An Analysis of TB Epidemiology from a Primary Care Perspective Using the General Practice Research Database

by

Sarah R. Anderson

MA MB BChir MRCGP FFPHM

A thesis submitted for the degree of Doctor of Medicine (Research)

University College London

2011

Department of Primary Care and Population Health

University College London

Declaration

I, Sarah Anderson confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

Background

TB is a growing public health problem. Most research focuses on secondary care. This, the largest study of TB in primary care to date, aims to investigate aspects of TB epidemiology by examining the role of primary care in case detection, management and outcomes.

Methods

Using data from the General Practice Research Database a case-control and a self-controlled study were undertaken to investigate TB pre-diagnosis and a self-controlled and cohort study to describe and quantify TB morbidity and mortality post-diagnosis.

Results

3032 TB cases and 15,160 matched controls were analysed. The pre-TB studies showed that cases had higher consultation rates, consulted on multiple occasions often with recurrent respiratory infections and experienced a diagnostic delay of up-to-4months. Strengths of association between symptoms, respiratory diseases and TB were also noted.

The case-control study also showed that male gender, smoking and chronic disease were independent risk factors for TB.

Post-TB, patients who survived at least 18months had no increase in morbidity but TB led to a slight increase in all-cause mortality (adjusted HR:1.53) and a case fatality rate at one year of 6.5%. Mortality was highest in the first year post-diagnosis. Undiagnosed pulmonary TB, older age, male gender, malignancy and chronic disease were all

associated with a worse prognosis. At 10 years, TB cases had a 70% chance of survival compared to 79% for patients' without-TB.

Conclusions

Diagnosing TB in primary care remains a challenge due to its variable and non-specific nature but a window of opportunity exists from 4months following first presentation. Combining knowledge of the strength of association between TB, its symptoms and respiratory diagnoses, TB risk factors and raising awareness of newer investigations and local referral processes could help GPs make an earlier diagnosis.

This research indicates a need for improved patient and professional awareness of TB and emphasises the need for earlier diagnosis.

Table of Contents

List of Figures

List of Tables

Acknowledgements

I wish to express my deepest thanks to the many people who have helped me undertake this research. I truly appreciate the role that they - and others not mentioned here - have played; and without whose support, encouragement and example I would not have completed this work.

I would especially like to thank my supervisors Professor Anne Johnson and Dr Greta Rait: Anne for her clarity of thought, her constructively critical eye and her support in developing my epidemiological thinking; and Greta for her guidance, encouragement and invaluable discussions as well as her primary care knowledge that helped ground this work.

In addition, I would like to thank:

- Dr Andrew Hayward, for helping initiate this research and his guidance as the research progressed.
- Shahed Murad, Mark Griffin and Sharon See Tai for their statistical advice and help in assisting me understand the sometimes complex world of statistics.
- Irene Peterson and Amir Islam for their early assistance in working with the GPRD.
- The British Infection Society for their fellowship that made the initiation of this research possible.
- My family for all their constant support and encouragement.

List of Abbreviations

A&E	Accident & Emergency Dept
BCG	Bacillus Calmette-Guérin vaccine against TB
BNF	British National Formulary
CI	Confidence interval
COPD	Chronic Obstructive Pulmonary Disease
CXR	Chest X-ray
ETS	Enhanced TB surveillance
GEE	Generalised Estimating Equation Model
GP	General Practitioner
GPRD	General Practice Research Database
HCW	Health Care Worker
HIV	Human Immunodeficiency Virus
HPA	Health Protection Agency
HR	Hazard Ratio
HSD	Health Service Delay
ICD-10	International Classification of Diseases - version 10
IGRA	Interferon-gamma release assays for TB
IRR	Incidence Rate Ratio
IVDU	Intravenous drug user
LRTI	Lower respiratory tract infection
MDR-TB	Multidrug resistant TB
MHRA	Medicines and Healthcare products Regulatory Agency
NOIDs	Statutory notifications of infectious diseases
ONS	Office of National Statistics
OPCS-4	Surgical Operations and Procedures - version 4
OR	Odds Ratio
OXMIS	Oxford Medical Information System
PCPS	Primary Care and Population Sciences Dept - UCL
PCT	Primary Care Trust
PD	Patient Delay
PPV	Positive Predictive Value
PYAR	Person-years-at-risk
RR	Rate Ratio
SD	Standard Deviation
SEAG	Scientific and Ethical Advisory Board
STATA	Statistical Software package from Statacorp
TB	Tuberculosis
TDD	Total Diagnostic Delay
UTS	Up-to-standard data
WHO	World Health Organization

