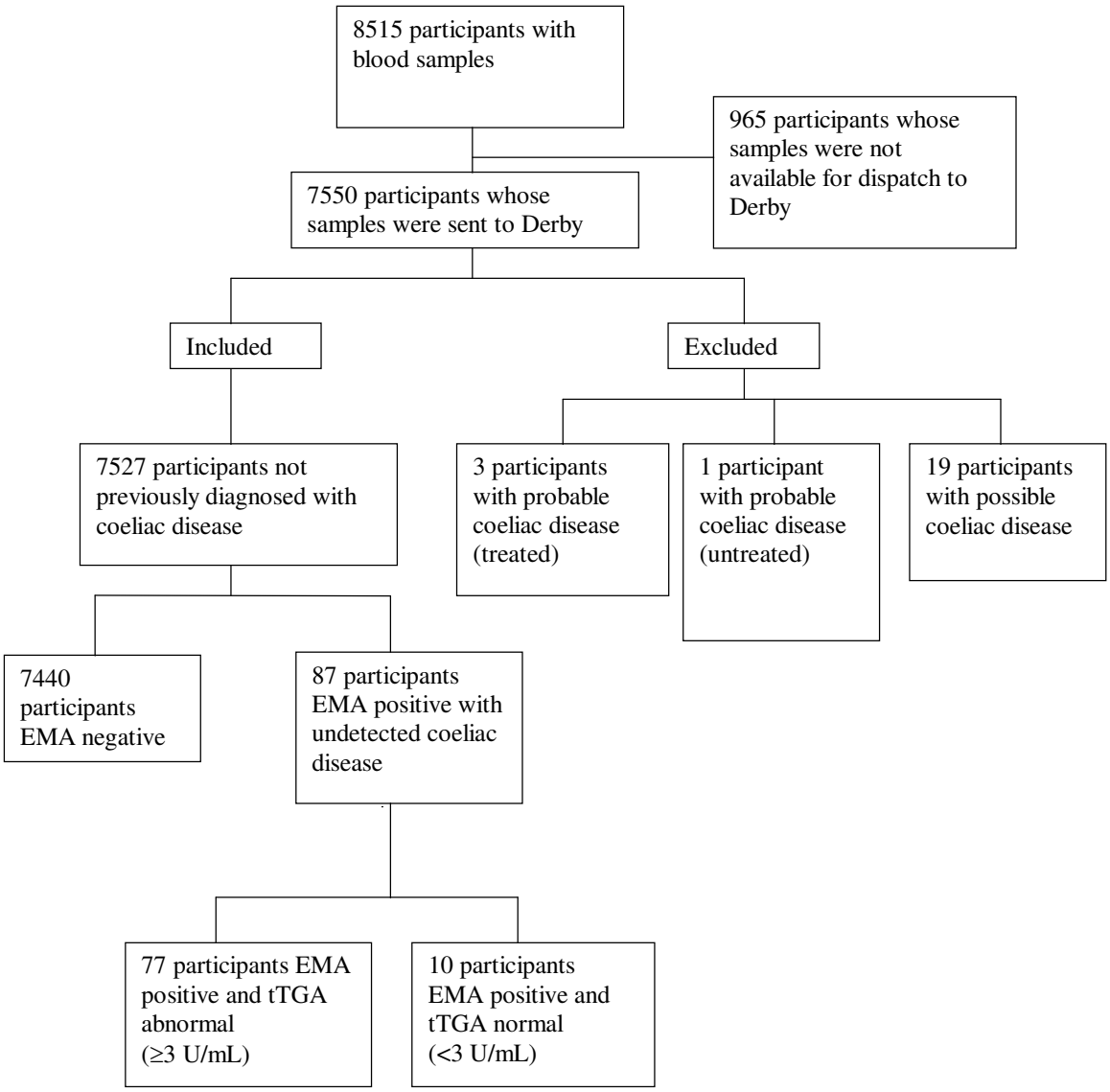
Fisher's Exact test, χ^2 tests and χ^2 tests for trend were used to examine the association between smoking, socio-demographic and other binary variables and EMA positivity. Comparisons between EMA positive and EMA negative participants with regard to laboratory, anthropometric and bone density variables were examined using independent samples t-tests. Multivariate

analyses were performed to adjust for age, gender, smoking, social class and other potential confounders using logistic regression for binary and multiple linear regression for continuous dependent variables. We tested biologically plausible interactions, particularly those with gender, by adding multiplicative interaction terms to the multivariate linear regression model. Where data on confounders was missing these data were modelled as separate categories to ensure nested models contained the same number of participants. All significance tests were two-sided. The assumptions of multiple linear regression were checked by examining histograms of residuals and normal probability plots.

2.4 Results

A total of 20,314 individuals were mailed an invitation to participate. Of those mailed 8515 (42%) agreed to participate. Of those who participated, there were 7550 (89%) with blood specimens available for serological testing. Of the 7440 EMA negative participants, 9 had evidence of selective IgA deficiency. An overview of the number of participants in the study is shown in Figure 2-1.

Figure 2-1. Overview of study participants



There were a total of 7527 participants not previously diagnosed with coeliac disease included in the analyses. The mean age of these participants was 59 years (SD 8.9) and 4444 (59%) were female. Estimates of the seroprevalence of undetected coeliac disease are shown in Table 2-1. The overall seroprevalence of undetected coeliac disease was 1.2% (95% CI 0.9-1.4) based on EMA positivity alone (n=87) and 1.0% (95% CI 0.8-1.3) based on EMA positive and abnormal tTGA results (n=77). Seroprevalence showed some variation by age and gender.

Table 2-1. Seroprevalence of undetected coeliac disease using antiendomysial antibody (EMA) and human anti-tissue transglutaminase (tTGA) tests by age and gender

	<55 years		55–64 years		>=65 years		Total	
	M	F	M	F	M	F	M	F
	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)
EMA +ve	12(1.1)	25 (1.4)	12(1.3)	18(1.3)	6 (0.6)	14(1.1)	30(1.0)	57(1.3)
EMA +ve & tTGA +ve	10(0.9)	20(1.1)	11(1.2)	16(1.2)	6(0.6)	14(1.1)	27(0.9)	50(1.1)
Total	1087	1744	925	1384	1071	1316	3083	4444

Table 2-2 shows the distribution of tTGA results in the EMA positive participants considered to have undetected coeliac disease. Seventy five percent of the EMA positive participants' tTGA results were unequivocally abnormal (>6 U/mL).

Table 2-2. tTGA results in the EMA positive participants considered to have undetected coeliac disease

tTGA	EMA +ve	
Category	n	%
<3	10	11.5
3-6	11	12.6
>6-10	8	9.2
>10-20	23	26.4
>20	35	40.2
Total	87	100

EMA positive participants were less likely to have reported being ex or current smokers (Table 2-3), compared to EMA negative participants, and also showed a trend towards higher social class, although this was not significant at the 5% level (χ^2 for trend, p=0.08). Mutual adjustment and adjusting for age and gender did not appreciably change these associations.

Table 2-3. Sociodemographic and smoking characteristics of study participants by antiendomysial antibody (EMA) result

	EMA-ve	EMA+ve	Odds ratio for positive EMA test (95% CI)	
	n (%)	n (%)	Univariate	Multivariate*
Gender				
Female	4387(59.0)	57(65.5)	1	1
Male	3053(41.0)	30(35.5)	0.76(0.49-1.18)	0.83(0.52-1.32)
Age group				
< 55 years	2794(37.6)	37(42.5)	1	1
55-64 years	2280(30.6)	30(34.5)	0.99(0.61-1.61)	1.01(0.62-1.65)
>= 65 years	2366(31.8)	20(23.0)	0.64(0.37-1.10)	0.67(0.38-1.16)
Social class group				
Professional	3573(48.0)	49(56.3)	1	1
Skilled	3011(40.5)	32(34.5)	0.78(0.50-1.21)	0.82(0.52-1.29)
Partly skilled and unskilled	649(8.7)	4(4.6)	0.45(0.16-1.25)	0.51(0.18-1.43)
Missing values	207(2.8)	2(2.3)		
Smoking status				
Never	3371(45.3)	52(59.8)	1	1
Ex	3024(40.6)	30(34.5)	0.64(0.41-1.01)	0.71(0.45-1.14)
Current	958(12.9)	5(5.7)	0.34(0.14-0.85)	0.36(0.14-0.90)
Missing values	87(1.2)	0(0)		

* Mutually adjusted for other variables in the table

In the univariate analyses EMA positive participants had lower mean haemoglobin, total protein, corrected calcium, cholesterol, low density lipoprotein, triglyceride, diastolic blood pressure and weight (all $p<0.05$, Table 2-4). Mean BMI, hip and spine BMD, systolic blood pressure, mean cell volume and albumin were all slightly lower in participants with undetected coeliac disease, but these differences were not significant at the 5% level. Both alanine aminotransferase and platelet count were higher among EMA positive participants. After adjustment for age, gender, smoking status and social class the differences in weight and diastolic blood pressure were not significant at the 5% level.

In the multivariate analyses undetected coeliac disease was associated with a reduction of 0.5 mmol/L (95% CI 0.3 to 0.8 mmol/L) in cholesterol and 0.3 g/dL (95% CI 0.1 to 0.5 g/dL) in haemoglobin. For BMD the mean difference at the hip was -0.02 g/cm^2 (95% CI -0.05 to +0.02 g/cm^2) and at the spine 0.03 g/cm^2 (95% CI -0.07 to +0.02 g/cm^2). There were significant interactions between the effects of EMA result and gender upon corrected calcium ($p=0.001$) and high density lipoprotein ($p=0.001$) only. Mean corrected calcium was reduced in EMA positive participants among women only (–0.05 mmol/L 95% CI –0.07 to –0.02 mmol/L) and there was a similar effect for high density lipoprotein (–0.14 mmol/L 95% CI –0.24 to –0.04 mmol/L).

There were no statistically significant associations between undetected coeliac disease and reported morbidity but reporting having high blood pressure, high cholesterol, angina or heart attack, diabetes or bronchitis/emphysema all appeared to be less common in those who were EMA positive (Table 2-5).

Table 2-4. Distribution of selected physiologic variables by antiendomysial antibody (EMA) test result

Dependent variable	EMA-ve	EMA+ve	Mean difference (standard error)	
	Mean (n)	Mean (n)	Univariate	Multivariate§
Haemoglobin (g/dL)	13.7(7425)	13.3(87)	-0.4(0.2)**	-0.3(0.1)**
Mean cell volume (fl)	89.4(7423)	88.9(87)	-0.6(0.6)	-0.4(0.5)
Platelet count ($x10^9$/L)	260.9(7406)	279.7(87)	18.8(9.2)*	18.2(6.5)**
Total protein (g/L)	71.1(7436)	70.2(87)	-0.9(0.5)*	-1.0(0.5)*
Albumin (g/L)	42.0(7438)	42.0(87)	-0.1(0.4)	-0.3(0.4)
Alkaline phosphatase (IU/L) $^{\infty}$	161.9(7438)	163.6(87)	1.0(0.9-1.1)	1.0(1.0-1.1)
Alanine aminotransferase (IU/L)	16.7(7438)	20.8(87)	1.2(1.1-1.4)**	1.2(1.1-1.4)**
Corrected calcium (mmol/L)	2.33(7431)	2.30(87)	-0.03(0.01)*	-0.02(0.01)*
Cholesterol (mmol/L)	6.4(7430)	5.8(86)	-0.6(0.1)**	-0.5(0.1)**
Hdl (mmol/L)	1.2(6838)	1.1(82)	-0.02(0.04)	-0.05(0.04)
Ldl (mmol/L)	4.4(6654)	4.1(81)	-0.3(0.1)*	-0.3(0.1)*
Triglyceride (mmol/L) $^{\infty}$	1.6(6850)	1.3(81)	0.8(0.7-0.9)**	0.8(0.7-0.9)**
Weight (kg)	71.6(7428)	68.6(87)	-2.9(1.4)**	-2.2(1.2)
Height (cm)	166.4(7427)	166.0(87)	-0.5(1.0)	-0.1(0.7)
Body mass index (kg/m^2)	25.8(7425)	25.0(87)	-0.8(0.4)	-0.7(0.4)
Systolic blood pressure (mm/Hg)	137.0(7417)	134.1(87)	-3.0(2.2)	-1.6(2.0)
Diastolic blood pressure (mm/Hg)	82.5(7417)	80.1(87)	-2.4(1.1)*	-1.9(1.2)
Total spine BMD (g/cm^2)	0.97(5035)	0.94(58)	-0.03(0.02)	-0.03(0.02)
Total hip BMD (g/cm^2)	0.90(5024)	0.88(58)	-0.02(0.02)	-0.02(0.02)

§ adjusted for age group, gender, social class group and smoking status, * p<0.05, ** p<0.01, $^{\infty}$ geometric mean, ratio of geometric means and 95% CIs

Table 2-5. Relation between participants with positive antiendomysial antibody (EMA) test result and reported morbidity

	Reported morbidity		Odds ratio (95% CI)	
EMA result	No	Yes	Univariate	p value*
	n (%)	n (%)		
	High blood pressure§			
Negative	5876(81.4)	1347(18.6)	1	
Positive	73(83.9)	14(16.1)	0.84(0.47-1.49)	0.6
	High blood cholesterol			
Negative	6188(87.6)	875(12.4)	1	
Positive	81(93.1)	6(6.9)	0.52(0.23-1.20)	0.2
	Stroke			
Negative	7149(98.5)	109(1.5)	1	
Positive	85(97.7)	2(2.3)	1.54(0.37-6.35)	0.4
	Heart attack and/or angina			
Negative	6778(93.3)	487(6.7)	1	
Positive	85(97.7)	2(2.3)	0.33(0.08-1.34)	0.2
	CVS combined			
Negative	5570(76.6)	1701(23.4)	1	
Positive	70(80.5)	17(19.5)	0.79(0.47-1.35)	0.5
	Thyroid disease			
Negative	6806(93.8)	446(6.2)	1	
Positive	80(92.0)	7(8.0)	1.33(0.61-2.91)	0.6
	Diabetes			
Negative	7049(97.2)	202(2.8)	1	
Positive	86(98.9)	1(1.1)	0.41(0.06-2.93)	0.7
	Fracture of the wrist			
Negative	6806(93.8)	452(6.2)	1	
Positive	80(92.0)	7(8.0)	1.32(0.60-2.87)	0.6
	Bronchitis/emphysema			
Negative	6125(84.7)	1110(15.3)	1	
Positive	79(90.8)	8(9.2)	0.56(0.27-1.16)	0.2
	Asthma			
Negative	6520(90.1)	720(9.9)	1	
Positive	78(89.7)	9(10.3)	1.04(0.52-2.09)	1
	Cancer			
Negative	6853(94.5)	396(5.5)	1	
Positive	82(94.3)	5(5.7)	1.05(0.43-2.62)	0.8

* χ^2 test with continuity correction or Fisher's exact test, § except during pregnancy

In 75 of 85 (88%) EMA positive participants and 5527 of 7015 (79%) EMA negative participants, general health was reported as being “good or excellent” giving an odds ratio of 1.76 (95% CI 0.90 to 3.46) adjusted for age, gender, social class and smoking status. There was no difference in the number of deaths recorded (EMA negative 591, 7.9%; EMA positive 5, 5.7%; χ^2 p=0.579). Underlying cause of death given for the five EMA positive participants were carcinoma of the pancreas, acute myeloid leukaemia, ischaemic heart disease, carcinoma of the cervix and B cell lymphoma (not otherwise specified).

Fourteen of 87 (16.1%) EMA positive participants and 315 of 7425 (4.2%) EMA negative participants were found to be anaemic, giving an odds ratio of 4.56 (95% CI 2.53 to 8.21) adjusted for age, smoking status, gender and social class. Of the EMA positive anaemic participants, 9 were women (haemoglobin range 10.1 to 11.3 g/dl) and 5 were men (haemoglobin range 11.6 to 12.8 g/dl). In 7 of 58 (12.1%) EMA positive participants and 311 of 5024 (6.2%) EMA negative participants BMD at the hip showed osteoporosis, giving an odds ratio of 3.08 (95% CI 1.31 to 7.25) adjusted for age, BMI, smoking status and social class.

2.5 Discussion

2.5.1 Principal findings

The study shows that undetected coeliac disease as assessed by EMA positivity affects approximately 1% of this general population sample aged 45-76 years. In comparison with earlier screening studies the number of EMA positive participants was sufficiently large to use data collected at recruitment for comparisons with the EMA negative participants. Although the positive participants were more likely to assess their own health as good or excellent than the negative participants, some were mildly anaemic (16%) or had evidence of osteoporosis (12%). In contrast they had a more favourable cardiovascular risk profile in terms of having lower cholesterol levels, slightly lower blood pressure and smoking less than the EMA negative participants.

2.5.2 Limitations and merits

Unlike earlier studies it was not possible to confirm the diagnosis of coeliac disease by intestinal biopsy in the EMA positive participants. The validity of the findings therefore is mainly dependent on the specificity and, to a lesser extent, the sensitivity of the EMA test. In routine clinical practice the test has proved to be extremely accurate with a sensitivity of 94% and a specificity of 99% quoted recently.[67] Recent data from the laboratory in Derby (unpublished) has estimated a specificity of 99.8% based on 1468 EMA tests. In earlier screening studies a total of 57 participants have been found to be EMA positive and in all but three a diagnosis of coeliac disease has been supported by abnormal biopsy findings.[76, 79-84] The recent introduction of the human tTGA assay in the laboratory allowed us to confirm that 89% (77/87) of the EMA positives also had an abnormal tTGA level.[94] The tTGA assay we used has also been shown to have high concordance with EMA Therefore while it is possible that a few of the EMA positive participants do not have abnormal intestinal histology, the results from the tTGA assay indicate this is likely to be at most 11%. In addition the sensitivity of the EMA test means that about 5% of those with coeliac disease will not

have been detected which is therefore likely to be less than 5 missed cases in this study.

I chose not to restrict the analyses to those positive for both tests for two reasons. Firstly the sensitivity of human tTGA is not yet well-established; a figure of 96% has been estimated recently.[68] Secondly the tTGA testing was performed only on the EMA positive subjects whereas the EMA testing was performed on the whole sample. When I compared those with both EMA positive and abnormal tTGA results to the EMA negative participants the associations presented here were of similar magnitude and in the same direction. None of the associations were changed towards the null.

Although this was a general population sample there were greater numbers of women and those from higher social class groups, which is likely to be a consequence of the response rate to the original mailed invitation. I had no information about the non-responders, but it seems unlikely that those people who agreed to participate were, in some way, more likely to have undetected coeliac disease and therefore biased the findings towards an overestimate of seroprevalence.

2.5.3 Comparison with other studies

The seroprevalence estimate is similar to the findings of others who have performed smaller screening studies in the UK.[77, 133] I found a trend towards lower seroprevalence with lower social class, which is unexplained and in contrast to many diseases. It would appear not to be completely explained by smoking as the adjusted odds ratios show the same trend. The numbers involved are small but the trend could possibly reflect events in early life such as infant feeding practices that vary by social class and have been shown to influence the development of coeliac disease in childhood.[36, 134]

As might be expected with undetected coeliac disease the mean haemoglobin, corrected calcium and total protein levels were lower in the EMA positive participants and 16% had a mild anaemia compared with 4% of the EMA negative participants. Overall most differences were small. The

increases in serum alanine aminotransferase and in platelet counts were small but similar to those found in clinically diagnosed coeliac disease.[50, 57] While there have been reports of significant decreases in bone mineral density in screen-detected or subclinical coeliac disease, I have found only a small, non-significant decrease in bone mineral density.[103, 105] However, the prevalence of osteoporosis defined according to World Health Organisation criteria was 12% in the EMA positives, twice that of the EMA negatives.

The EMA positive participants did not regard themselves as unwell; indeed the numbers reporting good or excellent health were greater than in the EMA negatives, although this difference was not statistically significant. Reported morbidity was no greater in the positives and there was a trend towards less cardiovascular morbidity. The finding of an 8% reduction in serum cholesterol among EMA positive participants would be expected to have a significant impact on cardiovascular morbidity. It has been estimated that a 0.6 mmol/L lower cholesterol will confer a 25% reduction in incidence of ischaemic heart disease.[135] The data suggest that undetected coeliac disease may afford protection from ischaemic heart disease, a hypothesis first raised by Whorwell et al some 25 years ago.[116]

I found that few of the EMA positives were smokers and that as a group they reported much less smoking in the past. Four published case control studies have examined the relationship between cigarette smoking and coeliac disease. While three have found a positive association between not smoking or never smoking and having coeliac disease, the one carried by Patel et al did not.[39-42] In these studies it has not been clear how much of the association might be accounted for by selection and reporting biases and by cases stopping smoking after diagnosis. In addition while the proportion of cases (people with coeliac disease) reporting current smoking (~40%) has remained fairly consistent the amount of smoking in the control groups varied considerably. The problem of selection bias for this relationship should not

apply to my study and the findings therefore suggest that this is indeed a causal relationship probably analogous to that seen in ulcerative colitis.