Chapter 1

Introduction and Background

1.1 Introduction

Tuberculosis (TB) is a potentially fatal infectious disease caused by the bacterium *Mycobacterium tuberculosis*. It is a worldwide public health emergency.[1] Rates are rising once again in the UK,[2] (after having fallen dramatically through most of the 20th century) and ways to improve TB prevention and control are being debated.[2-8]

Epidemiology is "the study of the distribution and determinants of health-related states or events in specified populations, and the application of this study to the control of health problems."[9] Most published studies on TB focus on secondary care and the role of primary care is poorly understood.[8, 10-13] To improve TB prevention and control in the UK a better understanding of the epidemiology of TB from a primary care perspective is essential. In the UK, sick patients usually consult their GP first and are likely to consult on a number of occasions prior to a diagnosis of TB because TB is often an insidious disease.[14] Primary care therefore has an important role in trying to diagnose and treat TB early and so reduce ongoing disease transmission. To diagnose early it is important to understand the epidemiology of TB, the diagnostic process in primary care and the relationship between the onset of symptomatic disease and the start of treatment.

Late diagnosis of active TB is thought to be responsible for increased transmission, poor outcomes and an increased risk of death.[15-18] Research on diagnostic delay, classified as a delay from symptom onset to diagnosis or start of treatment,[19, 20] has been conducted and shows large variation.[19-25]

Long-term morbidity following a diagnosis of active TB has been little studied and mortality rates in TB patients vary considerably.[16, 26-33] Analysing TB-patient deaths, and assessing those attributable to TB and those from other causes, as well as risk factors for death should increase our understanding of the epidemiology of TB. In

addition, it could help improve the care and outcomes of TB patients and thus reduce this potentially preventable cause of death.

In the UK, TB remains mostly managed by secondary care. Although a diagnosis may be suspected in primary care, confirmatory investigations tend to occur in secondary care followed by prolonged treatment. An understanding of the pre-diagnostic period and the role of primary care is crucial to controlling TB in the UK.

A better understanding of TB epidemiology in primary care could be used to inform ways to diagnose TB earlier and improve patient management and outcomes. My thesis therefore aims to investigate TB from a primary care perspective and to study different aspects of TB epidemiology, control and prevention by assessing primary care consultations, TB morbidity and mortality.

1.2 Aims and Objectives

1.2.1 Aims

The overall aim of this research was to investigate the epidemiology of TB from a primary care perspective by examining the role of primary care in detection, management and patient outcomes.

1.2.2 Objectives

The objectives of this research were:

1. To investigate consultation behaviour in primary care using a case-control study prior to a diagnosis of active TB and to assess diagnostic delay.
2. To investigate consultation behaviour in primary care using a self controlled study to minimise the impact of confounders, such as ethnicity and social status, and compare the results of this to the case-control study
3. To describe and quantify morbidity from TB in primary care.
4. To describe and quantify mortality from TB.
5. To describe risk factors for death in TB patients.

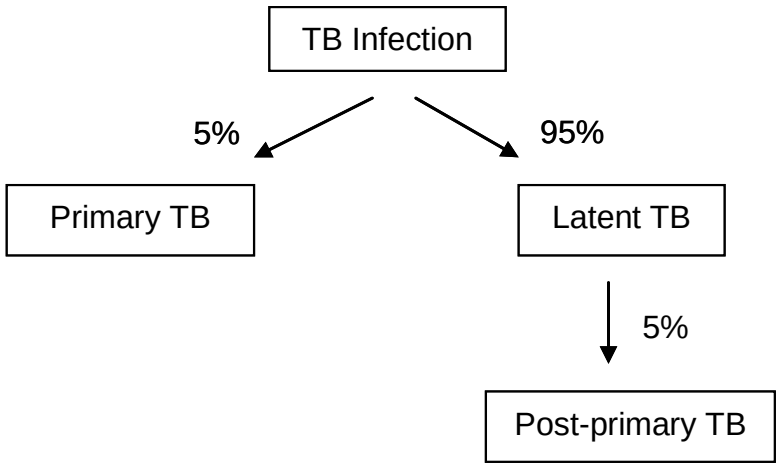
1.3 Overview of TB

1.3.1 The Natural History of TB

Tuberculosis (TB) is a potentially fatal infectious disease caused by the bacterium *Mycobacterium tuberculosis*. It is transmitted via the inhalation of infected respiratory droplets coughed or sneezed by a person infected with TB.[34] It most commonly infects the lungs, but can affect any part of the body.

Once a patient is infected with *Mycobacterium tuberculosis* they have 'TB infection' for life (unless treated) and their lifetime risk of developing 'active' or symptomatic TB is approximately 10%.[14, 35, 36] Progression from TB infection to TB disease (a symptomatic state) occurs when the *Mycobacterium tuberculosis* bacilli overcome the patient's immune response and multiply leading to symptomatic disease. A patient with TB infection usually remains asymptomatic and undiagnosed; however about 5% go on to develop primary TB disease. 'Primary TB' is when TB infection progresses to symptomatic TB disease in a relatively short period of time (a few weeks to months).[14, 37, 38] Another 5% of patients go on to develop post-primary or reactivation TB often many years later.[14, 37] 'Post-primary' or reactivation TB predominantly affects the lungs and symptoms once started usually develop quickly and can be both respiratory and constitutional.[6, 14]

Figure 1.1 The natural history of TB



After infection, *Mycobacterium tuberculosis* may remain dormant for many years before causing post-primary or reactivation TB. This dormant state is known as latent infection or latent TB. 'Latent TB' is defined as the absence of clinical disease associated with a positive tuberculin skin test (a diagnostic test for TB), a positive interferon-gamma blood test or a chest x-ray that shows the scars of old TB (i.e. the patient is well and has no symptoms or signs of TB disease).[39-42] In contrast to individuals with active/symptomatic (primary or post-primary) TB those with latent TB do not transmit the infection and so pose no public health risk. The likelihood of progression from latent to active TB is highest in the first year after infection.[40, 41, 43] Host factors such as age, immunosuppression and other medical problems (e.g. HIV infection) increase the chance of developing active TB.[43]

1.3.2 Transmission

TB is not highly infectious in everyday circumstances. Prolonged close contact with an infectious patient is usually necessary for transmission, for example, living in the same household as an infectious patient.[34, 44, 45] It has been estimated that an infectious patient (i.e. a patient whose sputum-smear is positive for *Mycobacterium tuberculosis* bacilli) will infect on average 10-15 people each year and over 20 people during the natural course of untreated disease.[4,46] The risk of a contact becoming infected with TB depends on the nature and duration of their exposure; in close household contacts the risk is considered to be 1 in 3 and for casual social contacts 1 in 100,000.[4,34,39] Transmission tends to occur from people with active, and infectious pulmonary or laryngeal TB, not latent TB or extra-pulmonary TB.

1.3.3 Clinical Presentation

Nearly two-thirds (62%) of patients with active TB, in England and Wales, had pulmonary TB in 1998[2] and 56% in 2007.[47] Cough is the most common respiratory symptom and is usually present for more than 2-3 weeks;[14] with haemoptysis in a small minority of cases.[6] The classic symptoms of pulmonary TB (active TB disease affecting the lungs) include: chronic cough, fever, lymphadenopathy, weight loss, night sweats and haemoptysis.[14] Not all forms of pulmonary TB are infectious. TB patients with "smear positive" TB are infectious and this state is defined by the fact that bacteria can be seen on microscopic examination of the patient's sputum. TB can affect other parts

of the body, this is known as extra-pulmonary TB (for example TB of the kidney or spine). Symptoms depend on the organ affected and intermittent fever or weight loss are often also present.[6] Persistent lymphadenopathy, present for longer than 4 weeks, in a non-white individual should be considered TB until proven otherwise.[6] Extra-pulmonary TB may present with systemic symptoms or site-specific symptoms.[48] Weight loss, fever and night sweats are particularly associated with disseminated (including miliary)[38] and gastro-intestinal TB. However, fever and night sweats are less common in other extra-pulmonary TB cases (e.g. peripheral lymph nodes, skin, bone and joint, genitourinary TB). Nearly half of all extra-pulmonary TB occurs in peripheral lymph nodes.[14]