2.5.4 Summary

I have found that undetected coeliac disease is likely to affect about 1% of the general population in England. Although these people are at increased risk of mild anaemia and osteoporosis, they do not regard themselves as unwell. The important finding of a favourable cardiovascular risk profile in these individuals suggests that any screening programme of the general population would need to be carefully evaluated in terms of risks and benefits before its introduction.

3 Description of the General Practice Research Database (GPRD) and the coeliac disease dataset

3.1 Introduction

This section will describe the General Practice Research Database and how the study population, disease status and other variables were defined that are common to subsequent chapters. It contains the descriptive results of the study population and the distribution of the other variables defined.

3.2 The General Practice Research Database

The General Practice Research Database (GPRD) is a longitudinal primary care database and contains the computerised medical records from general practice of more than 8 million of these registered people.[136-141] When people are seen in primary care in the UK the majority of significant medical diagnoses, information from hospital letters and discharge summaries, and prescriptions are entered onto a desktop computer. These data are then aggregated and anonymised to maintain patient confidentiality. The database was started in 1987 under the name Value Added Medical Products (VAMP), which was a geographically representative group of doctors that collected data according to a research protocol that included collection of all prescriptions and medically significant events.[142] At the end of 1994, Reuters acquired VAMP database and gave it to the Department of Health, who renamed the database the GPRD. Currently the Medicine Control Agency manages the database. Practices are required to record at least 95% of prescribing and relevant patient-encounter events and there are regular routine validity checks.[141] The data practices contribute on this basis is named “up-to-standard” data. Events that occur prior to the “up-to-standard” period are also recorded if they are considered important medical diagnoses.

Structurally the GPRD is a relational database divided into four data tables linked by a unique identification string variable. The four tables contain patient records, medical records, therapy records and prevention records. The patient records contain information on date of birth, family identification

code, gender, marital status, registration date, code for usual doctor, and prescription exemption status. The medical records contain all the records on medical diagnoses coded using Oxford Medical Information System (OXMIS) and Read codes, each of which is coded with the event date. OXMIS and Read codes are hierarchical codes commonly used in general practices in England. Most of the data within the GPRD are coded with OXMIS codes, but more recently many practices have converted to Read codes. Data from the GPRD uses a combination of these codes depending on the practice. Information from speciality consultations is also recorded in this section. The third section is the therapy records and this contains all the information on prescribing including the date of the prescription, the drug details are coded using the prescription pricing authority system, quantity, and the dosage instructions. The last section of the GPRD is the prevention records, which contains information on other aspects of medical care such as smoking status, contraception use and vaccines given.

3.2.1 Ethical approval

Ethical approval was given by the GPRD Scientific and Ethical Advisory Group (Protocol number 361).

3.3 Limitations and merits of the GPRD

The GPRD has good qualities for use as a tool for disease based epidemiology, but as with any data source it has limitations. The following sections will address some of these merits and limitations. When designing, carrying out and reporting the studies presented in the following chapters, these factors were taken into consideration.

3.3.1 Size

For studies of coeliac disease one of the greatest strengths of GPRD is its size. Although as many as 1% of the general population may have undetected coeliac disease, as shown in section 1, the prevalence of diagnosed coeliac disease is estimated to be approximately 0.1 to 0.3%. To study the impact of such a relatively rare condition on specific (also relatively rare) outcomes, such as fracture or malignancy requires a cohort of

considerable size. The GPRD provides an opportunity to assemble such a cohort and therefore has the potential to achieve precise estimates of risk.

3.3.2 Representative.

Although the practices in GPRD are self-selected they are from a wide variety of different areas of the UK, and have been shown to have levels of morbidity that are very similar to national estimates[137]. It is therefore reasonable to generalise results from the GPRD to the population of the UK as a whole.

3.3.3 Prospectively collected

The identification of each subject's entry date to the dataset ensures that it is easy to ascertain which data are prospectively, and which retrospectively recorded. This is of particular importance for studies where measurement of outcome may be biased, for example fracture. The data reflect the occurrence of disease and subsequent management in primary care, so when an outcome (fracture) occurs, or is notified to the general practitioner, it is recorded. Although there is the potential for some misclassification of the date of recording of medical diagnoses, this is unlikely to be the case for the writing of a prescription, which are generated on the general practitioner's desktop computer contemporaneously.

3.3.4 Contemporary

For the studies involving survival analysis for outcomes in this thesis, for example time to fracture and time to death I have used data from June 1987 to April 2002. These data represent the contemporary experience of people with coeliac disease and their matched control cohort. For the outcomes I have studied this represents an improvement on previous studies that, for the most part, have used data further away in time from the present.

3.3.5 Validity

Although as outlined above strict data standards are maintained for the GPRD, systems also exist for the independent validation of the data by third parties. The mechanism for this is via the requesting of anonymised copies of paper records, or the completion of questionnaires by general practitioners. This system although it is sufficiently expensive (£70-200

approximately per validated subject) to restrict its use does permit validation of small samples. Using this system the GPRD has been extensively validated for a wide range of diagnoses and consistently found to be accurate[138, 143-146]. Described in section 3.5 is a small validation study of the diagnosis of coeliac disease using this methodology that I have carried out.

3.3.6 Duration of follow up

Although the GPRD is the largest available prospectively collected general practice dataset, it is not as large as it first appears. There are no prospective records prior to 1987. In addition both practices and people have stopped contributing over the duration of the database. Therefore in total there are a large number of people contributing to the database but for relatively short periods. In practice this leaves relatively little person time following diagnosis in incident cohorts.

3.3.7 Incomplete recording

One problem with the use of any routinely collected data such as the GPRD is that what is recorded is determined not by the needs of the research, but by what is felt relevant to the primary purpose of the data recording. This means that recording is determined by the general practitioner's assessment of what is relevant to the ongoing primary medical care of subjects. Hence not only is data incomplete, but it is likely that there is bias as to which data are missing. For example it is likely that a general practitioner will record that a patient drinks heavily if they know this as it may adversely affect health, it perhaps less likely that they would record the knowledge that the patient was teetotal unless they suffered from a condition which might be attributable to alcohol.

3.4 The coeliac disease dataset: study population and definitions

3.4.1 Coeliac disease and control status

Records were extracted of all persons within the GPRD between June 1987 and April 2002 with a recorded diagnosis of coeliac disease using the codes listed in Table 3-1.

Table 3-1. Diagnostic codes for coeliac disease

Description	Code	Type of code
COELIAC DISEASE	2690B	OXMIS
COELIAC DISEASE	J690.00	READ
COELIAC DISEASE NOS	J690z00	READ
GLUTEN ENTEROPATHY	J690.13	READ
INFANTILE COELIAC DISEASE	2690D	OXMIS
ACQUIRED COELIAC DISEASE	J690100	READ

Where possible 5 control subjects were selected matched to each individual with coeliac disease by age, gender, general practice and follow up time. When selecting control subjects I excluded individuals who had any record of coeliac disease, dermatitis herpetiformis, a gluten-free prescription or a non-specific reference to coeliac disease e.g. “gluten free diet”, “gluten sensitivity” using the codes listed in Table 3-2. Prescriptions were identified for gluten free products from the GPRD drug index. Each control had to be alive and contributing data on the date (index date) of the first occurrence in their matched case’s record of any of the coeliac disease or gluten free product codes, within up-to-standard data.

Table 3-2. Medical codes used to exclude control selection

Code	Description
2690B	COELIAC DISEASE
J690.00	COELIAC DISEASE
J690z00	COELIAC DISEASE NOS
J690.11	COELIAC RICKETS
J690000	CONGENITAL COELIAC DISEASE
ZC2C200	DIETARY ADVICE FOR COELIAC DISEASE
J690.13	GLUTEN ENTEROPATHY
13B2.00	GLUTEN FREE DIET
8B55.00	GLUTEN-FREE DIET
2690D	INFANTILE COELIAC DISEASE
8CA4200	PT ADVISED RE GLUTEN FREE DIET
2690E	SENSITIVITY GLUTEN
J690100	ACQUIRED COELIAC DISEASE
J690.14	SPRUE - NONTROPICAL
693	DERMATITIS HERPETIFORMIS
M140.00	DERMATITIS HERPETIFORMIS
M142.00	JUVENILE DERMATITIS HERPETIFORMIS
M145200	SENILE DERMATITIS HERPETIFORMIS

3.4.2 Using a more restricted case definition

It is probable that not all of the individuals selected as having coeliac disease in the way described in section 3.4.1 have the disease. To take account of this possibility two additional more restricted case definitions were used. The first method was to include only those people who in addition to having a diagnostic code for coeliac disease had received at least one gluten free prescription. The second method was to select as cases only those individuals who had the use of any of the coeliac disease codes (Table 3-1) at least twice throughout their whole general practice record (within and not in up to standard data).

3.4.3 Date of diagnosis

Each person with coeliac disease was assigned a date of diagnosis defined as the date of the first record of a coeliac disease code as defined in Table

3-1 or a dermatitis herpetiformis code (see Table 3-2). Since general practitioners enter some data for important historical events retrospectively this date preceded the start of their GPRD up-to-standard record in some cases. Each control was assigned a pseudo-diagnosis date identical to the diagnosis date of their case.

3.4.4 Incident and prevalent status

I defined "incident" subjects with coeliac disease as those individuals whose diagnosis date of coeliac disease or first prescription for a gluten free product occurred at least 1 year after the beginning of their up-to-standard GPRD record as this has been shown to reflect incident diagnoses in inflammatory bowel disease, probably as around the time of registration with the practice there is increased recording of chronic conditions[144]. All other subjects with coeliac disease were defined as "prevalent".

3.4.5 Age

Age was calculated for each subject in subsequent studies at the beginning of follow up in each study. For the studies on fracture risk and vascular disease, age at the start of up to standard data was used. For the study on malignancy and mortality age at index date was used. Age was then grouped into 8 categories.

3.4.6 Gender

All subjects have their gender recorded.

3.4.7 Smoking status

Smoking status is not recorded for all subjects in the GPRD. It is recorded by the general practitioner if and when they enquire about it, and is therefore based on the individual's response to the doctor's enquiry. Subjects' smoking status was classified based upon the coding during up to standard data in the medical and prevention tables of GPRD as unknown, non-smoker, ex-smoker or current smoker. Smoking status was referred to by Oxmis or Read codes. Categorisation of these codes and the codes used are listed in **Error! Reference source not found.**. Subjects who appeared in more than one of these categories at differing times were coded in the category suggesting greatest smoking experience.

3.4.8 Body mass index (BMI)

All data coding height (metres) and weight (kilograms) during up to standard data recording were identified from the prevention table. Since BMI measured across time in this way is not necessarily applicable in children it was not calculated for subjects 15 years of age or younger. Inspection of the data showed many impossible or highly unlikely values. Records of height over 2.5 metres or under 1 metre were therefore ignored as being probable errors as well as records coding a weight under 30kg. BMI (kgm^{-2}) was then calculated using the median values of the recorded heights and weights.

3.4.9 Gluten-free prescription rate

The number of gluten-free prescriptions was extracted from the therapy file and divided by the observation time to calculate the rate of gluten-free product prescription. This value was grouped into categories: none, up to 9.99 and 10 or more.

3.4.10 Visit rate

Visit rate was calculated as the number of unique calendar dates with a medical diagnosis code divided by the observation time for each subject to estimate the amount each person visited their general practitioner. This rate was categorised by tertiles.

3.5 Validation of coeliac disease diagnosis

To evaluate the validity of the coeliac disease diagnosis in the studies using the GPRD a stratified (by prevalent/incident status and age group) random sample (n=34) of people with coeliac disease (as identified in section 3.4.1) were selected. In order to maximise responses the people with coeliac disease selected were, according to the data available, not dead and continuing to contribute to the GPRD. Using Comasco Computer Services Ltd (the company licensed for processing requests to general practitioners for more information) the general practitioners for each individual were contacted by letter. Each general practitioner was asked to provide, where possible, any confirmatory information available in the paper record relating to the diagnosis of coeliac disease. They were provided with individual identifiers and approximate age at and year of diagnosis. The returned

documentation was then read and each person was assigned a definition of his or her coeliac disease status as follows:

- Definitely not: no evidence of coeliac disease or a clear statement saying that the individual did not have the disease
- Definitely yes: clear evidence of coeliac disease, for example an appropriate histology report of a duodenal biopsy, or a clear statement indicating that the individual has coeliac disease
- Probably yes: no clear evidence against the assumption of a diagnosis of coeliac disease

3.6 Statistical analysis

Proportions were compared using χ^2 tests. The binomial distribution was used to calculate confidence intervals (CI).

3.7 Results describing the study population

The results for this section are shown in Table 3-3.

3.7.1 Numbers in the cohorts and person years at risk

The cohorts included 4732 subjects with coeliac disease and 23620 matched controls contributing 27116 and 149896 observed years at risk respectively.

3.7.2 Numbers with two diagnostic codes for coeliac disease

Of the 4732 people with coeliac disease there were 2761 (58.4%) with at least two diagnostic codes in their whole general practitioner record.

3.7.3 Age

The mean age at diagnosis of the “incident” subjects with coeliac disease was 44.7 years (SD 20.6). The cohorts were closely matched on age at the start of up to standard data.

3.7.4 Gender

Of the people with coeliac disease 67% were female. The control cohort was closely matched on gender.

3.7.5 Smoking status

Overall recorded smoking status varied between the people with coeliac disease and controls (χ^2 95.5, $p<0.001$). There were more current smokers

in the control cohort compared to the coeliac disease cohort (15.4% vs 13.0%).

3.7.6 Body mass index

Overall BMI varied between the people with coeliac disease and controls (χ^2 648.9, $p<0.001$). More individuals were underweight (BMI $\leq$ 18.5) in the coeliac disease cohort (4.2% vs 1.2%). Only 3.1% of people with coeliac disease were obese compared to 8.1% in the control cohort.

3.7.7 Gluten-free prescriptions

Only 10% of the coeliac disease cohort had never received a gluten-free prescription and over 54% had more than 10 prescriptions per year for gluten free products.

3.7.8 Visit rate

The cut points for the creation of tertiles were at a visit rate of 8.2 and 15.5 visits per year of follow up. Greater than 50% of the coeliac disease cohort were above the highest tertile of visit rate compared to only 29% of the control cohort (χ^2 1303.1, $p<0.001$).

Table 3-3. Description of the study population

	Coeliac disease cohort (n=4732)	%	Control cohort (n=23620)	%
Median observed time (years)	5.7		6.4	
Total observed time (years)	27116		149896	
Female	3095	65.4	18545	65.4
Age groups at start of UTS record (years)				
0-3	257	5.4	1521	6.4
>3-15	362	7.7	1664	7.0
>15-25	523	11.1	2813	11.9
>25-35	779	16.5	3782	16.0
>35-45	809	17.1	4066	17.2
>45-55	833	17.6	4113	17.4
>55-65	550	11.6	2707	11.5
>65-75	415	8.8	1994	8.4
>75	204	4.3	960	4.1
Smoking status				
Non smoker	2082	44.0	8623	36.5
Ex-smoker	221	4.7	1265	5.4
Current smoker	613	13.0	3630	15.4
Unknown	1816	38.4	10102	42.8
Body mass index, kgm^{-2}				
Less than or equal to 18.5	197	4.2	273	1.2
18.51 to 25	2010	42.5	6949	29.4
25.01 to 30	653	13.8	4319	18.3
Greater than 30	148	3.1	1920	8.1
Unknown	1724	36.4	10159	43.0
Tertiles of visit rate				
1st tertile	716	15.1	8874	37.6
2nd tertile	1503	31.8	7953	33.7
3rd tertile	2513	53.1	6793	28.8
Gluten free prescriptions per year of follow up				
None	452	9.6		
Between 0 and 10	1719	36.3		
10 or more	2561	54.1		
Individuals with at least two diagnostic codes for coeliac disease	2761	58.4		

3.7.9 Validation of the coeliac disease diagnosis

Of the 34 individuals selected for validation there were 32 (94%) responses from general practitioners. Of the 32 for which there was additional information 26 (81.3% 95% CI 63.6%-92.7%) had definite or probable coeliac disease (Table 3-4). Of those 26 only 3 had not had a prospectively recorded gluten free prescription. Therefore using the definition of having one diagnostic code and at least one prescription 23 (88.5% 95% CI 70.0%-97.6%) had definite or probable coeliac disease. That proportion increased when the definition requiring two diagnostic codes was used to 100% (one sided 97.5% CI 78.2%). None of the 6 people who definitely did not have coeliac disease had more than 1 coeliac code used, although three of them had had a gluten free prescription.

Table 3-4. Validation of the three definitions of a coeliac disease diagnosis in 32 people with additional information available

	One diagnostic code		One diagnostic code and at least one gluten free prescription		Two diagnostic codes	
Coeliac disease confirmed	Number	Percent	Number	Percent	Number	Percent
Definitely not	6	18.8	3	11.5	0	0
Definitely yes	21	65.6	19	73.1	13	86.7
Probably yes	5	15.6	4	15.4	2	13.3
Total	32	100	26	100	15	100

3.8 Overall study design for subsequent studies

Having defined the study populations in sections 3.4 and 3.7 the subsequent studies use this base population. In the studies on fracture risk, malignancy and mortality the whole population is used. In the study on vascular disease only those subjects over the age of 25 years are included. In general a cohort analysis has been performed as each individual potentially enters and exits prospective follow up (up to standard data) at different times. For the cohort studies as some subjects are censored, either at the end of the follow

up period (April 2002) or at the time they leave a contributing practice (at the end of up to standard data), survival analysis was used to estimate the rate of occurrence of events taking account of these unequal lengths of follow up. In addition survival analysis, in particular the Cox proportional hazard's model, allows the baseline hazard to vary over time (which it conceivably may do with the outcomes chosen) and therefore is appropriate (as long as the assumptions of such modelling are met). The specific analysis strategies are described in more detail for each separate study. Data manipulation and analyses were carried out using the software Access 2000 and Stata 7 (Stata corportation, Texas, USA).

4 Fracture risk in people with coeliac disease

4.1 Introduction

This section will describe a study of the fracture risk in people with coeliac disease compared to the general population.

4.2 Methods

4.2.1 Study population

I used the whole study population for this study as defined in section 3.4.

4.2.2 Outcomes and confounders

The main outcome measure was any fracture, and the observation time at risk was between the beginning of the up to standard record and the end of data collection for the overall and subgroup analyses. I also examined first hip and radius/ulna fracture by the same method, and then multiple fractures by extracting all diagnosis codes for any fracture. I extracted data for drug exposures (e.g. oral and injected corticosteroids) for the time period before outcome or the end of data collection and calculated a prescription rate for each drug (number of prescriptions/observation time). I categorised individuals into those who had none of the specific prescription, a few and many. The latter two categories were split at the median of the prescription rate for those who had one or more prescriptions. Potential drug confounders included steroids, antidepressants, bisphosphonates and hormone replacement therapy. A number of other potential confounders including recorded falls and other co-morbid diagnoses (e.g. chronic obstructive pulmonary disease) were extracted for the time period before outcome or the end of data collection.

4.3 Statistical analysis

Initially I calculated crude age and gender specific fracture rates for the two cohorts and then used Cox regression modelling to estimate the hazard ratio (HR) of fracture in the coeliac disease cohort compared to the matched control cohort, checking the proportional hazards assumption using the diagnostic section within Stata. Kaplan-Meier graphs were plotted for each

of the main outcome measures. For the multiple fractures analysis I used a conditional risk set model[147]. The impact of potential confounders was assessed using a series of multivariable models, retaining variables that led to a change in the hazard ratio for coeliac disease of 10% or more. For confounder variables, missing data were fitted as a separate category to ensure that nested models contained the same number of individuals. To assess possible interaction between coeliac disease and age group (age at the beginning of up to standard data and at diagnosis), gender and prevalent/incident status we performed stratified analyses and fitted multiplicative interaction terms. Finally we explored any change in fracture rates before and after diagnosis of coeliac disease in the “incident” subjects only, by comparing the rates for these periods adjusted for age group.

4.4 Results

The results for the study population are shown in section 3.7. The proportions of subjects with various potential confounding co- morbidities and drug prescription rates are shown in Table 4-1. There was no excess of recorded falls in the coeliac disease cohort and 2.2% had a diagnosis of chronic obstructive pulmonary disease compared to 2.0% of the control cohort.