1.3.4 Diagnosis

A diagnosis of TB can be confirmed by finding *Mycobacterium tuberculosis* in a sputum[49] or biopsy specimen[50] (either microscopically or on culture); and is supported by an abnormal chest x-ray,[51,52] a positive tuberculin skin test (Heaf or Mantoux)[40,41,53] or more recently, for latent infection, a positive interferon-gamma test.[14,42,54,55]

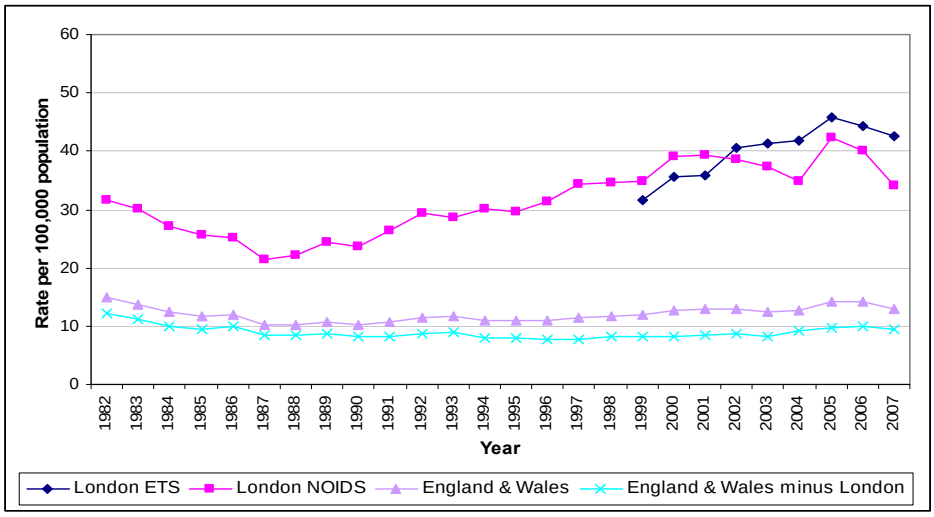
1.3.5 Treatment

TB is curable if treated early with appropriate drugs and for a minimum of six months.[5, 55-57] During treatment a patient takes 4 drugs: rifampicin, isoniazid, pyrazinamide and ethambutol for 2 months then rifampicin and isoniazid alone for the subsequent 4 months. A minority of cases are resistant to at least one standard drug. For these patients different treatment regimes and duration are used. Patients with pulmonary TB are considered non-infectious after two weeks of treatment.

1.4 The Epidemiology of TB in the UK

TB is a notifiable disease, i.e. it is a statutory requirement that cases of active TB are notified to the 'Proper Officer' for Health, the local Consultant in Communicable Disease Control. Data on notified TB cases are collected via Enhanced Tuberculosis Surveillance (ETS).[58,1] TB nurses and respiratory consultants complete ETS forms which are then collated and analysed by the Health Protection Agency Centre for Infections. ETS data is linked to other datasets such as those from the mycobacterial surveillance network (MycobNet) that collates microbiological data on *Mycobacterium tuberculosis* complex isolates and their drug susceptibilities from reference laboratories across the UK.