Table 4-1. Details of selected potential confounders

	Coeliac disease cohort (n=4732)	%	Control cohort (n=23620)	%
Recorded number of falls				
None	4560	96.4	22815	96.6
One	137	2.9	676	2.9
More than one	35	0.7	129	0.6
Chronic obstructive pulmonary disease				
No	4626	97.8	23146	98.0
Yes	106	2.2	474	2.0
Diabetes (Type 1 and type 2)				
No	4595	97.1	23018	97.5
Yes	137	2.9	602	2.6
Recorded number of prescriptions for:				
Oral or injected steroids				
None	3956	83.6	20909	88.5
A few	316	6.7	1409	6.0
Many	460	9.7	1302	5.5
SSRIs				
None	4297	90.8	21966	93.0
A few	206	4.4	824	3.5
Many	229	4.8	827	3.5
Tricyclic antidepressants				
None	4012	84.8	20747	87.8
A few	338	7.1	1456	6.2
Many	382	8.1	1415	6.0
Bisphosphonates				
None	4629	97.8	23461	99.3
A few	47	1.0	80	0.3
Many	56	1.2	79	0.3
Hormone replacement therapy				
None	3924	82.9	20240	85.7
A few	337	7.1	1756	7.4
Many	471	10.0	1622	6.9

4.4.1 Numbers and rates of fracture

There were 356 recorded first fractures in the coeliac disease cohort, and 24 recorded hip fractures. The overall rate of any fracture for the coeliac disease cohort was 137.9 per 10000 person years compared to 105.9 per 10000 person years in the control cohort (Table 4-2). The crude rates of all fracture by age group, gender and disease status are shown in Figure 4-1. There was approximately a 30% increase in the risk of any fracture for the coeliac disease cohort compared to the control cohort (HR 1.30 95% CI 1.16-1.46)). When restricted to an analysis of hip fracture or ulna/radius fracture the hazard ratios were increased to 1.90 (95% CI 1.20-3.02) and 1.77 (95% CI 1.35-2.34) respectively.

In the multivariate analyses, none of the potential confounders I assessed made any substantial impact on the coefficients for the coeliac disease cohorts, so they were not included in the final models. The absolute difference in the rate of any fracture overall was 3.20 fractures per 1000 person years. For hip fracture, in those over 45 years of age, the rate difference was 0.97 per 1000 person years and in those over 75 years 2.35 per 1000 person years. Figure 4-2 shows the overall Kaplan-Meier plots for any fracture, hip and ulna/radius fracture by disease status.

4.4.2 Incident/prevalent cases

When I analysed the subjects with an “incident” diagnosis of coeliac disease I found slightly reduced hazard ratios in comparison with both the overall hazard ratios for all subjects and compared to those in the “prevalent” group. However I found no statistically significant evidence of interaction. The fracture rate after diagnosis of coeliac disease in those with an incident diagnosis was 145.2 per 10000 person years in comparison with 122.8 per 10000 person years before diagnosis (Table 4-3). When adjusted for age the hazard ratio for the period after diagnosis compared to before was 1.07 (95% CI 0.77-1.50).

4.4.3 Multiple fractures

Of the control cohort 183 (0.8%) had more than one fracture compared to 50 (1.0%) of the coeliac disease cohort. The hazard ratio estimated for risk of multiple fractures within each individual showed no substantial difference from the overall findings for any fracture (HR 1.28 (95% CI 1.15-1.42)). In addition there was no interaction between age group and coeliac disease status with respect to multiple fracture.

4.4.4 Restriction analyses

When I repeated the analyses restricted to only those subjects with coeliac disease who had had at least one gluten-free prescription I found no important differences in the risk estimates (Table 4-4). Nor were there any differences when I restricted to including only cases with at least two coeliac disease medical codes.

Figure 4-1. Crude rates (95% CI) of any fracture by age group, gender and disease status

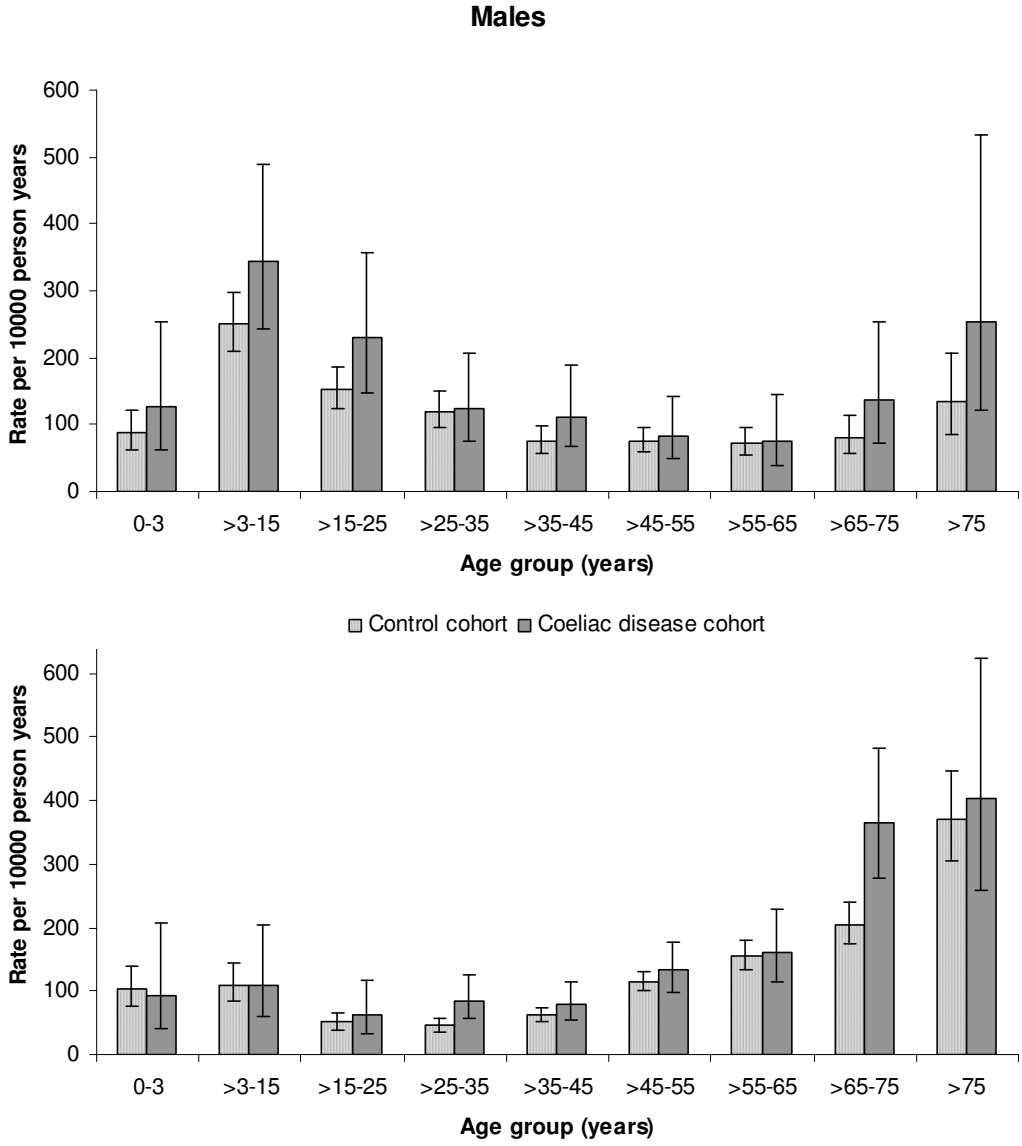


Table 4-2. Number, rate and crude hazard ratios for the coeliac disease cohort compared to the control cohort, overall and limited to prevalent and incident subjects with coeliac disease

	Overall					Prevalent subjects with coeliac disease$				Incident subjects with coeliac disease$			
	N**	Number of fractures	Rate ($10000yr^{-1}$)	Hazard ratio	95% CI	N**	Rate ($10000yr^{-1}$)	Hazard ratio	95% CI	N**	Rate ($10000yr^{-1}$)	Hazard ratio	95%CI
Any fracture													
Control cohort*	23616	1524	105.9	1		15671	103.8	1		7945	109.3	1	
Coeliac disease cohort	4732	356	137.9	1.30	(1.16-1.46)	3143	144.2	1.40	(1.20-1.62)	1589	129.5	1.19	(0.99-1.42)
Hip fracture													
Control cohort*	23620	71	4.7	1		15675	4.6	1		7945	4.9	1	
Coeliac disease cohort	4732	24	8.9	1.90	(1.20-3.02)	3143	11.1	2.41	(1.37-4.23)	1589	6.0	1.23	(0.54-2.81)
Ulna/radius fracture													
Control cohort*	23619	210	14.1	1		15674	12.2	1		7945	17.3	1	
Coeliac disease cohort	4732	67	24.9	1.77	(1.35-2.34)	3143	25.5	2.13	(1.48-3.07)	1589	24.2	1.40	(0.92-2.14)

* baseline category

** total numbers vary as those individuals who had a fracture on the same date as the start of their GPRD record began were excluded

$ for these analyses only the matched controls of those subjects with coeliac disease included were used

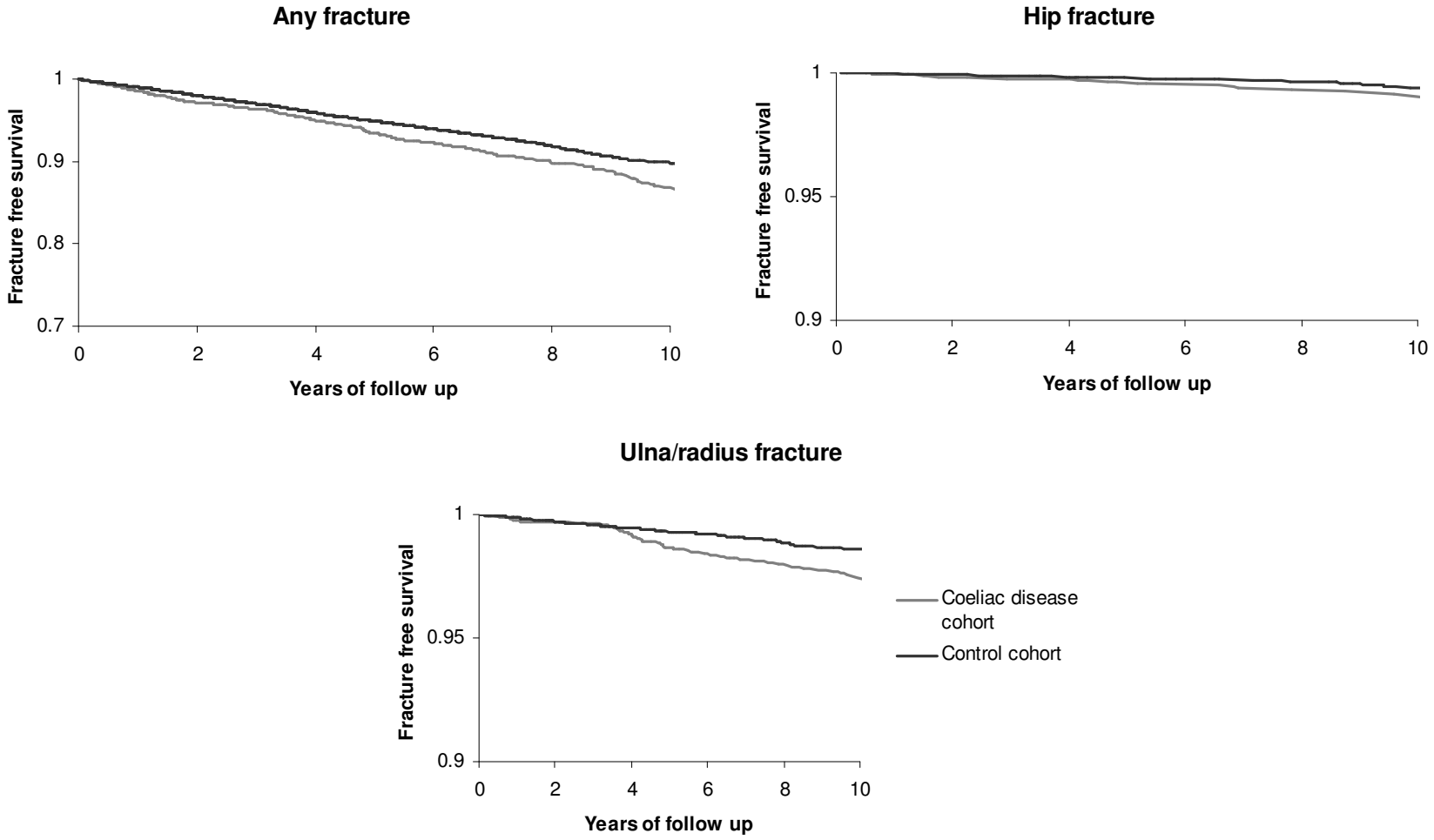


Figure 4-2. Kaplen-Meier survival plots for any fracture, hip fracture and ulna/radius fracture

Table 4-3. Number, rate and hazard ratio of any fracture before and after diagnosis in 1589 subjects with an incident diagnosis of coeliac disease

	Number of fractures	Person years	Rate ($10000yr^{-1}$)	95% CI		Hazard ratio**	95% CI
Before diagnosis*	87	7082	122.8	99.6	151.6	1	
After diagnosis	61	4201	145.2	113.0	186.6	1.07	(0.77-1.50)

* baseline category

** adjusted for age at start of observation time

Table 4-4. Any fracture, hip fracture and ulna/radius fracture analyses restricted to those cases with 1 coeliac code plus at least one gluten free prescription and to those cases with 2 coeliac codes

	Subjects with 1 coeliac code plus at least one gluten free prescription (n=4280)$		Subjects with 2 coeliac codes (n=2761)$	
	Hazard ratio	95% CI	Hazard ratio	95% CI
Any fracture				
Control cohort*	1		1	
Coeliac disease cohort	1.31	(1.16-1.48)	1.30	(1.12-1.50)
Hip fracture				
Control cohort*	1		1	
Coeliac disease cohort	1.99	(1.25-3.18)	1.97	(1.07-3.64)
Ulna/radius fracture				
Control cohort*	1		1	
Coeliac disease cohort	1.84	(1.39-2.44)	1.81	(1.35-2.34)

* baseline category

$ for these analyses only the matched controls

of those subjects with coeliac disease included were used

4.4.5 Proportional hazard's assumption

In the analysis of hip fracture the ph test for evidence against the proportional hazard's assumption gave $p<0.9$. The log –log plot for the same analysis is shown in Figure 4-3. In the plot of Schoenfeld residuals there was only small variation in the hazard over time seen (Figure 4-4). As with the hip fracture analysis there was little evidence against the proportional hazard's assumption in any of the other presented models.

Figure 4-3. Log -log plots of hip fracture analysis

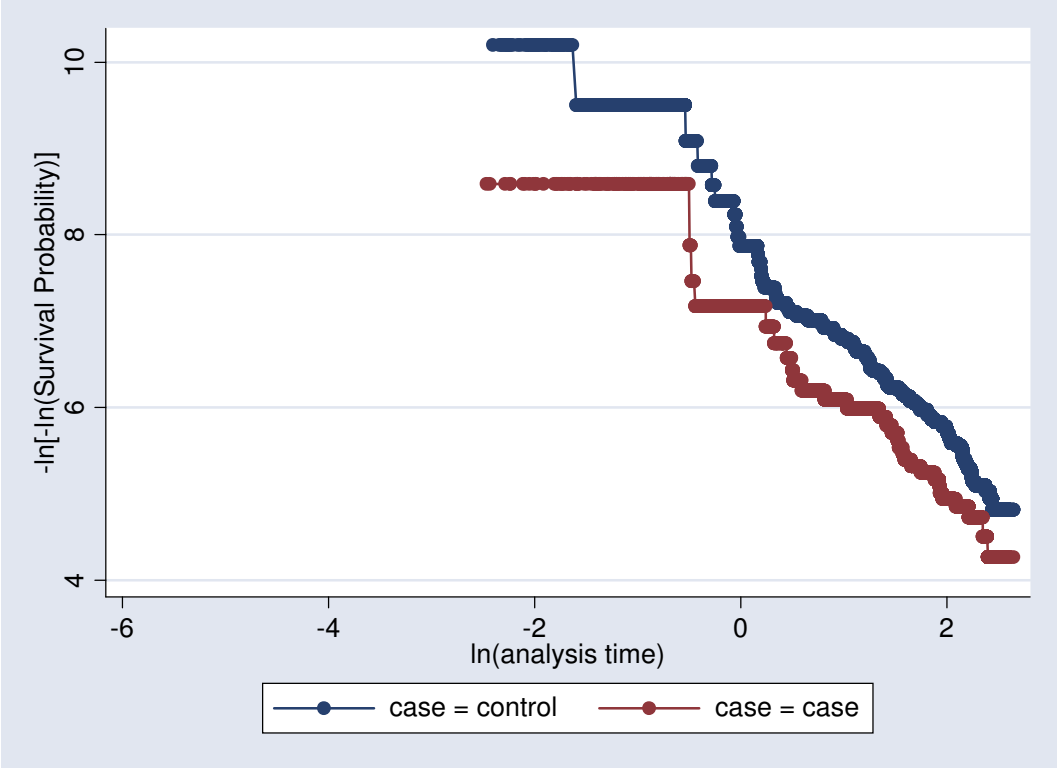
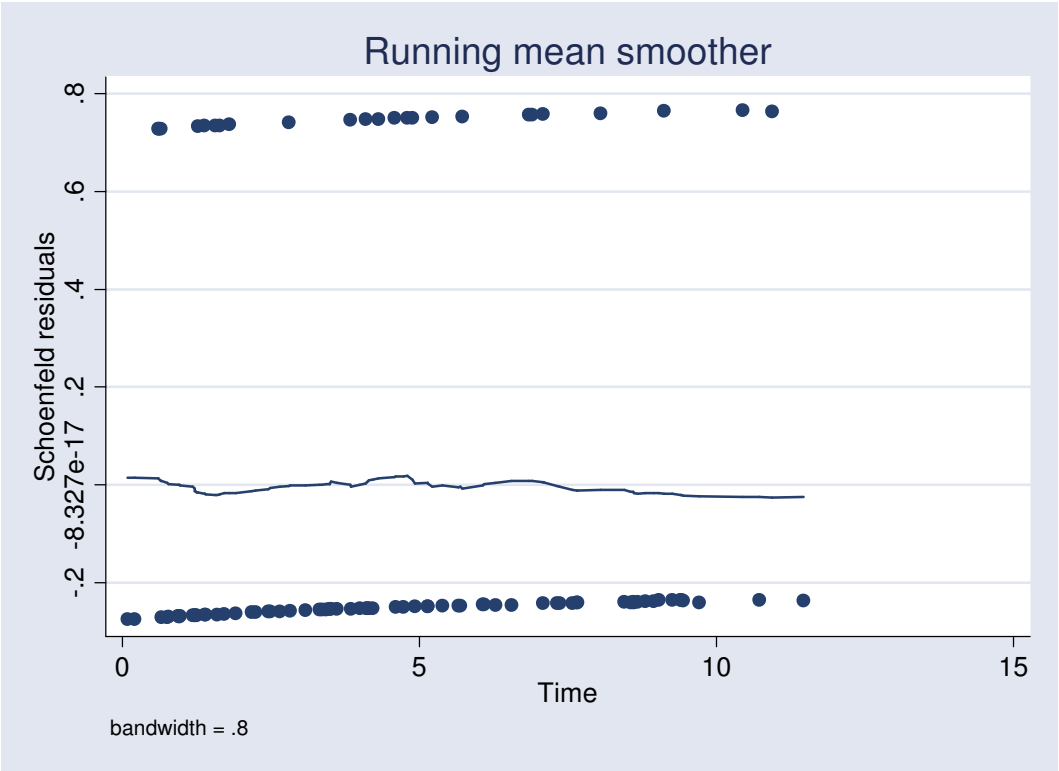


Figure 4-4. Plot of schoenfeld residuals against time for the hip fracture analysis



4.5 Discussion

4.5.1 Principal findings

The results of the study show modest increases in the relative risk of any fracture (30% increase), hip fracture (90% increase) and ulna/radius fracture (77% increase) among people with coeliac disease compared to the general population. Nonetheless, the increases in absolute risk of fracture were modest being 3.19 fractures per 1000 person years and 0.97 per 1000 person years in those over age 45 for hip fracture alone. These increases in risk were slightly less in individuals with coeliac disease diagnosed more recently. Additionally, in the subjects with "incident" coeliac disease, I found no difference in the risk of fracture in the period after diagnosis compared to before diagnosis (HR 1.07 (95% CI 0.77-1.50)).

4.5.2 Limitations and merits

In the study people with coeliac disease were more frequent attenders at their general practitioner than members of the general population. It is therefore possible that I have overestimated the relative rate of any fracture in the coeliac disease cohort compared to the control cohort as a consequence of less complete recording of medical events in a "healthy" group. However this is unlikely to have been a problem with hip fracture, where ascertainment is likely to be high. In contrast, the risks of fracture may be an underestimate due to random error in the definition of both coeliac disease status and fracture diagnoses. However, the accuracy of medical diagnoses within the GPRD is known to be high, specifically with respect to fracture[145] and the crude age and gender specific rates for any fracture are similar to those reported by others in the United Kingdom[148]. Furthermore, there have been several validation studies of the GPRD, including one that has looked in detail at the accuracy of the diagnosis of inflammatory bowel disease (IBD)[137, 138, 144]. In this study Lewis et al found that the IBD diagnosis within the GPRD was highly probable or probable in 92% (95% CI 86 to 96%) of their surveyed cases. IBD is

analogous to coeliac disease in that it is a diagnosis made in secondary care and their findings are likely to be generalisable to the study.