Figure 1.2 TB notifications in London, England and Wales 1982-2007



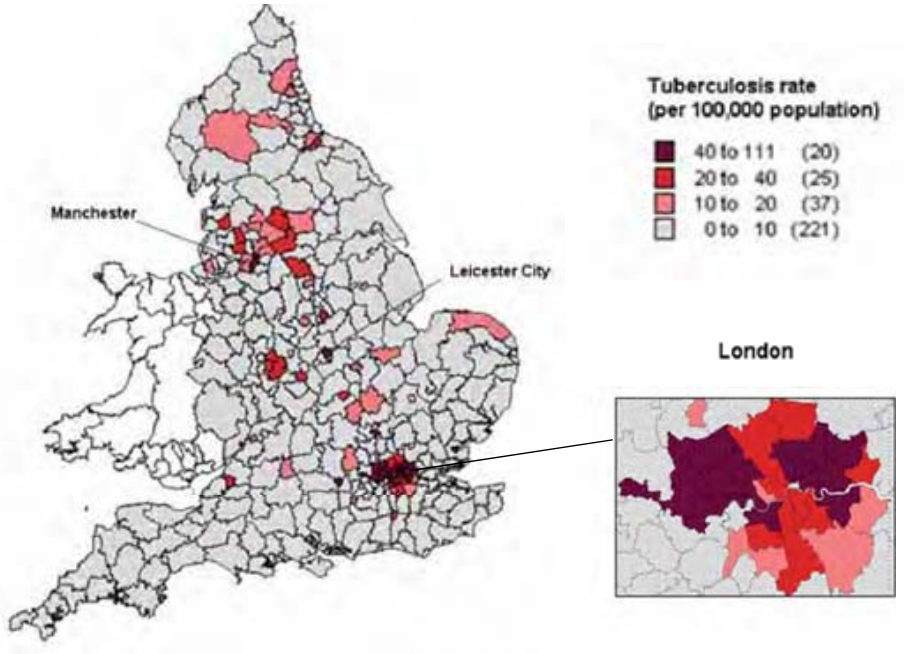
ETS = Enhanced TB Surveillance Data
NOIDs = statutory notifications of infectious diseases

---

[1] **Enhanced Tuberculosis Surveillance** is a surveillance system established in 1999 in England and Wales to collect more detailed and timely information on patients with TB. It provides an annual corrected analysis of TB case reports by age, sex, ethnic group, country of birth, site of disease and NHS region.

In 2007, 8417 people were newly diagnosed with active TB in the UK. This represents an incidence rate of 13.8 per 100,000 population.[47] Since the mid-1980's there has been a steady rise in the number of UK TB cases (see Figure 1.2). TB remains primarily a disease linked to poverty and social exclusion.[59] Contributing to the rise have been changing patterns of immigration,[60] increased homelessness[3,60,61] and HIV infection[60] as well as an ageing population. In 2007 the majority of cases were reported in England (92%) with London accounting for 39% (a rate of 43.2 per 100,000).[47] Published research shows there is significant local variation in TB incidence across London[62] and that it is especially high in new immigrants who have recently entered the UK.[2,63] An analysis in 2003, showed that the majority (85%) of Primary Care Trusts (PCTs) had less than 20 cases per 100,000; 7% (24 PCTs) had a rate of ≥40 per 100,000 and most of these were in London.[64]

Figure 1.3 Rates of TB in England by Local Authority, 2007



Source: Tuberculosis in the UK: Annual report on tuberculosis surveillance in the UK 2008 [47]

In 2007, 56% of UK TB cases had pulmonary TB and more than half of all TB cases were male (55%).[47] TB occurs in all age groups but the majority (60%) of cases were diagnosed in people aged 15-44 years; 6% were aged 0-14 years, 19% aged 45-64 years and 14% aged over 65 years.[47] Seventy two percent of those diagnosed with TB were born abroad, predominantly in South Asia and sub-Saharan Africa.[47] The ethnic groups

with the highest rates of TB are Black Africans (309 per 100,000), followed by Indian, Pakistani and Bangladeshi (212 per 100,000).[47] UK surveillance data shows that about half of all TB patients who are born abroad were diagnosed with TB within the first five years of first entering the UK.[47]

In 2007, 7.4% of TB cases in the UK had resistance to at least one first line drug and 1.2% were multidrug-resistant (resistant to at least isoniazid and rifampicin).[47] This represents a small increase in resistance over a 10 year period to at least one first line drug and multidrug-resistance.[2, 65]