In addition, I believe that there will be very few people with a recorded diagnosis of coeliac disease and coexisting prescriptions for one or more gluten free products who do not have coeliac disease. One reason for this is that general practitioners in the UK have a limited prescribing budget, which means that they are unlikely to write an unnecessary and expensive prescription, such as for a gluten free product, unless they have good reason. When I restricted the analyses to those people with coeliac disease who also had at least one gluten free prescription, to increase the specificity of the coeliac disease diagnosis, there was no substantial change in the effect. When I was even more restrictive in the coeliac disease definition, by including only people with at least two diagnostic codes for coeliac disease, the estimates were again very similar to the overall findings.

The definition of an incident case is a pragmatic one based on the assumption that an important medical diagnosis such as coeliac disease should be coded accurately in the historical records or within a year of the person registering with a GPRD practice and it is similar to the “incident” definition of IBD that was validated by Lewis et al. It is possible that a person with coeliac disease might not see their GP for 5 years or more and therefore be incorrectly categorised as “incident”, however I think the numbers are likely to be small. Should there be misclassification between prevalent and incident cases, this may have led to an underestimate of the differences in fracture risk between the prevalent and incident groups. The finding of a lack a gluten free prescription in 10% of those people recorded as having coeliac disease might suggest that they have been incorrectly labelled. I think it is as likely that either they purchase their gluten free products over the counter, or that they have mild disease and do not comply with a gluten free diet.

As the GPRD does not contain information on socio-economic status I have been unable to control for this in the analysis and since people with coeliac disease have tended to be in higher socio-economic groups (section 1), this could have led to an underestimate of fracture risk. This would only be the case if there is a strong link between socio-economic status and fracture risk for which there is not strong evidence. The control cohort was closely matched in terms of age, gender and community to minimise the potential for confounding by these factors and I had the ability to assess the impact of potential confounders such as BMI, smoking status, co- morbidity and drug exposures on fracture risk. In the event I found no evidence of substantial confounding, although I acknowledge the presence of missing data for some of these variables. It seems unlikely that I have therefore greatly under or over estimated the fracture risk due to any residual confounding.

The study included 4732 people with coeliac disease who contributed more than 27000 person years at risk and had over 350 fractures. It is the largest study of fracture risk in coeliac disease published to date. As a consequence of its size and the cohort design I have been able to estimate, with reasonable precision, both the absolute and relative risks of any fracture and also for specific fracture subgroups. I have also assessed multiple fractures and explored the effect of treatment on fracture risk. I believe that the results are therefore likely to be generalisable to people with diagnosed coeliac disease elsewhere and can be considered to reflect contemporary risk of fracture. When I took multiple fractures into account in the analysis there was no substantial change in the risk estimate nor were older people with coeliac disease more at risk of multiple fractures. This suggests that I have not underestimated the relative incidence of fracture in people with coeliac disease by restricting the analyses to their first recorded fracture occurrence. Many studies have shown that the clinical presentation of diagnosed coeliac disease has changed in recent times compared to that seen in the 1960's and 70's[50-53, 98]. The finding of slightly greater relative risks in the prevalent group may reflect that change i.e. that those people diagnosed with coeliac disease in recent times have not had such

severe malabsorption and consequent malnutrition prior to diagnosis, and therefore have less risk of osteoporosis and fracture. Alternatively it may be that prevalent cases with a longer duration of treated disease have an increased fracture risk, although there is no evidence that the rate of decrease of bone mineral density in coeliac disease whilst on a gluten free diet is greater than the general population.

4.5.3 Comparison with other studies

The results can be compared with the findings of previous studies of fracture risk in coeliac disease. Vasquez et al compared the fracture experience of 165 patients (median age 40 years) with that of controls with functional gastrointestinal disorders and found a three-fold increase in overall fracture risk (Odds ratio OR 3.5 95% CI 1.8-7.2) based on 25% of patients reporting fractures and only 8% of controls[112]. A subsequent study from the same group found that the increase in fracture experience was confined to people with coeliac disease presenting with “classical malabsorption” (OR 5.2 95% CI 2.8-9.8)[115]. Fickling et al reported a “relative risk” of fracture of 7 based on a survey of 75 patients with a mean age of 52 years and age and sex matched controls selected from patients who had attended for bone densitometry[114]. While 21% of their patients with coeliac disease reported a past fracture, only two (3%) of the controls did. The differences between the findings and these two studies are likely to be due to a combination of the over-representation of more severe disease in their subjects with coeliac disease, and low fracture rates in the control groups. The results are more in keeping with two recent population based studies of fracture risk in coeliac disease. Both Vestergaard et al, in a database study of 1021 hospital diagnosed subjects with coeliac disease, and Thomason et al, using a mailed questionnaire survey of 244 cases, found no increase in the overall fracture risk compared to the general population but with wide confidence limits[111, 113].

4.5.4 Summary

The findings confirm that overall people with diagnosed coeliac disease have a small increased risk of fracture and that the excess risk was lower in those diagnosed more recently. The risks of “osteoporotic” fracture such as hip

and ulna/radius are higher than the overall risk but are, at most, moderate. Although the results do not relate to people with previously undetected coeliac disease it seems unlikely that the fracture risks in this group will be substantially greater than I have found for clinically diagnosed disease, as I have previously shown that the risk of osteoporosis in the former group is small[149]. Some groups have suggested that all newly diagnosed adults with coeliac disease should be screened for osteoporosis, either at diagnosis or following one year of treatment with a gluten free diet[43, 109, 110]. More data are needed on the safety, efficacy and cost-effectiveness of such screening programs in coeliac disease before they are universally recommended.

5 Risk of vascular disease in adults with coeliac disease

5.1 Introduction

This section will describe a study of the risk of hypertension, high cholesterol, atrial fibrillation, myocardial infarction and stroke in adults with coeliac disease compared to the general population.

5.2 Methods

I used two approaches in this study: firstly I compared the risk of "persistent" conditions (hypertension, high cholesterol and atrial fibrillation) in coeliac disease compared to the general population using all available GPRD data (cross-sectional design), and secondly I used a historical matched cohort study design for acute myocardial infarction and stroke where the observation time (person years at risk) for individuals included in the study started at the beginning of their up to standard GPRD record.

5.2.1 Study population

I used the study population as defined in section 3.4. Selection for this study was restricted to those people with coeliac disease aged 25 years or over at the beginning of their up-to-standard data period, and their matched controls. The age cut off was chosen arbitrarily to limit the study population to adults.

5.2.2 Outcomes and confounders

I investigated the risk of a diagnosis of hypertension, high cholesterol or atrial fibrillation at any time in the available data and rate of first myocardial infarction or stroke during up to standard GPRD data for the cohort analysis. In addition I calculated a composite measure of hypertension that was positive only if subjects had both a diagnosis of hypertension and had ever had a prescription for an anti-hypertensive medication. Similarly I calculated a composite measure of hypercholesterolaemia that was positive if subjects had both a diagnosis of high cholesterol and had a prescription for a lipid-lowering medication. Potential confounders including the recorded presence or absence of diabetes and other co-morbid diagnoses (e.g. thyroid disease) were extracted from all the available data (including data not in the up to standard period).

5.3 Statistical analysis

To compare the risk of each “persistent” condition in adults with coeliac disease to the control population I calculated odds ratios for ever having a diagnosis of hypertension, high cholesterol and atrial fibrillation using conditional logistic regression. For the cohort analysis I calculated crude age and gender specific myocardial infarction and stroke rates for the two cohorts and then used Cox regression modelling to estimate the hazard ratio (HR) of myocardial infarction or stroke in the coeliac disease cohort compared to the matched control cohort. I plotted Kaplan-Meier graphs and checked the proportional hazards assumption of the models. The impact of potential confounders was assessed using a series of bivariable models, retaining variables that led to a change in the hazard or odds ratios for coeliac disease of 10% or more. For confounder variables, missing data were fitted as a separate category to ensure that nested models contained the same number of individuals. I checked for any evidence of interaction between disease status and both body mass index and prevalent/incident status by fitting multiplicative interaction terms.

5.4 Results

The study included 3590 subjects with coeliac disease and 17925 matched controls contributing 21248 and 117210 observed years at risk respectively. The groups were closely matched on age at the start of the GPRD record and gender (Table 5-1). There were more current smokers in the control cohort (16.9% vs 13.7%). There was no excess of recorded diabetes in the coeliac disease group (3.5% vs 3.7%) and 7.0% had a diagnosis of thyroid disease compared to 3.2% of controls. Mean systolic blood pressure was 5 mmHg lower in the coeliac disease group.

Table 5-1. Characteristics of the adults with coeliac disease and their matched controls

	Coeliac disease (n=3590)	%	Controls (n=17925)	%
Median observed time (years)	5.9		6.6	
Total observed time (years)	21248		117210	
Female	2461	68.6	12285	68.5
Age groups at start of GPRD record (years)				
=<35	779	21.7	4085	22.8
>35-45	809	22.5	4066	22.7
>45-55	833	23.2	4113	23.0
>55-65	550	15.3	2707	15.1
>65-75	415	11.6	1994	11.1
>75	204	5.7	960	5.4
Smoking status				
Non smoker	1818	50.6	7553	42.1
Ex-smoker	203	5.7	1156	6.5
Current smoker	490	13.7	3029	16.9
Unknown	1079	30.1	6187	34.5
Body mass index, kgm^{-2}				
Less than or equal to 18.5	155	4.3	202	1.1
18.51 to 25	1746	48.6	5923	33.0
25.01 to 30	608	16.9	4057	22.6
Greater than 30	140	3.9	1808	10.1
Unknown	941	26.2	5935	33.1
Diabetes ever recorded				
No	3463	96.5	17258	96.3
Yes	127	3.5	667	3.7
Thyroid disease ever recorded				
No	3340	93.0	17358	96.8

Yes	250	7.0	567	3.2
Systolic blood pressure, mmHg (mean, sd)	129.5	19.4	134.6	20.0
Diastolic blood pressure, mmHg (mean, sd)	77.0	9.5	80.0	9.9
Blood pressure not recorded	821	22.9	4701	26.2

5.4.1 Numbers, proportions and risks of hypertension, high cholesterol and atrial fibrillation

Overall 408 (11%) of the adults with coeliac disease had ever had a diagnosis of hypertension compared to 2765 (15%) in the control group giving an unadjusted odds ratio (OR) of 0.68 (95% CI 0.60 to 0.76) as shown in Table 5-2. This relationship was partly explained by body mass index, as after adjusting for this variable the odds ratio was 0.78 (95% CI 0.69 to 0.87). When I repeated the analyses using the composite measure of diagnosed hypertension (including prescriptions), the results were similar. When I repeated the analyses using the composite measure of hypercholesterolaemia the results were again similar. The unadjusted odds ratios for atrial fibrillation and high cholesterol diagnoses were 1.26 (95% CI 0.97 to 1.64) and 0.58 (95% CI 0.47 to 0.72) respectively. Adjusting for body mass index had no appreciable effect on these latter two estimates.

Table 5-2. Analysis of risk of hypertension, atrial fibrillation and high cholesterol diagnoses in 3590 people with coeliac disease compared to 17925 controls using conditional logistic regression

	Overall number with disease (%)	Unadjusted odds ratio	95% CI	Adjusted** odds ratio	95% CI
Hypertension					
Control*	2765 (15.4)	1		1	
Coeliac disease	408 (11.4)	0.68	(0.60 to 0.76)	0.78	(0.69 to 0.87)
Atrial fibrillation					
Control*	305 (1.7)	1		1	
Coeliac disease	76 (2.1)	1.26	(0.97 to 1.64)	1.28	(0.98 to 1.67)
High cholesterol					
Control*	866 (4.8)	1		1	
Coeliac disease	107 (3.0)	0.58	(0.47 to 0.72)	0.60	(0.49 to 0.74)

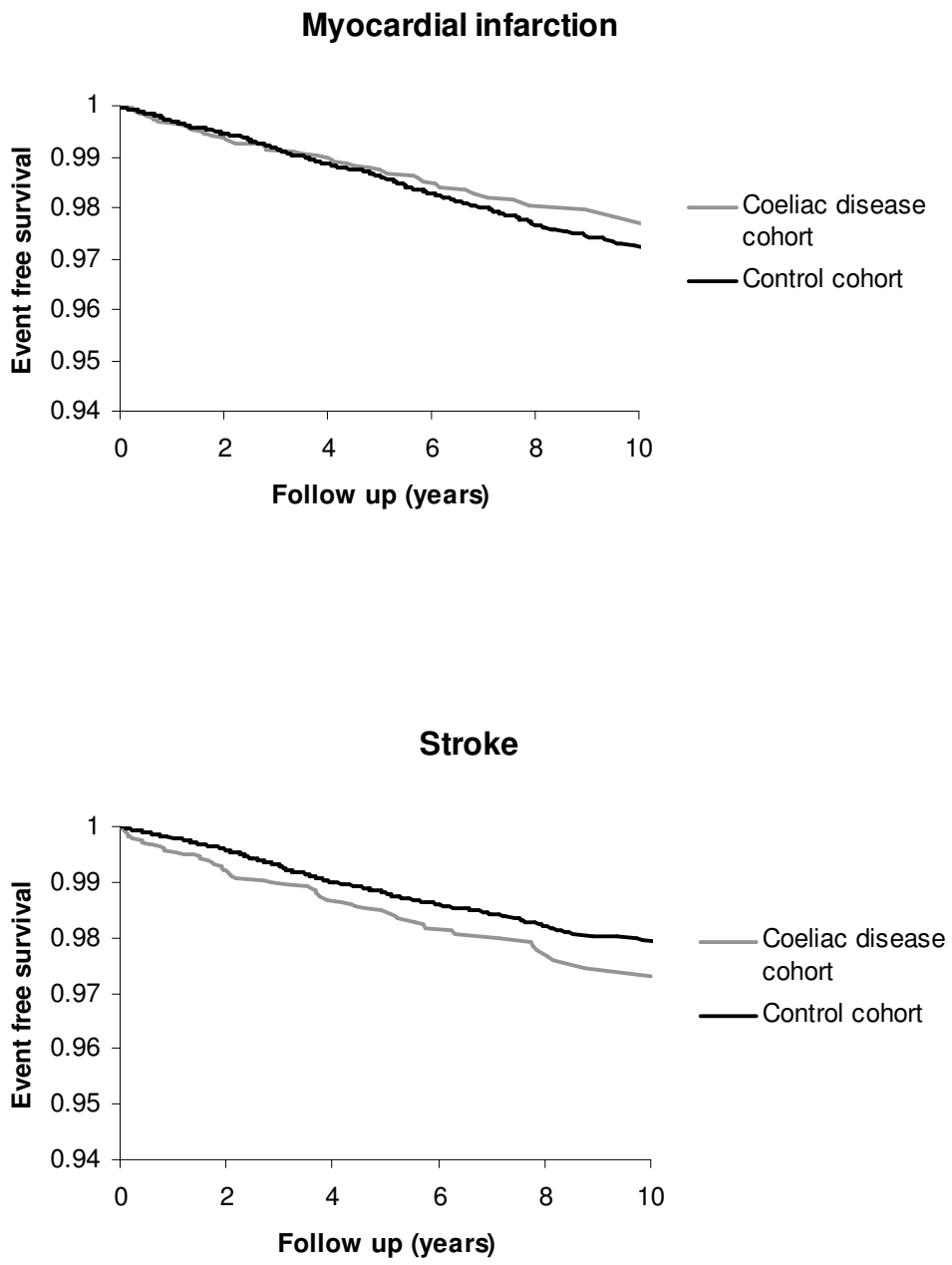
* baseline category

** adjusted for body mass index

5.4.2 Numbers and rates of myocardial infarction and stroke

There were 52 recorded first myocardial infarctions in the coeliac disease cohort, and 62 recorded first strokes. The overall rate of myocardial infarction for the coeliac disease cohort was 24.7 per 10000 person years compared to 29.2 per 10000 person years in the control cohort (Table 5-3). There was approximately a 15% decrease in the risk of myocardial infarction for the coeliac disease cohort compared to the control cohort (HR 0.85 (95% CI 0.63-1.13)). The overall rate of stroke was 29.4 per 10000 person years in the coeliac disease group and the hazard ratio was 1.29 (95% CI 0.98-1.70). In the multivariate analyses, only body mass index of the potential confounders I assessed made any substantial impact on the coefficient for myocardial infarction in coeliac disease altering the overall hazard ratio to 0.95 (95% CI 0.71 to 1.27). For stroke, only the presence or absence of diagnosed hypertension altered the hazard ratio appreciably, to 1.40 (95% CI 1.06 to 1.84). In contrast, when adjusted for smoking status, the hazard ratios were 0.87 (95% CI 0.65-1.16) and 1.33 (95% CI 1.00-1.75) respectively. Figure 5-1 shows the Kaplan-Meier plots for myocardial infarction and stroke (crude analysis).

Figure 5-1. Kaplan-Meier survival plots for myocardial infarction and stroke



5.4.3 Incident/prevalent cases

When I analysed only the subjects with an “incident” diagnosis of coeliac disease I found a slightly reduced hazard ratio for myocardial infarction (HR 0.75 (95% CI 0.46-1.23)) and a slightly increased hazard ratio for stroke (HR 1.60 (95% CI 0.99-2.59)) in comparison with both the overall hazard ratios for all subjects and compared to those in the “prevalent” group. However we found no statistically significant evidence of interaction.

Table 5-3. Number, rate and crude hazard ratios for the coeliac disease cohort compared to the control cohort, overall and limited to prevalent and incident subjects with coeliac disease

	Overall					Prevalent subjects with coeliac disease$				Incident subjects with coeliac disease$			
	N	Number of events	Rate (10000yr^{-1})	Hazard ratio	95% CI	N	Rate (10000yr^{-1})	Hazard ratio	95% CI	N	Rate (10000yr^{-1})	Hazard ratio	95%CI
Myocardial infarction													
Control cohort*	17925	339	29.2	1		11775	31.3	1		6150	25.8	1	
Coeliac disease cohort	3590	52	24.7	0.85	(0.63-1.13)	2360	28.8	0.92	(0.64-1.32)	1230	19.5	0.75	(0.46-1.23)
Stroke													
Control cohort*	17925	265	22.7	1		11775	27.9	1		6150	14.8	1	
Coeliac disease cohort	3590	62	29.4	1.29	(0.98-1.70)	2360	33.7	1.19	(0.85-1.68)	1230	23.8	1.60	(0.99-2.59)

* baseline category

$ for these analyses only the matched controls of those subjects with coeliac disease included were used

5.4.4 Restriction analyses

When I repeated the analyses restricted to only those adults with coeliac disease who had at least one gluten free prescription all the results were similar. When I only included those cases with two or more diagnostic codes for coeliac disease the majority of results were slightly different, although all within the previous confidence intervals. In the latter case the odds ratios for hypertension, high cholesterol and atrial fibrillation were 0.67, 0.58, 1.07. For myocardial infarction and stroke the hazard ratios were 0.99 and 1.23 respectively.

5.4.5 Proportional hazard's assumption

There was no clear evidence against the proportional hazard's assumption in this study. The log –log plots and schoenfeld residual plots for the myocardial infarction analysis are shown in Figure 5-2 and Figure 5-3 respectively.

Figure 5-2. Log -log plot for myocardial infarction analysis

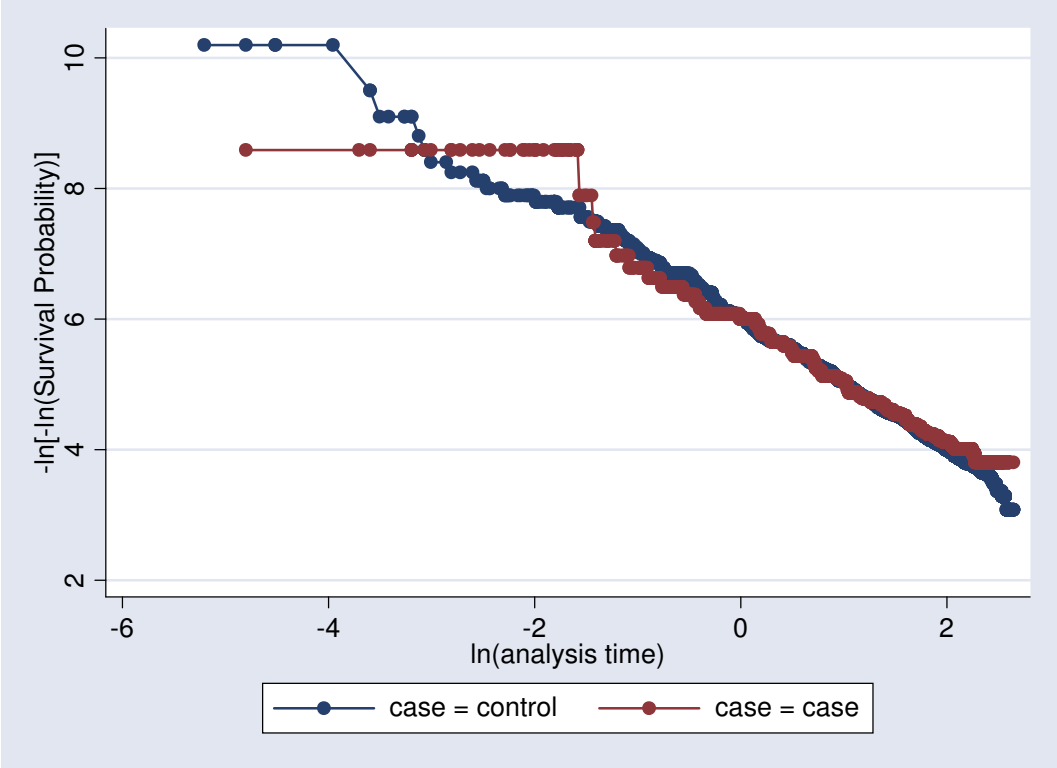
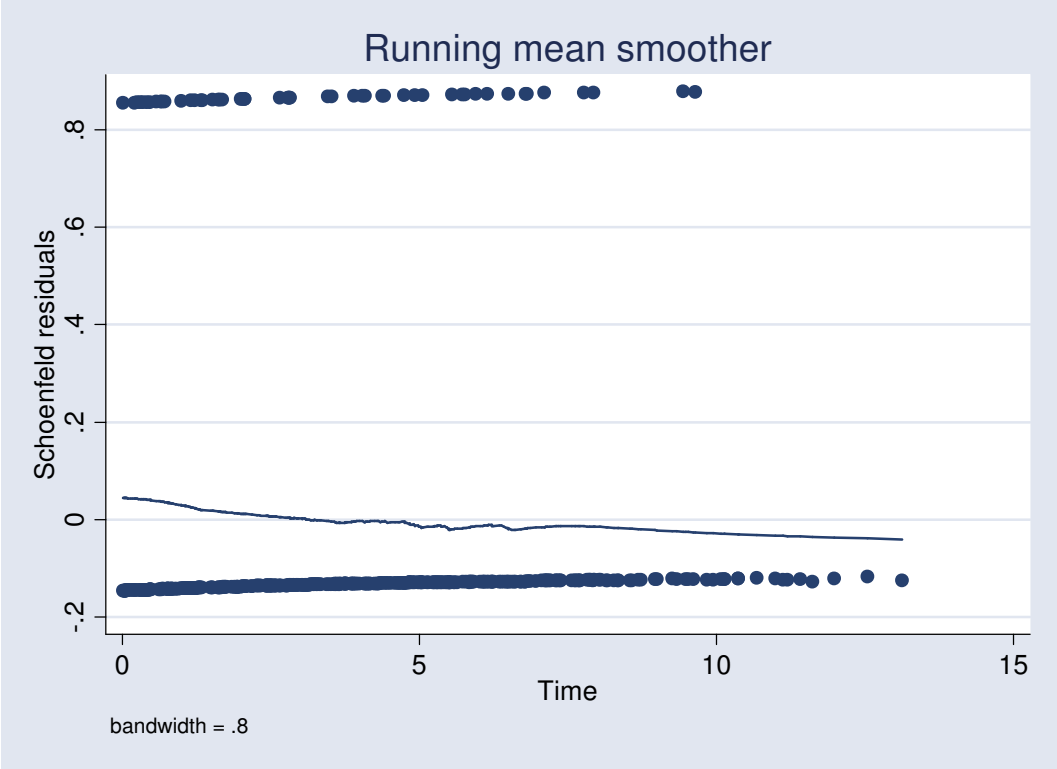


Figure 5-3. Plot of schoenfeld residuals against time for myocardial infarction analysis



5.5 Discussion

5.5.1 Principal findings

The results of the study show a marked decrease in the risk of diagnosed hypertension and high cholesterol, yet a slight increase in the risk of atrial fibrillation in adults with coeliac disease compared to the general population. These factors appear to have slightly different effects on vascular disease, as the point estimates suggest that there is approximately a 15% reduction (95% CI 37% reduction to 16% increase) in the risk of myocardial infarction but approximately a 30% increase (95% CI 2% reduction to 70% increase) in the risk of stroke. The contrasts in these findings were slightly greater in individuals with coeliac disease diagnosed more recently.

5.5.2 Limitations and merits

Adults with coeliac disease are more frequent attenders at their general practitioner than members of the general population (section 3.7.8). As a consequence, if ascertainment bias is present, it is possible that I have underestimated the risk of stroke and atrial fibrillation in the control cohort due to less complete recording of medical events in a "healthy" group. Similarly I may have underestimated the decrease in relative risk of hypertension, high cholesterol and myocardial infarction. It is also possible that I have underestimated the effect sizes through misclassification of unrecorded coeliac disease i.e. that some of the control group may have coeliac disease. To check the validity of my diagnostic definitions I repeated the analyses with the more restrictive definitions described earlier. When I restricted the analyses to those adults with coeliac disease who also had at least one gluten free prescription, to increase the specificity of the coeliac disease diagnosis, there was no substantial change in the effect estimates. When I included only cases who had at least two diagnostic codes for coeliac disease, the majority of the findings were similar to those found overall.

As the GPRD does not contain information on socio-economic status I have, as before, been unable to control for this important variable in the analyses

and since adults with coeliac disease have tended to be in higher socio-economic groups, this could explain part of their marginally lower incidence of cardiovascular disease (section 1). The control cohort was closely matched in terms of age, gender and community to minimise the potential for confounding by these factors and I had the ability to assess the impact of potential confounders such as body mass index, smoking status and co-morbidity on vascular disease risk. Although adjusting for body mass index and the diagnosis of hypertension altered the effect estimates for myocardial infarction and stroke respectively, both can be considered as intermediate steps between coeliac disease and the outcome. Rather than being alternative explanations for the observed association, they are more likely to be on the causal pathway between coeliac disease and vascular disease. I found no other evidence of substantial confounding.

5.5.3 Interpretation

Although not statistically significant at the 5% level, the point estimates for the risks of stroke and myocardial infarction are intriguing. Based on my findings in section 1 where people with undetected coeliac disease had lower serum cholesterol and slightly lower blood pressure a lower risk for myocardial infarction might be expected. However the reduction in risk was not as great as might have been predicted by the earlier study (i.e. a 10% reduction in cholesterol leading to a 25% reduction in the incidence of myocardial infarction). It is possible, for example, that the observed relationship with myocardial infarction may have been attenuated by the effect of treatment with a gluten free diet. Although I have previously shown that people with undetected coeliac disease have lower serum cholesterol than the general population (section 1) it is possible that following treatment, and the subsequent improvement in intestinal absorption, serum cholesterol may increase therefore attenuating any protective effect. The relationship is clearly complex in view of the finding that plasma homocysteine levels remain high in people with treated coeliac disease even after many years of gluten exclusion[150] which may counteract any beneficial effect of lower cholesterol[151]. The analysis of incident and prevalent cases adds some support to the idea that treatment with a gluten free diet may alter the

vascular disease risk profile of these individuals. The finding of an increased risk of stroke was not explained by the greater prevalence of thyroid disorders or atrial fibrillation in the coeliac disease group, although it is possible that the latter condition is under recorded in primary care as only 13% of the subjects who had had a stroke in the study had atrial fibrillation. With less hypertension, lower body mass index and presumably lower cholesterol, it was surprising to find, albeit modest, contrasting relationship to that with myocardial infarction. One explanation may be the previously documented finding of a high prevalence of coeliac disease in idiopathic cardiomyopathy[152], suggesting that the mechanism of increased risk in coeliac disease might be arrythmogenic or thromboembolic. Alternatively the recently postulated neurotoxic effects of gluten may play an important role[153].

5.5.4 Summary

The findings confirm the hypothesis that adults with diagnosed coeliac disease have a decreased risk of hypertension and hypercholesterolaemia. The finding of a slightly decreased risk of myocardial infarction and a small increased risk of stroke in coeliac disease are intriguing and lead to speculation about the mechanisms of vascular disease, particularly in relation to nutritional status.

6 Malignancy and mortality in people with coeliac disease

6.1 Introduction

This section describes a study of the risk of malignancy and mortality in people with coeliac disease compared to the general population. This section also includes an estimate of life expectancy in people with coeliac disease compared to the general population.

6.2 Methods

6.2.1 Study population

I used the study population already described in sections 3.4 and 3.7.

6.2.2 Outcome data

For the outcomes, I extracted data that included the date of first occurrence of any malignancy, first occurrence of specific malignancy subgroups and date of death. I defined all malignancies by using the relevant codes in the GPRD database mapped to International Classification of Diseases (ICD) 9 codes[132]: 140-208 and 230-234. The specific malignancy groups I chose were: all gastrointestinal cancer (ICD 9 150-154), lung cancer (ICD 9 162-163), breast cancer (ICD 9 174-175), prostate cancer (ICD 9 185) and lymphoproliferative disease (ICD 9 200-202). To identify death I used a combination of Oxmis and Read coding and the subject's registration status within GPRD to assess whether they had died. Where a subject was multiply recorded as having died I used the earliest recorded date to define the date of death. The origin of the time axis and the entry of the subject to the study were both set as the index date (the matched case's relevant date was used for controls).

6.3 Statistical analysis

Initially I calculated crude cancer incidence and mortality rates for the coeliac disease and control cohorts. I used Cox regression modelling to estimate the hazard ratio (HR) comparing outcomes in the coeliac disease cohort compared to the control cohort plotting appropriate Kaplan-Meier graphs. I

checked the proportional hazards assumption of each model. The possible confounding effects of body mass index and smoking status were assessed using a series of multivariable models, retaining variables that led to a change in the hazard ratio for coeliac disease of 10% or more. For confounder variables, missing data were fitted as a separate category to ensure that nested models contained the same number of individuals. To assess possible interaction between coeliac disease status and age group or gender I performed stratified analyses and fitted multiplicative interaction terms as appropriate.

To assess the robustness of the initial analyses, I performed a series of sensitivity analyses. To assess ascertainment bias, i.e. whether any increase in cancer risk was related to increased investigation as a result of having either a diagnosis of coeliac disease or cancer made, I examined the hazard ratios for each outcome within the first year after diagnosis and during subsequent follow up. To assess the validity of the findings with respect to possible misclassification of coeliac disease status, I restricted the analyses to only those subjects with coeliac disease who had had at least one prescription for a gluten-free product. I also restricted to those with two codes for coeliac disease. To assess the possibility of survival bias, as the cohort contained prevalent cases of coeliac disease, I stratified the censored analysis by prevalent/incident status.

6.3.1 Indirect standardisation analysis

As the majority of previous mortality studies of coeliac disease have used population data as their comparison I additionally compared the mortality experience of both the coeliac cohort and the cohort control to that of the population of England and Wales. To do this I carried out an age, sex and period indirect standardisation of each cohort to where the expected deaths were estimated from the population of England and Wales. This analysis was truncated at 1/01/2000 as that was the extent of the available data.

6.3.2 Life table analysis

To further aid interpretation of any mortality associated with coeliac disease I additionally used a life table analysis. To determine the life expectancy of

people with coeliac disease compared to the general population I constructed life tables from the age-specific mortality rates.

6.4 Results

The cohorts included 4732 people with coeliac disease and 23620 matched controls contributing 18923 and 94323 person years at risk respectively (Table 6-1). There was less person time at risk in this study (compared to earlier sections in this thesis (3.7.1 and 3.4)) as subjects were entered at index date.

Table 6-1. Description of study population

	Coeliac disease cohort		Control cohort	
	(n=4732)	%	(n=23620)	%
Median observed time (years)	3.4		3.5	
Total observed time (years)	18923		94323	
Female	3095	65.4	18545	65.4
Age groups at entry to follow up (years)				
0-3	196	4.1	985	4.2
>3-15	385	8.1	1915	8.1
>15-25	455	9.6	2270	9.6
>25-35	676	14.3	3375	14.3
>35-45	812	17.2	4045	17.1
>45-55	858	18.1	4285	18.1
>55-65	610	12.9	3050	12.9
>65-75	469	9.9	2342	9.9
>75	271	5.7	1353	5.7

6.4.1 Malignancy

Among people with coeliac disease 134 had at least one malignancy. The overall rate of any malignancy for the coeliac disease cohort was 72·0 per 10000 person years compared to 55·9 per 10000 person years in the control cohort (Table 6-2), giving approximately a 30% increase in the risk of any malignancy among people with coeliac disease (HR 1·29 95% CI 1·06-1·55). The absolute excess rate of any malignancy was 1·6 per 1000 person years. In the analyses of specific malignancy subgroups I found an increase in the risk of gastrointestinal cancer (HR 1·85) and lymphoproliferative disease (HR 4·80) and decreases in the risk of both breast cancer (HR 0·35) and lung cancer (HR 0·34) in the coeliac disease group compared to the control population. When I

restricted the analyses to the first year after diagnosis I found that most of the hazard ratios were increased (any malignancy HR 1·97 (95% CI 1·39-2·80), gastrointestinal malignancy HR 3·20 (95% CI 1·38-7·39)). After excluding events within the first year of follow up after diagnosis the risks were, in general, decreased (any malignancy HR 1·10 (95% CI 0·87-1·39), gastrointestinal malignancy HR 1·56 (95% CI 0·95-2·58)). The absolute excess rate of any malignancy in this period was 0·6 per 1000 person years.

Table 6-2. Number of events, rates and hazard ratios for maligancy overall and resticted to before and after the first year of follow up after diagnosis

	Overall					Analysis restricted to the first year of follow up after diagnosis				Analysis restricted to follow up beyond one year after diagnosis			
	N**	Number of events**	Rate (10000yr^{-1})	Hazard ratio	95% CI	Number of events**	Rate (10000yr^{-1})	Hazard ratio	95%CI	Number of events**	Rate (10000yr^{-1})	Hazard ratio	95%CI
Any malignancy													
Control cohort*	23433	519	55.9	1		111	52.7	1		395	56.5	1	
Coeliac disease cohort	4695	134	72.0	1.29	(1.06-1.55)	44	104.2	1.97	(1.39-2.80)	87	62.2	1.10	(0.87-1.39)
Gastrointestinal cancer													
Control cohort*	23605	81	8.6	1		14	6.6	1		64	9.0	1	
Coeliac disease cohort	4724	30	15.9	1.85	(1.22-2.81)	9	21.1	3.20	(1.38-7.39)	20	14.1	1.56	(0.95-2.58)
Breast cancer													
Control cohort*	23562	113	12.0	1		24	11.3	1		87	12.3	1	
Coeliac disease cohort	4725	8	4.2	0.35	(0.17-0.72)	3	7.0	0.62	(0.19-2.06)	5	3.5	0.29	(0.12-0.70)
Lung cancer													
Control cohort*	23616	58	6.2	1		14	6.6	1		43	6.0	1	
Coeliac disease cohort	4728	4	2.1	0.34	(0.13-0.95)	1	2.3	0.36	(0.05-2.70)	3	2.1	0.35	(0.11-1.13)
Lymphoproliferative disease													
Control cohort*	23612	24	2.5	1		6	2.8	1		17	2.4	1	
Coeliac disease cohort	4724	23	12.2	4.80	(2.71-8.50)	11	25.8	9.12	(3.37-24.65)	12	8.4	3.55	(1.70-7.43)
Prostate cancer													
Control cohort*	23614	30	3.2	1		4	1.9	1		25	3.5	1	
Coeliac disease cohort	4730	6	3.2	0.99	(0.41-2.38)	1	2.3	1.24	(0.14-11.14)	5	3.5	1.00	(0.38-2.60)

* baseline category, ** numbers vary as those individuals who had an event on the same date or before the start of follow-up were excluded

6.4.2 Mortality

There were 237 deaths among people with coeliac disease and 902 in the control cohort giving overall crude mortality rates of 125·3 per 10000 person years and 95·7 per 10000 person years respectively (Table 6-3). These rates corresponded to a hazard ratio of 1·31 (95% CI 1·13-1·51). The absolute excess rate was 3·0 per 1000 person years. The risk in the first year after diagnosis was considerably higher (HR 1·97 (95% CI 1·50-2·59)) compared to that subsequently (HR 1·17 (95% CI 0·98-1·38)). The absolute excess rate when I excluded deaths within the first year of follow up after diagnosis was 1·7 per 1000 person years. In the multivariate analyses, none of the potential confounding factors I assessed altered the coefficients for the coeliac disease by more than 10%, so they were not included in the final models.

Table 6-3. Number of deaths, rates and hazard ratios overall and restricted to before and after 1 year of follow up after diagnosis

	Overall					Analysis restricted to the first year of follow up after diagnosis				Analysis restricted to follow up beyond one year after diagnosis			
	N**	Number of deaths**	Rate ($10000yr^{-1}$)	Hazard ratio	95% CI	Number of deaths**	Rate ($10000yr^{-1}$)	Hazard ratio	95%CI	Number of deaths**	Rate ($10000yr^{-1}$)	Hazard ratio	95%CI
Mortality													
Control cohort*	23609	902	95.7	1		184	86.7	1		697	98.0	1	
Coeliac disease cohort	4728	237	125.3	1.31	(1.13-1.51)	73	171.0	1.97	(1.50-2.59)	163	114.6	1.17	(0.98-1.38)

* baseline category

** numbers vary as those individuals who had an event on the same date or before the start of follow-up were excluded

6.4.3 Incident/prevalent cases

When I stratified the analyses by prevalent/incident status, having excluded events in the first year after diagnosis, the hazard ratios for overall malignancy were 1·11 (95% CI 0·86-1·44) and 1·03 (95% CI 0·59-1·79) respectively. For mortality the hazard ratio for the prevalent group was 1·09 (95% CI 0·90-1·33) and for the incident group 1·46 (95% CI 1·04-2·07).

6.4.4 Restriction analyses

When I repeated the analyses restricted to only those subjects with coeliac disease who had had at least one gluten-free prescription I found no important differences in the risk estimates (overall malignancy HR 1·20 (95% CI 0·97-1·45), mortality HR 1·20 (95% CI 1·07-1·45)). When I restricted to including only cases with two diagnostic codes for coeliac disease the overall hazard ratios were marginally increased (overall malignancy HR 1.48 (95% CI 1.10-2.00), mortality HR 1.66 (95% CI 1.35-2.05)).

6.4.5 Proportional hazard's assumption

The log –log plots and plots of schoenfeld residuals against time are shown for the overall malignancy analysis and the mortality analysis in Figure 6-1, Figure 6-2, Figure 6-3 and Figure 6-4. The hazard decreases slightly over time for both analyses. I dealt with this by splitting the follow up time into before and subsequent to one year after diagnosis. In general, there was no evidence against the proportional hazards assumption in any of the presented models.

Figure 6-1. Log -log plot for overall malignancy analysis

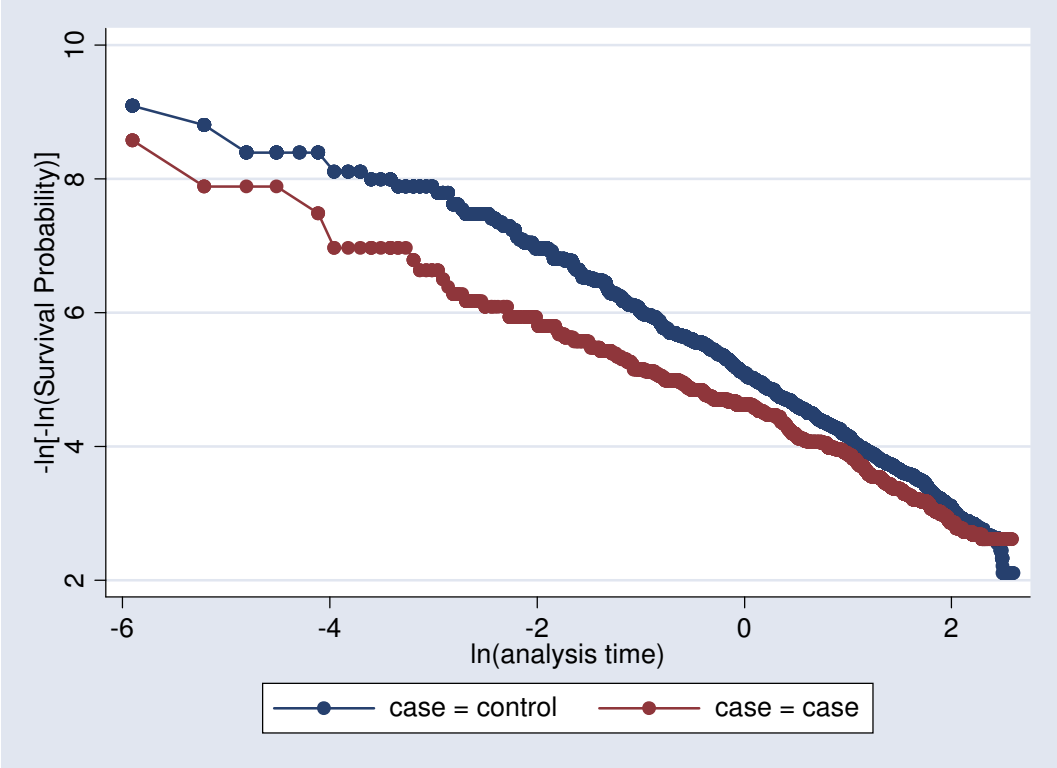


Figure 6-2. Plot of schoenfeld residuals against time for the overall malignancy analysis

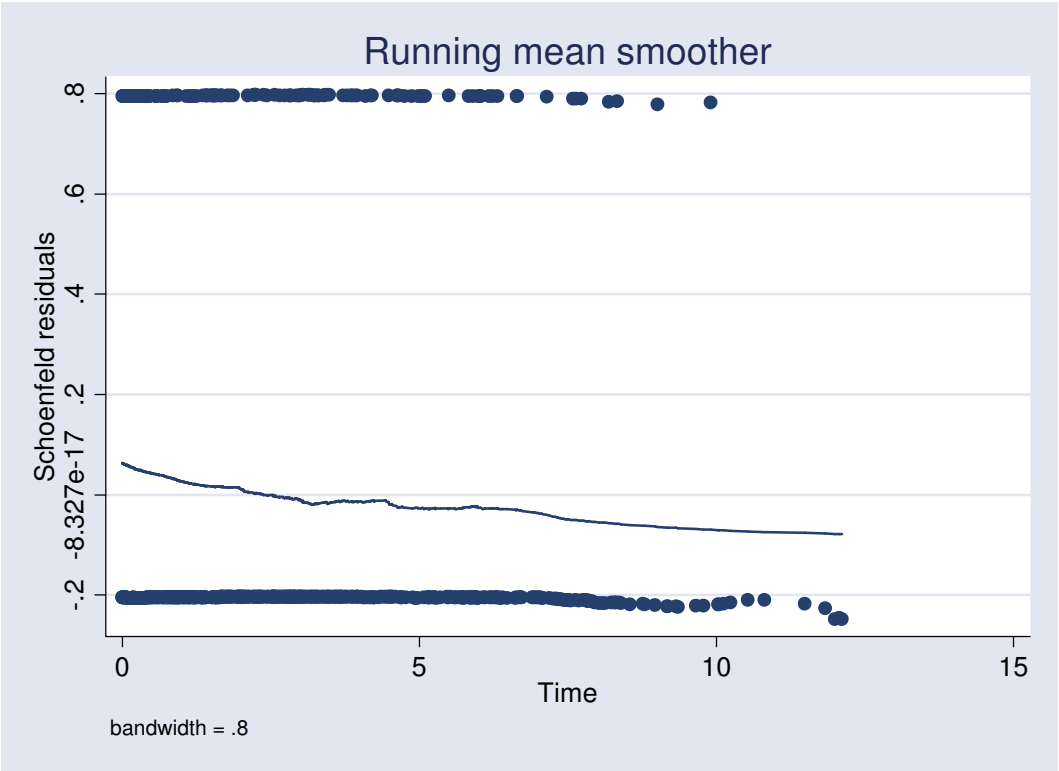


Figure 6-3. Log -log plots for the mortality analysis

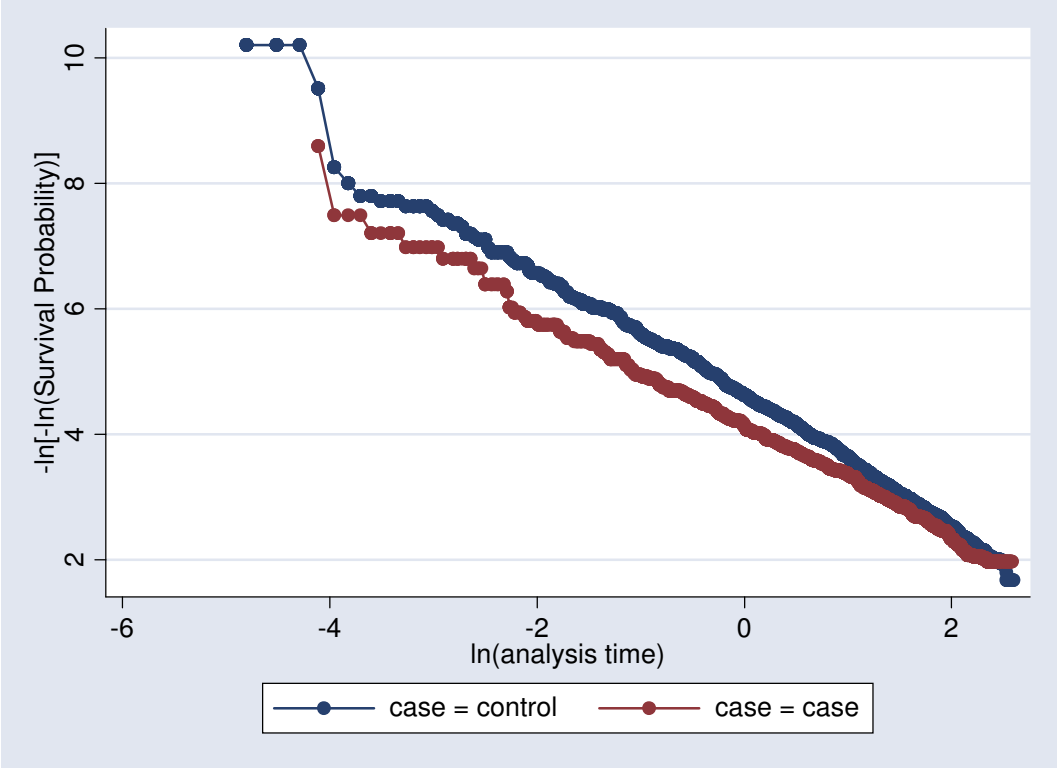
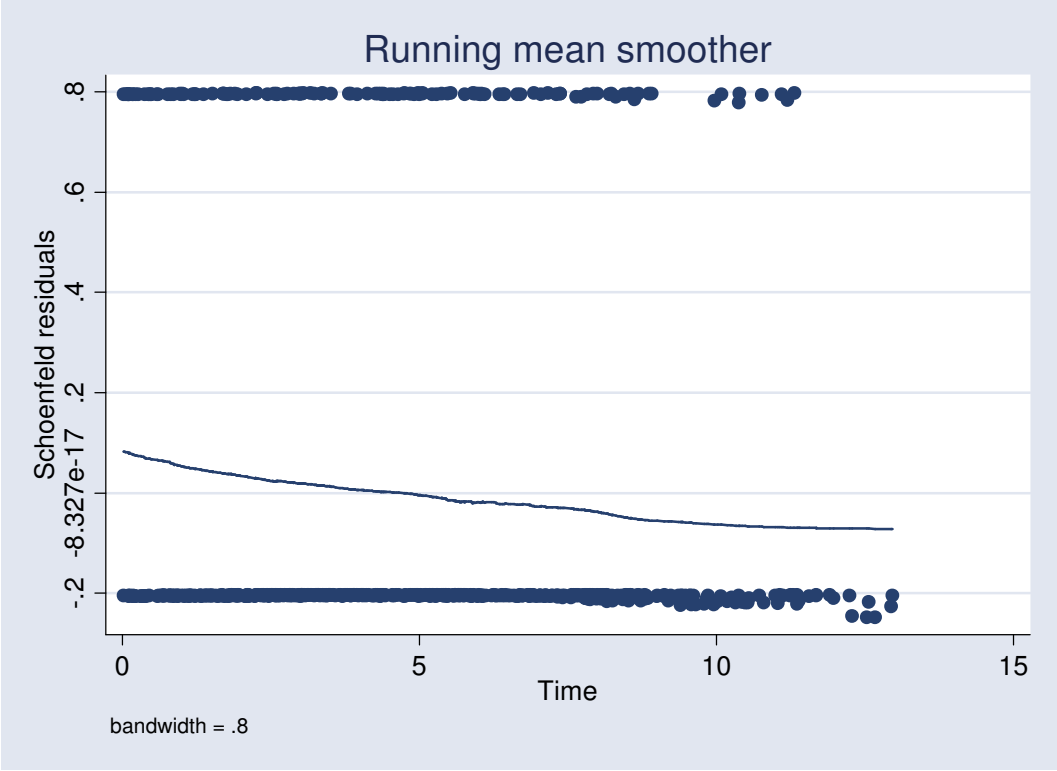


Figure 6-4. Plot of schoenfeld residuals for mortality analysis



6.4.6 Indirect standardisation analysis

The results of the indirect standardisation to the population of England and Wales for both the coeliac and control cohort are shown in Table 6-4. The control population are at substantially reduced risk of death compared to the general population. The estimate of mortality risk is similar, if slightly reduced, for the coeliac cohort compared with the cox proportional hazard's model described earlier.

Table 6-4. Indirect standardisation of both the coeliac disease and control cohorts to the population of England and Wales

	Observed	Expected	Standardised Mortality Ratio	95% CI	
Coeliac disease	223	177.88	1.25	1.09	1.43
Control population	848	924.62	0.92	0.86	0.98

6.4.7 Life table analysis

In the life table analysis (Table 6-5 and Table 6-6) the people with coeliac disease overall have a roughly 2.8 year lower life expectancy than their controls. With advancing age this difference diminishes so that by the age of 60 the life expectancy of a person with coeliac disease is only about 0.7 years lower than that of a control.

Table 6-5. Life table for control cohort

Age group	Deaths observed	Years observed	Mortality rate	Qx	Px	lx	dx	Lx	Tx	ex
0	1	2694.5	0.0004	0.0019	0.9981	1000.0	1.9	4995.4	74991.5	75.0
5	0	4046.0	0.0000	0.0000	1.0000	998.1	0.0	4990.7	69996.1	70.1
10	0	3281.0	0.0000	0.0000	1.0000	998.1	0.0	4990.7	65005.4	65.1
15	1	3028.5	0.0003	0.0016	0.9984	998.1	1.6	4986.6	60014.6	60.1
20	3	4492.7	0.0007	0.0033	0.9967	996.5	3.3	4974.2	55028.0	55.2
25	6	5540.5	0.0011	0.0054	0.9946	993.2	5.4	4952.5	50053.8	50.4
30	4	6413.7	0.0006	0.0031	0.9969	987.8	3.1	4931.4	45101.3	45.7
35	1	7339.2	0.0001	0.0007	0.9993	984.7	0.7	4922.0	40169.9	40.8
40	15	7728.2	0.0019	0.0097	0.9903	984.1	9.5	4896.6	35247.9	35.8
45	20	9230.7	0.0022	0.0108	0.9892	974.6	10.5	4846.6	30351.3	31.1
50	28	9378.4	0.0030	0.0148	0.9852	964.1	14.3	4784.6	25504.8	26.5
55	42	8070.0	0.0052	0.0257	0.9743	949.8	24.4	4687.9	20720.2	21.8
60	70	6390.7	0.0110	0.0533	0.9467	925.4	49.3	4503.6	16032.3	17.3
65	93	5413.6	0.0172	0.0824	0.9176	876.1	72.1	4199.9	11528.7	13.2
70	142	4726.3	0.0300	0.1397	0.8603	803.9	112.3	3738.7	7328.8	9.1
75	186	3581.9	0.0519	0.2298	0.7702	691.6	158.9	3060.6	3590.1	5.2
80	290	2921.3	0.0993	0.3977	0.6023	532.6	211.8	529.5	529.5	1.0

Table 6-6. LIfe table for coeliac disease cohort

Age group	Deaths observed	Years observed	Mortality rate	Qx	Px	lx	dx	Lx	Tx	ex
0	1	544.1	0.0018	0.0091	0.9909	1000.0	9.1	4977.1	72195.6	72.2
5	0	795.3	0.0000	0.0000	1.0000	990.9	0.0	4954.3	67218.5	67.8
10	0	664.9	0.0000	0.0000	1.0000	990.9	0.0	4954.3	62264.2	62.8
15	0	605.1	0.0000	0.0000	1.0000	990.9	0.0	4954.3	57310.0	57.8
20	1	870.7	0.0011	0.0057	0.9943	990.9	5.7	4940.1	52355.7	52.8
25	0	1120.0	0.0000	0.0000	1.0000	985.2	0.0	4925.9	47415.6	48.1
30	2	1284.8	0.0016	0.0078	0.9922	985.2	7.6	4906.8	42489.7	43.1
35	2	1495.9	0.0013	0.0067	0.9933	977.5	6.5	4871.4	37582.9	38.4
40	8	1550.8	0.0052	0.0255	0.9745	971.0	24.7	4793.3	32711.5	33.7
45	6	1895.0	0.0032	0.0157	0.9843	946.3	14.9	4694.3	27918.2	29.5
50	12	1924.8	0.0062	0.0307	0.9693	931.4	28.6	4585.7	23223.8	24.9
55	16	1631.9	0.0098	0.0478	0.9522	902.8	43.2	4406.2	18638.1	20.6
60	17	1293.7	0.0131	0.0636	0.9364	859.6	54.7	4161.5	14231.9	16.6
65	28	1053.9	0.0266	0.1246	0.8754	805.0	100.3	3774.1	10070.4	12.5
70	35	890.2	0.0393	0.1790	0.8210	704.7	126.1	3208.1	6296.2	8.9
75	38	723.8	0.0525	0.2320	0.7680	578.6	134.3	2557.2	3088.1	5.3
80	71	565.2	0.1256	0.4780	0.5220	444.3	212.4	530.9	530.9	1.2

6.5 Discussion

6.5.1 Principal findings

The results of the study show that people with coeliac disease have modest increases in the risks of malignancy and mortality compared to the general population. In addition the life table analysis shows that when compared to the general population people with coeliac disease have a life expectancy reduced on average by 2.8 years. The increased risks were most apparent in the first year after diagnosis and the decreased risks thereafter suggest that some of the overall excess risk is likely to be due to ascertainment. I also found that people with coeliac disease were at approximately a third of the risk of breast or lung cancer compared to the general population in contrast to their increase in risk of gastrointestinal and lymphoproliferative malignancy.

6.5.2 Limitations and merits

A potential weakness of epidemiological studies using routinely collected data such as the GPRD is the validity of diagnostic data for each individual subject involved, particularly with respect to histological status. As discussed before there have been many validation studies of the diagnostic accuracy of the GPRD. Specifically this has included cancer diagnoses which have been found to be accurate [137, 138, 144, 154]. Furthermore, I have carried out my own validation of the diagnosis of coeliac disease (section 3.7.9) and when I restricted the analyses to those people with coeliac disease who also had at least one gluten free prescription, to increase the specificity of the coeliac disease diagnosis, there were no substantial changes in the effect estimates. Although the relative risks were marginally greater in the analysis restricted to those cases with two codes, this is not unexpected as it is likely to include those with most severe disease.

Death recording has also not been specifically validated in the GPRD and the findings of the indirect standardisation suggest that it is possible that death has been under recorded in the control population. An alternative explanation for the low SMR of the control population is that the GPRD

contains a population sample of England and Wales that is not wholly representative of the entire population to the extent that it is, on average, at less risk of death. If this is true then it would appear more appropriate to use the internal comparison as presented in the cox proportional hazard's model.

As discussed before people with coeliac disease may attend general practitioners more frequently than the general population (section 3.4.10) and therefore there is the possibility of differences in ascertainment of some malignancies such as breast or prostate cancer as a result of opportunistic or systematic screening. In addition when people are first investigated for coeliac disease the likelihood of detecting an occult or overt malignancy may be increased or else coeliac disease may be more likely to be detected during the investigation of cancer. If this potential ascertainment bias exists then it would suggest that the risks of lung and breast cancer in the study are underestimates of any reduction through more complete ascertainment among people with coeliac disease. In contrast, the excess risk of gastrointestinal malignancy is likely, in part, to be contributed to by the more detailed investigation of gastrointestinal symptoms particularly at presentation. Unlike previous studies, I had some ability to assess the impact of potential confounders such as body mass index and smoking status on both malignancy and mortality risk. Notwithstanding the incomplete data for these variables I found no suggestion of confounding.

6.5.3 Comparision with other studies

The findings with respect to the risks of overall malignancy and mortality suggest much more modest increases in comparison to other studies. The most recent of these, from Sweden, found slightly greater risks of both malignancy (SIR 1·3) and mortality (SMR 2·0) compared to mine[123, 124]. The slightly greater risks they estimated may reflect greater severity of disease at presentation and/or a period effect as all their subjects with coeliac disease had been hospitalised at least once and follow up ended at least 6 years earlier than in the study. The majority of other studies have found overall increased risks for malignancy or mortality of 2-fold or greater[118, 119, 121, 122]. These studies have been in cohorts of people

with coeliac disease diagnosed and followed up some time ago, or from specialist referral centres. Generally they have been limited by not being population based or too small to provide robust estimates.

My analysis using the life table methodology has not been reported before for coeliac disease. Its advantage over the presentation of hazard ratios is that from the tables I can estimate the life expectancy of the people with coeliac disease compared to the control group at different ages. Notwithstanding the two main assumptions in constructing a life table i.e. that there has been no secular change in survivorship over calendar time and that those censored are assumed to have the same survival experience as those followed up, the findings indicate only a small decrease in life expectancy.

The finding of a marked reduction in the incidence of breast cancer among people with coeliac disease is consistent with two previous studies[121, 124]. A lack of breast cancer in people with coeliac disease was observed in both the Lothian and Swedish cohorts yet the former study was too small to be sure of the association and Askling et al were concerned that, among many comparisons, it may have been a chance finding. It is most unlikely that socio-economic status is an important confounder in this relationship as breast cancer has been consistently shown to be associated with higher socio-economic groups, and there is no good evidence that people with diagnosed coeliac disease are of lower socio-economic status[155]. I have recently shown that people with undetected coeliac disease have low serum lipids and as both this measurable physiological characteristic and dietary fat intake have been implicated in the aetiology of breast cancer, studying the mechanism of protection in coeliac disease may give an opportunity to clarify their role in its aetiology[149, 156-158].

As the study is the first to report a significant reduction in lung cancer incidence this finding must be treated with caution, indeed whether this is a causal relation remains unclear. People with coeliac disease appear to

smoke less than the general population, even when “undiagnosed”[40, 41, 149]. While smoking habit was included in the multivariate analysis some data were missing and I was unable to assess past smoking accurately. Indeed, although adjusting for smoking status did not substantially change the results residual confounding remains a possibility. Nonetheless there has been speculation that factors other than smoking, such as nutritional status and dietary intake of carotinoids, are important in the causation of lung neoplasms[159].

6.5.4 Summary

The findings show that people with diagnosed coeliac disease have modest increases in the relative and absolute risk of malignancy and mortality, with life expectancy reduced on average only by 2.8 years. Most of the excess risk occurs in the first year after diagnosis and although there are markedly increased risks of some malignancies such as gastrointestinal cancers and lymphoma there are substantial reductions in the risk of other, common, cancers such as those of the lung and breast. The latter findings are of particular interest with respect to the possible genetic, nutritional or other environmental factors that may protect people with coeliac disease against developing certain malignancies. Indeed by understanding the mechanism of protection against breast cancer in people with coeliac disease we may gain insight into its aetiology.

7 Conclusions

7.1 Principal findings

The principal findings of this thesis are that:

- Undetected coeliac disease is likely to affect about 1% of the population of England aged 45-76.
- Those affected by undetected coeliac disease have an increased risk of osteoporosis and mild anaemia but also have a favourable cardiovascular risk profile, compared to those without evidence of the disease.
- In people with clinically diagnosed coeliac disease, compared to the general population, there are small increases in both the absolute and relative overall risk of fracture with about a 2-fold increase in the risk of hip fracture.
- Adults with treated coeliac disease have a favourable vascular disease risk factor profile but numbers having heart attacks or strokes are modest and rates of heart attack and stroke are not reduced.
- There are modest increases in the overall risks of malignancy and mortality in people with coeliac disease and most of this excess risk occurred in the first year of follow up after diagnosis.
- There is a marked reduction in the risk of breast and lung cancer in people with coeliac disease.

7.2 Interpretation

The findings of this thesis suggest that the impact on health of both undetected coeliac disease and clinically diagnosed coeliac disease is

important. Although there are clearly some negative health effects, for example mild anaemia and osteoporosis, the possible benefits of having undetected coeliac disease have yet to be fully clarified. The implications on risk of cardiovascular disease of lower body mass index, lower blood pressure, and lower serum cholesterol are potentially large, and as yet, unresolved. From an individual perspective, people with clinically diagnosed, treated, coeliac disease now have information to inform them of the, reasonably precise, actual risks to their health in terms of a range of morbidities and life expectancy. From a population perspective, the suggestion of identifying as many people as possible with the disease through either mass screening or targeted case finding is, in my view, not supported.

7.3 Recommendations for future work

7.3.1 Undetected coeliac disease

Further approaches to clarify the impact of both undetected and clinically diagnosed coeliac disease are suggested by this work. Firstly the cohort described in section 1 has continued follow-up for cause-specific mortality. In the future, a survival analysis carried out on this cohort would give information about the mortality risk in those people with serological evidence of coeliac disease compared to those without evidence of the disease. This plan has the disadvantage of being dependent on a great deal more follow-up before a meaningful analysis could be undertaken. An alternative approach would be to identify a separate cohort that has both serum available for serological testing and prospective follow up for both mortality and perhaps other outcomes (for example malignancy) that has already accrued enough person time and events to give precise estimates.

7.3.2 Diagnosed coeliac disease

Further work that has been recently completed using the GPRD data in this thesis includes the fertility experience of women with coeliac disease compared to the general population. The impact of contemporary clinically diagnosed disease could be further assessed using a similar historical cohort using a larger database similar to the GPRD. Such databases may be available in the future. This would allow an extension of the methodology I have used where comparisons of the risk of various outcomes are assessed before and after diagnosis.

7.3.3 Breast and lung cancer aetiology

The findings in relation to both breast and lung cancer are intriguing. Other study designs to address the apparent "protective" effect of coeliac disease may be of help. For example nested case-control studies from cohorts designed for follow up of cancer may give enough power to look at coeliac disease as an exposure, and in addition be able to explore the way other environmental exposures and physiological measures interact.

7.3.4 Cholesterol and other risk factors for cardiovascular disease

The findings of this thesis appear to be consistent in relation to cholesterol and blood pressure. That cholesterol could change after treatment is a distinct possibility however it is possible that early life factors that determine lifetime cholesterol levels may override any effect of gluten exclusion. As an initial step measurement of cholesterol and other physiological markers could be carried out before and after introduction of a gluten free diet in newly diagnosed people with coeliac disease. If data were available in such people of early life factors such

as birth weight and in utero exposures then our understanding of the mechanisms of cardiovascular disease could be greatly enhanced.

8 References

1. Gee, S., *On the coeliac affection.* St Bartholomew's Hospital Reports, 1888. **24**: p. 17-20.

2. Hass, S.V., *The value of the banana diet in the treatment of coeliac disease.* American Journal of Diseases of Childhood, 1924. **xxviii**: p. 421-437.

3. Dicke, D.K., H.A. Weijers, and J.H. Van de Kamer, *Coeliac disease II; The presence in wheat of a factor having a deleterious effect in cases of coeliac disease.* Acta Paediatr, 1953. **42**: p. 34-42.

4. Schein, J., *Syndrome of non-tropical sprue with hither undescribed lesions of the intestine.* Gastroenterology, 1947. **8**: p. 438-460.

5. Adlersberg, D. and J. Schein, *Clinical and pathological studies in sprue.* Jama, 1947. **134**: p. 1459-67.

6. Himes, H.W. and D. Adlersberg, *Pathological studies in idiopathic sprue.* Mount Sinai Journal of Medicine, 1957. **24**: p. 251-72.

7. Paulley, J.W., *Observations on the aetiology of idiopathic steatorrhoea, jejunal and lymph node biopsies.* Bmj, 1954. **2**: p. 1318-21.

8. Van de Kamer, J.H., H.A. Weijers, and D.K. Dicke, *Coeliac disease IV; An investigation into the injurious constituents of wheat in connection with their action on patients with coeliac disease.* Acta Paediatr, 1953. **42**: p. 223-227.

9. Taylor, K.B., S.C. Truelove, D.L. Thomson, and R. Wright, *An immunological study of coeliac disease and idiopathic steatorrhoea.* Bmj, 1961. **2**: p. 1727-31.

10. Houlston, R.S. and D. Ford, *Genetics of coeliac disease.* Qjm, 1996. **89**(10): p. 737-43.

11. Sollid, L.M., G. Markussen, J. Ek, H. Gjerde, F. Vartdal, and E. Thorsby, *Evidence for a primary association of celiac disease to a particular HLA-DQ alpha/beta heterodimer.* J Exp Med, 1989. **169**(1): p. 345-50.

12. Louka, A.S. and L.M. Sollid, *HLA in coeliac disease: Unravelling the complex genetics of a complex disorder.* Tissue Antigens, 2003. **61**(2): p. 105-17.

13. Nilsen, E.M., K.E. Lundin, P. Krajci, H. Scott, L.M. Sollid, and P. Brandtzaeg, *Gluten specific, HLA-DQ restricted T cells from coeliac mucosa produce cytokines with Th1 or Th0 profile dominated by interferon gamma.* Gut, 1995. **37**(6): p. 766-76.

14. Anderson, R.P., P. Degano, A.J. Godkin, D.P. Jewell, and A.V. Hill, *In vivo antigen challenge in celiac disease identifies a single transglutaminase-modified peptide as the dominant A-gliadin T-cell epitope.* Nat Med, 2000. **6**(3): p. 337-42.

15. Arentz-Hansen, H., R. Korner, O. Molberg, H. Quarsten, W. Vader, Y.M. Kooy, K.E. Lundin, F. Koning, P. Roepstorff, L.M. Sollid, and S.N. McAdam, *The intestinal T cell response to alpha-gliadin in adult celiac disease is focused on a single deamidated glutamine targeted by tissue transglutaminase.* J Exp Med, 2000. **191**(4): p. 603-12.

16. Arentz-Hansen, H., S.N. McAdam, O. Molberg, B. Fleckenstein, K.E. Lundin, T.J. Jorgensen, G. Jung, P. Roepstorff, and L.M. Sollid, *Celiac lesion T cells recognize epitopes that cluster in regions of gliadins rich in proline residues.* Gastroenterology, 2002. **123**(3): p. 803-9.

17. Schuppan, D., *Current concepts of celiac disease pathogenesis.* Gastroenterology, 2000. **119**(1): p. 234-42.

18. Dieterich, W., T. Ehnis, M. Bauer, P. Donner, U. Volta, E.O. Riecken, and D. Schuppan, *Identification of tissue transglutaminase as the autoantigen of celiac disease.* Nat Med, 1997. **3**(7): p. 797-801.

19. Lock, R.J., M.C. Pitcher, and D.J. Unsworth, *IgA anti-tissue transglutaminase as a diagnostic marker of gluten sensitive enteropathy.* J Clin Pathol, 1999. **52**(4): p. 274-7.

20. Sardy, M., U. Odenthal, S. Karpati, M. Paulsson, and N. Smyth, *Recombinant human tissue transglutaminase ELISA for the diagnosis of gluten-sensitive enteropathy.* Clin Chem, 1999. **45**(12): p. 2142-9.

21. Sulkanen, S., T. Halttunen, K. Laurila, K.L. Kolho, I.R. Korponay-Szabo, A. Sarnesto, E. Savilahti, P. Collin, and M. Maki, *Tissue transglutaminase autoantibody enzyme-linked immunosorbent assay in detecting celiac disease.* Gastroenterology, 1998. **115**(6): p. 1322-8.

22. Shan, L., O. Molberg, I. Parrot, F. Hausch, F. Filiz, G.M. Gray, L.M. Sollid, and C. Khosla, *Structural basis for gluten intolerance in celiac sprue.* Science, 2002. **297**(5590): p. 2275-9.

23. Koning, F., *The molecular basis of celiac disease.* J Mol Recognit, 2003. **16**(5): p. 333-6.

24. Marsh, M.N., *Mucosal pathology in gluten sensitivity*, in *Coeliac disease*, M.N. Marsh, Editor. 1992, Blackwell Scientific Publications: Oxford. p. 136-191.

25. Ferguson, A. and D. Murray, *Quantitation of intraepithelial lymphocytes in human jejunum.* Gut, 1971. **12**(12): p. 988-94.

26. Groh, V., A. Steinle, S. Bauer, and T. Spies, *Recognition of stress-induced MHC molecules by intestinal epithelial gammadelta T cells.* Science, 1998. **279**(5357): p. 1737-40.

27. Bauer, S., V. Groh, J. Wu, A. Steinle, J.H. Phillips, L.L. Lanier, and T. Spies, *Activation of NK cells and T cells by NKG2D, a receptor for stress-inducible MICA.* Science, 1999. **285**(5428): p. 727-9.

28. Anderson, D.H. and P.A. Di Sant'Agnese, *Idiopathic celiac disease: I. Mode of onset and diagnosis.* Pediatrics, 1953. **11**: p. 207-22.

29. Stevens, F.M., B. Egan-Mitchell, E. Cryan, C.F. McCarthy, and B. McNicholl, *Decreasing incidence of coeliac disease.* Arch Dis Child, 1987. **62**(5): p. 465-8.

30. Littlewood, J.M., A.J. Crollick, and I.D.G. Richards, *Childhood coeliac disease is disappearing [Letter].* Lancet, 1980. **ii**: p. 1350.

31. Dossetor, J.F.B., A.A.M. Gibson, and A.S. McNeish, *Childhood coeliac disease is disappearing [Letter].* Lancet, 1981. **i**: p. 322-3.

32. Logan, R.F., E.A. Rifkind, A. Busuttil, H.M. Gilmour, and A. Ferguson, *Prevalence and "incidence" of celiac disease in Edinburgh and the Lothian region of Scotland.* Gastroenterology, 1986. **90**(2): p. 334-42.

33. *Present-day practice in infant feeding.*, in *Report of a working party of the panel on child nutrition*. 1974: Committee on Medical Aspects of Food Policy. Her Majesty's Stationary Office; London. p. 24-26.

34. Linberg, T., *Coeliac disease and infant feeding practices [Letter].* Lancet, 1981. **i**: p. 449.

35. Auricchio, S., D. Follo, G. de Ritis, A. Giunta, D. Marzorati, L. Prampolini, N. Ansaldi, P. Levi, D. Dall'Olio, A. Bossi, and et al., *Does breast feeding protect against the development of clinical symptoms of celiac disease in children?* J Pediatr Gastroenterol Nutr, 1983. **2**(3): p. 428-33.

36. Greco, L., S. Auricchio, M. Mayer, and M. Grimaldi, *Case control study on nutritional risk factors in celiac disease.* J Pediatr Gastroenterol Nutr, 1988. **7**(3): p. 395-9.

37. Ivarsson, A., L.A. Persson, L. Nystrom, H. Ascher, B. Cavell, L. Danielsson, A. Dannaeus, T. Lindberg, B. Lindquist, L. Stenhammar, and O. Hernell, *Epidemic of coeliac disease in Swedish children.* Acta Paediatr, 2000. **89**(2): p. 165-71.

38. Ivarsson, A., O. Hernell, H. Stenlund, and L.A. Persson, *Breast-feeding protects against celiac disease.* Am J Clin Nutr, 2002. **75**(5): p. 914-21.

39. Austin, A.S., R.F. Logan, K. Thomason, and G.K. Holmes, *Cigarette smoking and adult coeliac disease.* Scand J Gastroenterol, 2002. **37**(8): p. 978-82.

40. Patel, A.H., E.V. Loftus, Jr., J.A. Murray, W.S. Harmsen, A.R. Zinsmeister, and W.J. Sandborn, *Cigarette smoking and celiac sprue: a case-control study.* Am J Gastroenterol, 2001. **96**(8): p. 2388-91.

41. Snook, J.A., L. Dwyer, C. Lee-Elliott, S. Khan, D.W. Wheeler, and D.S. Nicholas, *Adult coeliac disease and cigarette smoking.* Gut, 1996. **39**(1): p. 60-2.

42. Vazquez, H., E. Smecuol, D. Flores, R. Mazure, S. Pedreira, S. Niveloni, E. Maurino, and J.C. Bai, *Relation between cigarette smoking and celiac disease: evidence from a case-control study.* Am J Gastroenterol, 2001. **96**(3): p. 798-802.

43. *American Gastroenterological Association medical position statement: Celiac Sprue.* Gastroenterology, 2001. **120**(6): p. 1522-5.

44. Fasano, A. and C. Catassi, *Current approaches to diagnosis and treatment of celiac disease: an evolving spectrum.* Gastroenterology, 2001. **120**(3): p. 636-51.

45. Feighery, C., *Fortnightly review: coeliac disease.* Bmj, 1999. **319**(7204): p. 236-9.

46. Ciclitira, P.J., A.L. King, and J.S. Fraser, *AGA technical review on Celiac Sprue. American Gastroenterological Association.* Gastroenterology, 2001. **120**(6): p. 1526-40.

47. Maki, M. and P. Collin, *Coeliac disease.* Lancet, 1997. **349**(9067): p. 1755-9.

48. Hawkes, N.D., G.L. Swift, P.M. Smith, and H.R. Jenkins, *Incidence and presentation of coeliac disease in South Glamorgan.* Eur J Gastroenterol Hepatol, 2000. **12**(3): p. 345-9.

49. Bode, S. and E. Gudmand-Hoyer, *Symptoms and haematologic features in consecutive adult coeliac patients.* Scand J Gastroenterol, 1996. **31**(1): p. 54-60.

50. Logan, R.F., G. Tucker, E.A. Rifkind, R.C. Heading, and A. Ferguson, *Changes in clinical features of coeliac disease in adults in Edinburgh and the Lothians 1960-79.* BMJ, 1983. **286**(6359): p. 95-7.

51. Sanders, D.S., D.P. Hurlstone, R.O. Stokes, F. Rashid, A. Milford-Ward, M. Hadjivassiliou, and A.J. Lobo, *Changing face of adult coeliac disease: experience of a single university hospital in South Yorkshire.* Postgrad Med J, 2002. **78**(915): p. 31-3.

52. Lo, W., K. Sano, B. Lebwohl, B. Diamond, and P.H. Green, *Changing presentation of adult celiac disease.* Dig Dis Sci, 2003. **48**(2): p. 395-8.

53. Hin, H., G. Bird, P. Fisher, N. Mahy, and D. Jewell, *Coeliac disease in primary care: case finding study.* BMJ, 1999. **318**(7177): p. 164-7.

54. Barry, R.E., P. Baker, and A.E. Read, *Coeliac disease. The clinical presentation.* Clin Gastroenterol, 1974. **3**(1): p. 55-69.

55. Volta, U., C. Bonazzi, A.M. Baldoni, M. Lenzi, F. Cassani, and E. Pisi, *Clinical presentation of adult coeliac disease.* Ann Med Interne (Paris), 1988. **139**(2): p. 123-4.

56. Vajro, P., A. Fontanella, M. Mayer, A. De Vincenzo, L.M. Terracciano, M. D'Armiento, and R. Vecchione, *Elevated serum aminotransferase activity as an early manifestation of gluten-sensitive enteropathy.* J Pediatr, 1993. **122**(3): p. 416-9.

57. Bardella, M.T., M. Fraquelli, M. Quatrini, N. Molteni, P. Bianchi, and D. Conte, *Prevalence of hypertransaminasemia in adult celiac patients and effect of gluten-free diet.* Hepatology, 1995. **22**(3): p. 833-6.

58. Novacek, G., W. Miehsler, F. Wrba, P. Ferenci, E. Penner, and H. Vogelsang, *Prevalence and clinical importance of hypertransaminasaemia in coeliac disease.* Eur J Gastroenterol Hepatol, 1999. **11**(3): p. 283-8.

59. Dickey, W., S.A. McMillan, J.S. Collins, R.G. Watson, J.C. McLoughlin, and A.H. Love, *Liver abnormalities associated with celiac sprue. How common are they, what is their significance, and what do we do about them?* J Clin Gastroenterol, 1995. **20**(4): p. 290-2.

60. Bardella, M.T., M. Vecchi, D. Conte, E. Del Ninno, M. Fraquelli, S. Pacchetti, E. Minola, M. Landoni, B.M. Cesana, and R. De Franchis, *Chronic unexplained hypertransaminasemia may be caused by occult celiac disease.* Hepatology, 1999. **29**(3): p. 654-7.

61. Greco, L., A.E. Tozzi, M. Mayer, M. Grimaldi, G. Silano, and S. Auricchio, *Unchanging clinical picture of coeliac disease presentation in Campania, Italy.* Eur J Pediatr, 1989. **148**(7): p. 610-3.

62. Gumaa, S.N., B. McNicholl, B. Egan-Mitchell, K. Connolly, and B.G. Loftus, *Coeliac disease in Galway, Ireland 1971-1990.* Ir Med J, 1997. **90**(2): p. 60-1.

63. George, E.K., M.L. Mearin, H.C. Franken, R.H. Houwen, R.A. Hirasing, and J.P. Vandenbroucke, *Twenty years of childhood coeliac disease in The Netherlands: a rapidly increasing incidence?* Gut, 1997. **40**(1): p. 61-6.

64. Visakorpi, J.K. and M. Maki, *Changing clinical features of coeliac disease.* Acta Paediatr Suppl, 1994. **83**(395): p. 10-3.

65. Paerregaard, A., M. Vilien, P.A. Krasilnikoff, and E. Gudmand-Hoyer, *Supposed coeliac disease during childhood and its presentation 14-38 years later.* Scand J Gastroenterol, 1988. **23**(1): p. 65-70.

66. Chorzelski, T.P., E.H. Beutner, J. Sulej, H. Tchorzewska, S. Jablonska, V. Kumar, and A. Kapuscinska, *IgA anti-endomysium antibody. A new immunological marker of dermatitis herpetiformis and coeliac disease.* Br J Dermatol, 1984. **111**(4): p. 395-402.

67. James, M.W. and B.B. Scott, *Endomysial antibody in the diagnosis and management of coeliac disease.* Postgrad Med J, 2000. **76**(898): p. 466-8.

68. Burgin-Wolff, A., I. Dahlbom, F. Hadziselimovic, and C.J. Petersson, *Antibodies against human tissue transglutaminase and endomysium in diagnosing and monitoring coeliac disease.* Scand J Gastroenterol, 2002. **37**(6): p. 685-91.

69. Horvath, K. and I.D. Hill, *Anti-tissue transglutaminase antibody as the first line screening for celiac disease: good-bye antigliadin tests?* Am J Gastroenterol, 2002. **97**(11): p. 2702-4.

70. Dieterich, W., E. Laag, H. Schopper, U. Volta, A. Ferguson, H. Gillett, E.O. Riecken, and D. Schuppan, *Autoantibodies to tissue transglutaminase as predictors of celiac disease.* Gastroenterology, 1998. **115**(6): p. 1317-21.

71. Hill, P.G., S.P. Thompson, and G.K. Holmes, *IgA anti-gliadin antibodies in adult celiac disease.* Clin Chem, 1991. **37**(5): p. 647-50.

72. West, J., C.A. Lloyd, P.G. Hill, and G.K. Holmes, *IgA-antitissue transglutaminase: validation of a commercial assay for diagnosing coeliac disease.* Clin Lab, 2002. **48**(5-6): p. 241-6.

73. Cataldo, F., V. Marino, G. Bottaro, P. Greco, and A. Ventura, *Celiac disease and selective immunoglobulin A deficiency.* J Pediatr, 1997. **131**(2): p. 306-8.

74. Hill, P.G., J.M. Forsyth, D. Semeraro, and G.K. Holmes, *IgA antibodies to human tissue transglutaminase: audit of routine practice confirms high diagnostic accuracy.* Scand J Gastroenterol, 2004. **39**(11): p. 1078-82.

75. *Revised criteria for diagnosis of coeliac disease. Report of Working Group of European Society of Paediatric Gastroenterology and Nutrition.* Arch Dis Child, 1990. **65**(8): p. 909-11.

76. Ivarsson, A., L.A. Persson, P. Juto, M. Peltonen, O. Suhr, and O. Hernell, *High prevalence of undiagnosed coeliac disease in adults: a Swedish population-based study.* J Intern Med, 1999. **245**(1): p. 63-8.

77. Johnston, S.D., R.G. Watson, S.A. McMillan, J. Sloan, and A.H. Love, *Prevalence of coeliac disease in Northern Ireland [letter].* Lancet, 1997. **350**(9088): p. 1370.

78. Kolho, K.L., M.A. Farkkila, and E. Savilahti, *Undiagnosed coeliac disease is common in Finnish adults.* Scand J Gastroenterol, 1998. **33**(12): p. 1280-3.

79. Riestra, S., E. Fernandez, L. Rodrigo, S. Garcia, and G. Ocio, *Prevalence of Coeliac disease in the general population of northern Spain. Strategies of serologic screening.* Scand J Gastroenterol, 2000. **35**(4): p. 398-402.

80. Volta, U., S. Bellentani, F.B. Bianchi, G. Brandi, L. De Franceschi, L. Miglioli, A. Granito, F. Balli, and C. Tiribelli, *High prevalence of celiac disease in Italian general population.* Dig Dis Sci, 2001. **46**(7): p. 1500-5.

81. Cook, H.B., M.J. Burt, J.A. Collett, M.R. Whitehead, C.M. Frampton, and B.A. Chapman, *Adult coeliac disease: prevalence and clinical significance.* J Gastroenterol Hepatol, 2000. **15**(9): p. 1032-6.

82. Couignoux, S., A. Ocmant, D. Cottel, F. Mascart, K. Geloes, A. Cortot, J. Colombel, F., and P. Amouyel, *Prevalence of adult coeliac disease in Northern France.* Gut, 2000. **47**(Suppl III): p. A196.

83. Hovell, C.J., J.A. Collett, G. Vautier, A.J. Cheng, E. Sutanto, D.F. Mallon, J.K. Olynyk, and D.J. Cullen, *High prevalence of coeliac disease in a population-*

based study from Western Australia: a case for screening? Med J Aust, 2001. **175**(5): p. 247-50.

84. Gomez, J.C., G.S. Selvaggio, M. Viola, B. Pizarro, G. la Motta, S. de Barrio, R. Castelletto, R. Echeverria, E. Sugai, H. Vazquez, E. Maurino, and J.C. Bai, *Prevalence of celiac disease in Argentina: screening of an adult population in the La Plata area.* Am J Gastroenterol, 2001. **96**(9): p. 2700-4.

85. Trevisiol, C., T. Not, I. Berti, E. Buratti, A. Citta, E. Neri, G. Torre, S. Martelossi, A. Tommasini, A. Alu, G. Barillari, S. Facchini, and A. Ventura, *Screening for coeliac disease in healthy blood donors at two immuno-transfusion centres in north-east Italy.* Ital J Gastroenterol Hepatol, 1999. **31**(7): p. 584-6.

86. Rostami, K., C.J. Mulder, J.M. Werre, F.R. van Beukelen, J. Kerchhaert, J.B. Crusius, A.S. Pena, F.L. Willekens, and J.W. Meijer, *High prevalence of celiac disease in apparently healthy blood donors suggests a high prevalence of undiagnosed celiac disease in the Dutch population.* Scand J Gastroenterol, 1999. **34**(3): p. 276-9.

87. Not, T., K. Horvath, I.D. Hill, J. Partanen, A. Hammed, G. Magazzu, and A. Fasano, *Celiac disease risk in the USA: high prevalence of antiendomysium antibodies in healthy blood donors.* Scand J Gastroenterol, 1998. **33**(5): p. 494-8.

88. Sjoberg, K. and S. Eriksson, *Regional differences in coeliac disease prevalence in Scandinavia?* Scand J Gastroenterol, 1999. **34**(1): p. 41-5.

89. Gandolfi, L., R. Pratesi, J.C. Cordoba, P.L. Tauil, M. Gasparin, and C. Catassi, *Prevalence of celiac disease among blood donors in Brazil [see comments].* Am J Gastroenterol, 2000. **95**(3): p. 689-92.

90. Catassi, C., I.M. Ratsch, E. Fabiani, M. Rossini, F. Bordicchia, F. Candela, G.V. Coppa, and P.L. Giorgi, *Coeliac disease in the year 2000: exploring the iceberg.* Lancet, 1994. **343**(8891): p. 200-3.

91. Meloni, G., A. Dore, G. Fanciulli, F. Tanda, and G.F. Bottazzo, *Subclinical coeliac disease in schoolchildren from northern Sardinia.* Lancet, 1999. **353**(9146): p. 37.

92. Bingley, P.J., A.J. Williams, A.J. Norcross, D.J. Unsworth, R.J. Lock, A.R. Ness, and R.W. Jones, *Undiagnosed coeliac disease at age seven: population based prospective birth cohort study.* Bmj, 2004. **328**(7435): p. 322-3.

93. Catassi, C., I.M. Ratsch, L. Gandolfi, R. Pratesi, E. Fabiani, R. El Asmar, M. Frijia, I. Bearzi, and L. Vizzoni, *Why is coeliac disease endemic in the people of the Sahara? [letter].* Lancet, 1999. **354**(9179): p. 647-8.

94. Maki, M., K. Mustalahti, J. Kokkonen, P. Kulmala, M. Haapalahti, T. Karttunen, J. Ilonen, K. Laurila, I. Dahlbom, T. Hansson, P. Hopfl, and M. Knip, *Prevalence of Celiac disease among children in Finland.* N Engl J Med, 2003. **348**(25): p. 2517-24.

95. Corazza, G.R., M.L. Andreani, F. Biagi, G. Corrao, S. Pretolani, G. Giulianelli, G. Ghironzi, and G. Gasbarrini, *The smaller size of the 'coeliac iceberg' in adults.* Scand J Gastroenterol, 1997. **32**(9): p. 917-9.

96. Collin, P., T. Reunala, M. Rasmussen, S. Kyronpalo, E. Pehkonen, P. Laippala, and M. Maki, *High incidence and prevalence of adult coeliac disease. Augmented diagnostic approach.* Scand J Gastroenterol, 1997. **32**(11): p. 1129-33.

97. Bode, S. and E. Gudmand-Hoyer, *Incidence and prevalence of adult coeliac disease within a defined geographic area in Denmark.* Scand J Gastroenterol, 1996. **31**(7): p. 694-9.

98. Murray, J.A., C. van Dyke, M.F. Plevak, R.A. Dierkhising, A.R. Zinsmeister, and L.J. Melton, *Trends in the identification and clinical features of celiac disease in a North American Community, 1950-2001.* Clinical Gastroenterology and Hepatology, 2003. **1**(1): p. 19-27.

99. Maki, M., K. Kallonen, M.L. Lahdeaho, and J.K. Visakorpi, *Changing pattern of childhood coeliac disease in Finland.* Acta Paediatr Scand, 1988. **77**(3): p. 408-12.

100. Dahan, S., P.E. Slater, M. Cooper, C. Brautbar, and A. Ashkenazi, *Coeliac disease in the Rehovot-Ashdod region of Israel: incidence and ethnic distribution.* J Epidemiol Community Health, 1984. **38**(1): p. 58-60.

101. McConnell, R.B., *Membership of the Coeliac Society of the United Kingdom*, in *The Genetics of Coeliac Disease*, R.B. McConnell, Editor. 1981, MTP Press: Lancaster. p. 65-9.

102. Logan, R.F., *Epidemiology of coeliac disease*, in *Coeliac disease*, M.N. Marsh, Editor. 1992, Blackwell Scientific Publications: London. p. 192-214.

103. Corazza, G.R., A. Di Sario, L. Cecchetti, R.A. Jorizzo, M. Di Stefano, L. Minguzzi, G. Brusco, M. Bernardi, and G. Gasbarrini, *Influence of pattern of clinical presentation and of gluten-free diet on bone mass and metabolism in adult coeliac disease.* Bone, 1996. **18**(6): p. 525-30.

104. Corazza, G.R., A. Di Sario, G. Sacco, G. Zoli, E.A. Treggiari, G. Brusco, and G. Gasbarrini, *Subclinical coeliac disease: an anthropometric assessment.* J Intern Med, 1994. **236**(2): p. 183-7.

105. Mustalahti, K., P. Collin, H. Sievanen, J. Salmi, and M. Maki, *Osteopenia in patients with clinically silent coeliac disease warrants screening [letter].* Lancet, 1999. **354**(9180): p. 744-5.

106. Valdimarsson, T., O. Lofman, G. Toss, and M. Strom, *Reversal of osteopenia with diet in adult coeliac disease.* Gut, 1996. **38**(3): p. 322-7.

107. Valdimarsson, T., *Three Years' Follow-up of Bone Density in Adult Coeliac Disease: Significance of Secondary Hyperparathyroidism.* Scand J Gastroenterol, 2000. **3**: p. 274-280.

108. McFarlane, X.A., A.K. Bhalla, D.E. Reeves, L.M. Morgan, and D.A. Robertson, *Osteoporosis in treated adult coeliac disease.* Gut, 1995. **36**(5): p. 710-4.

109. Bernstein, C.N., W.D. Leslie, and M.S. Leboff, *AGA technical review on osteoporosis in gastrointestinal diseases.* Gastroenterology, 2003. **124**(3): p. 795-841.

110. Scott, E.M., I. Gaywood, and B.B. Scott, *Guidelines for osteoporosis in coeliac disease and inflammatory bowel disease. British Society of Gastroenterology.* Gut, 2000. **46**(Suppl 1): p. i1-8.

111. Thomason, K., J. West, R.F. Logan, C. Coupland, and G.K. Holmes, *Fracture experience of patients with coeliac disease: a population based survey.* Gut, 2003. **52**(4): p. 518-22.

112. Vasquez, H., R. Mazure, D. Gonzalez, D. Flores, S. Pedreira, S. Niveloni, E. Smecuol, E. Maurino, and J.C. Bai, *Risk of fractures in celiac disease patients: a cross-sectional, case-control study.* Am J Gastroenterol, 2000. **95**(1): p. 183-9.

113. Vestergaard, P. and L. Mosekilde, *Fracture risk in patients with celiac Disease, Crohn's disease, and ulcerative colitis: a nationwide follow-up study of 16,416 patients in Denmark.* Am J Epidemiol, 2002. **156**(1): p. 1-10.

114. Fickling, W.E., X.A. McFarlane, A.K. Bhalla, and D.A. Robertson, *The clinical impact of metabolic bone disease in coeliac disease.* Postgrad Med J, 2001. **77**(903): p. 33-6.

115. Moreno, M.L., H. Vazquez, R. Mazure, E. Smecuol, S. Niveloni, S. Pedreira, E. Sugai, E. Maurino, J.C. Gomez, and J.C. Bai, *Stratification of bone fracture risk in patients with celiac disease.* Clin Gastroenterol Hepatol, 2004. **2**(2): p. 127-34.

116. Whorwell, P.J., M.R. Alderson, K.J. Foster, and R. Wright, *Death from ischaemic heart-disease and malignancy in adult patients with coeliac disease.* Lancet, 1976. **2**(7977): p. 113-4.

117. Vuoristo, M., S. Tarpila, and T.A. Miettinen, *Serum lipids and faecal steroids in patients with coeliac disease: Effects of gluten-free diet and cholestyramine.* Gastroenterology, 1980. **78**: p. 1518-1525.

118. Corrao, G., G.R. Corazza, V. Bagnardi, G. Brusco, C. Ciacci, M. Cottone, C.S. Guidetti, P. Usai, P. Cesari, M.A. Pelli, S. Loperfido, U. Volta, A. Calabro, and M. Certo, *Mortality in patients with coeliac disease and their relatives: a cohort study.* Lancet, 2001. **358**(9279): p. 356-61.

119. Cottone, M., A. Termini, L. Oliva, A. Magliocco, C. Marrone, A. Orlando, F. Pinzone, R. Di Mitri, M. Rosselli, A. Rizzo, and L. Pagliaro, *Mortality and causes of death in celiac disease in a Mediterranean area.* Dig Dis Sci, 1999. **44**(12): p. 2538-41.

120. Cooper, B.T., G.K. Holmes, R. Ferguson, and W.T. Cooke, *Celiac disease and malignancy.* Medicine (Baltimore), 1980. **59**(4): p. 249-61.

121. Logan, R.F., E.A. Rifkind, I.D. Turner, and A. Ferguson, *Mortality in celiac disease.* Gastroenterology, 1989. **97**(2): p. 265-71.

122. Holmes, G.K., P. Prior, M.R. Lane, D. Pope, and R.N. Allan, *Malignancy in coeliac disease--effect of a gluten free diet.* Gut, 1989. **30**(3): p. 333-8.

123. Peters, U., J. Askling, G. Gridley, A. Ekbom, and M. Linet, *Causes of death in patients with celiac disease in a population-based Swedish cohort.* Arch Intern Med, 2003. **163**(13): p. 1566-72.

124. Askling, J., M. Linet, G. Gridley, T.S. Halstensen, K. Ekstrom, and A. Ekbom, *Cancer incidence in a population-based cohort of individuals hospitalized with celiac disease or dermatitis herpetiformis.* Gastroenterology, 2002. **123**(5): p. 1428-1435.

125. Mulder, C.J., M.M. Hadithi, K. Rostami, and M.S. Goerres, *Coeliac disease--has the time come for routine mass screening? In 2002--2010--2020?* Rom J Gastroenterol, 2002. **11**(3): p. 179-82.

126. Kumar, P.J., *European and North American populations should be screened for coeliac disease.* Gut, 2003. **52**(2): p. 170-1.

127. Fasano, A., *European and North American populations should be screened for coeliac disease.* Gut, 2003. **52**(2): p. 168-9.

128. Logan, R.F., *Screening for coeliac disease--has the time come for mass screening?* Acta Paediatr Suppl, 1996. **412**: p. 15-9.

129. Trivedi, D.P. and K.T. Khaw, *Bone mineral density at the hip predicts mortality in elderly men.* Osteoporos Int, 2001. **12**(4): p. 259-65.

130. May, H., S. Murphy, and K.T. Khaw, *Bone mineral density and its relationship to skin colour in Caucasian females.* Eur J Clin Invest, 1995. **25**(2): p. 85-9.

131. May, H., S. Murphy, and K.T. Khaw, *Age-associated bone loss in men and women and its relationship to weight.* Age Ageing, 1994. **23**(3): p. 235-40.

132. *International Classification of Diseases: manual of the international statistical classification of diseases, injuries and causes of death*. 9th revision ed. 1975, Geneva: WHO.

133. Sanders, D.S., M.J. Carter, D.P. Hurlstone, A. Pearce, A.M. Ward, M.E. McAlindon, and A.J. Lobo, *Association of adult coeliac disease with irritable bowel syndrome: a case-control study in patients fulfilling ROME II criteria referred to secondary care.* Lancet, 2001. **358**(9292): p. 1504-8.

134. Peters, U., S. Schneeweiss, E.A. Trautwein, and H.F. Erbersdobler, *A case-control study of the effect of infant feeding on celiac disease.* Ann Nutr Metab, 2001. **45**(4): p. 135-42.

135. Law, M.R., N.J. Wald, and S.G. Thompson, *By how much and how quickly does reduction in serum cholesterol concentration lower risk of ischaemic heart disease?* Bmj, 1994. **308**(6925): p. 367-72.

136. Garcia Rodriguez, L.A. and S. Perez Gutthann, *Use of the UK General Practice Research Database for pharmacoepidemiology.* Br J Clin Pharmacol, 1998. **45**(5): p. 419-25.

137. Hollowell, J., *The General Practice Research Database: quality of morbidity data.* Popul Trends, 1997(87): p. 36-40.

138. Jick, H., S.S. Jick, and L.E. Derby, *Validation of information recorded on general practitioner based computerised data resource in the United Kingdom.* Bmj, 1991. **302**(6779): p. 766-8.

139. Jick, S.S., J.A. Kaye, C. Vasilakis-Scaramozza, L.A. Garcia Rodriguez, A. Ruigomez, C.R. Meier, R.G. Schlienger, C. Black, and H. Jick, *Validity of the general practice research database.* Pharmacotherapy, 2003. **23**(5): p. 686-9.

140. Lawson, D.H., V. Sherman, and J. Hollowell, *The General Practice Research Database. Scientific and Ethical Advisory Group.* Qjm, 1998. **91**(6): p. 445-52.

141. Walley, T. and A. Mantgani, *The UK General Practice Research Database.* Lancet, 1997. **350**(9084): p. 1097-9.

142. Lis, Y. and R.D. Mann, *The VAMP Research multi-purpose database in the U.K.* J Clin Epidemiol, 1995. **48**(3): p. 431-43.

143. Hubbard, R., A. Venn, S. Lewis, and J. Britton, *Lung cancer and cryptogenic fibrosing alveolitis. A population-based cohort study.* Am J Respir Crit Care Med, 2000. **161**(1): p. 5-8.

144. Lewis, J.D., C. Brensinger, W.B. Bilker, and B.L. Strom, *Validity and completeness of the General Practice Research Database for studies of inflammatory bowel disease.* Pharmacoepidemiol Drug Saf, 2002. **11**(3): p. 211-8.

145. van Staa, T.P., L. Abenhaim, C. Cooper, B. Zhang, and H.G.M. Leufkens, *The use of a large pharmocoepidemiological database to study exposure to oral corticosteroids and risk of fractures: validation of study population and results.* Pharmacoepidemiol & Drug Safety, 2000. **9**: p. 359-366.

146. Jick, H., B.Z. Terris, L.E. Derby, and S. Jick, *Further validation of information recorded on a General Practitioner Based Computerized Data Resource in the United Kingdom.* Pharmacoepidemiol & Drug Safety, 1992. **1**: p. 347-349.

147. Prentice, R.L., B.J. Williams, and A.V. Peterson, *On the regression analysis of multivariate failure time data.* Biometrika, 1981. **68**: p. 373-379.

148. van Staa, T.P., E.M. Dennison, H.G. Leufkens, and C. Cooper, *Epidemiology of fractures in England and Wales.* Bone, 2001. **29**(6): p. 517-22.

149. West, J., R.F. Logan, P.G. Hill, A. Lloyd, S. Lewis, R. Hubbard, R. Reader, G.K. Holmes, and K.T. Khaw, *Seroprevalence, correlates, and characteristics of undetected coeliac disease in England.* Gut, 2003. **52**(7): p. 960-5.

150. Hallert, C., C. Grant, S. Grehn, C. Granno, S. Hulten, G. Midhagen, M. Strom, H. Svensson, and T. Valdimarsson, *Evidence of poor vitamin status in coeliac patients on a gluten-free diet for 10 years.* Aliment Pharmacol Ther, 2002. **16**(7): p. 1333-9.

151. *Homocysteine and risk of ischemic heart disease and stroke: a meta-analysis.* Jama, 2002. **288**(16): p. 2015-22.

152. Curione, M., M. Barbato, L. De Biase, F. Viola, L. Lo Russo, and E. Cardi, *Prevalence of coeliac disease in idiopathic dilated cardiomyopathy.* Lancet, 1999. **354**(9174): p. 222-3.

153. Hadjivassiliou, M., R. Grunewald, B. Sharrack, D. Sanders, A. Lobo, C. Williamson, N. Woodroofe, N. Wood, and A. Davies-Jones, *Gluten ataxia in perspective: epidemiology, genetic susceptibility and clinical characteristics.* Brain, 2003. **126**(Pt 3): p. 685-91.

154. Jick, H., S. Jick, L.E. Derby, C. Vasilakis, M.W. Myers, and C.R. Meier, *Calcium-channel blockers and risk of cancer.* Lancet, 1997. **349**(9051): p. 525-8.

155. Faggiano, F., T. Partanen, M. Kogevinas, and P. Boffetta, *Socioeconomic differences in cancer incidence and mortality.* IARC Sci Publ, 1997(138): p. 65-176.

156. Manjer, J., R. Kaaks, E. Riboli, and G. Berglund, *Risk of breast cancer in relation to anthropometry, blood pressure, blood lipids and glucose metabolism: a prospective study within the Malmo Preventive Project.* Eur J Cancer Prev, 2001. **10**(1): p. 33-42.

157. Gaard, M., S. Tretli, and P. Urdal, *Risk of breast cancer in relation to blood lipids: a prospective study of 31,209 Norwegian women.* Cancer Causes Control, 1994. **5**(6): p. 501-9.

158. Bingham, S.A., R. Luben, A. Welch, N. Wareham, K.T. Khaw, and N. Day, *Are imprecise methods obscuring a relation between fat and breast cancer?* Lancet, 2003. **362**(9379): p. 212-4.

159. Goodman, G.E., *Lung cancer. 1: prevention of lung cancer.* Thorax, 2002. **57**(11): p. 994-9.

9 Papers published from the work in this thesis

Joe West, Logan RFA, Hill PG, Lloyd CA, Lewis S, Hubbard R, Reader R, Holmes GKT, Khaw KT. Seroprevalence, characteristics and correlates of undetected coeliac disease in England. *Gut* 2003 July;52(7):960-965

Joe West, Logan RFA, Card TR, Smith C, Hubbard R. Fracture risk in people with coeliac disease: a population-based cohort study *Gastroenterology* 2003;125:429-436

Joe West, Richard FA Logan, Tim R Card, Chris Smith, Richard Hubbard. Risk of vascular disease in adults with coeliac disease: a population-based study. *Alimentary Phar and Ther* 2004 Jul 1;20(1):73-9

Joe West, Richard FA Logan, Chris Smith, Richard Hubbard, Tim R Card. Malignancy and mortality in people with coeliac disease: population-based cohort study. *British Medical Journal* 2004 Sep 25;329(7468):716-9. Epub 2004 Jul 21