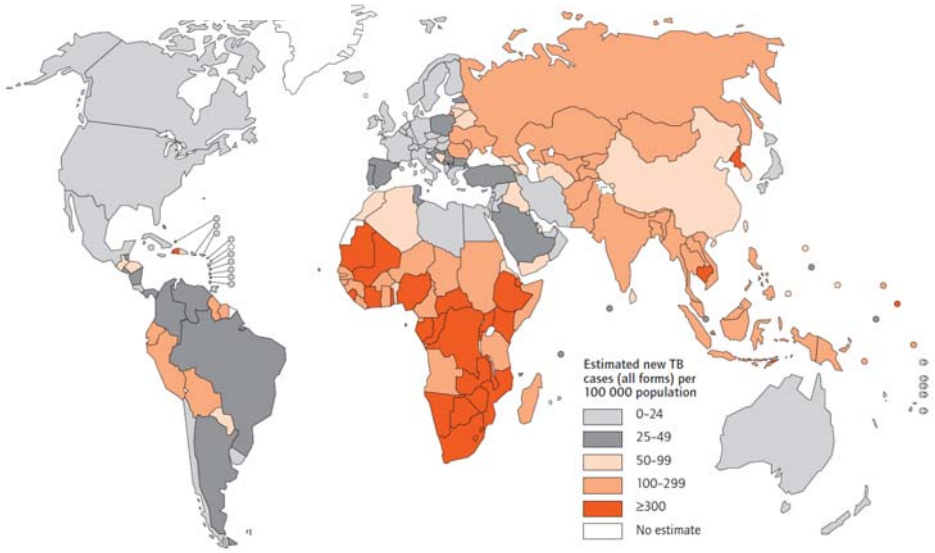
Co-infection of TB patients with HIV is a small but increasing problem in the UK. In 1993, 2.2% of TB patients were co-infected in England and Wales; by 1998 this had risen to 3.3%[66] and by 2003 to 8.3%.[67]

1.5 Public Health Issues and TB

1.5.1 The Rising Incidence of TB

TB is an international public health problem.[68] In 1993 the World Health Organization (WHO) declared TB a global emergency. WHO estimated that in the year 2000 approximately a third of the world's population, 2 billion people, were infected with *Mycobacterium tuberculosis*, that 8.3 million new cases of TB were diagnosed and that 1.8 million people died from the disease.[69] In 2007, 9.7 million new cases of TB were identified.[70] Much of the rising global incidence of TB is attributable to the spread of HIV in Africa.[68]

Figure 1.4 Estimated TB incidence by country 2007



Source: WHO report 2009 – Global tuberculosis control: Epidemiology, strategy, financing.[70]

The incidence of TB in the United Kingdom remains relatively low by global standards. Throughout most of the 20th century the incidence of TB declined in England and Wales;[71,72] but a rise in incidence since the mid-1980's[2,58] and high rates of TB in certain sub-groups now make TB an increasingly important public health problem.[3,47,60,72,73]

In the last 10 years England and Wales have seen a 25% rise in TB case numbers.[4] The distribution of TB in England and Wales varies markedly by geographical area.[73] Rates are highest in urban areas, particularly London,[2,7] whereas a continued decline has been seen in most rural areas.[58] Ethnic minorities contribute a substantial and increasing proportion of TB cases.[2,47,58,73]

1.5.2 The Risk of TB Transmission

The longer an active case of TB goes undetected, due to a delay in diagnosis, the higher the risk of transmission to close contacts. Research suggests that a patient with infectious TB will infect on average 10-15 people per year (i.e. 5-7 people if diagnosis is delayed by 6 months).[4,46] See section 1.3.2.

1.6 The Organisation of TB Services in the UK

The majority of patients who become unwell in the UK present to a GP in primary care. A GP's role includes taking a thorough history, examining the patient and ordering baseline investigations. If a GP suspects TB they either initiate investigation themselves or they refer the patient to the TB Service at their local hospital. The local TB Service then undertakes further investigations: such as, blood tests, sputum samples and a chest x-ray. If TB is confirmed a respiratory consultant reviews the patient and prescribes treatment. Standard TB treatment lasts for 6 months, this is often prescribed and dispensed by the hospital and patients return to hospital once a month throughout their treatment for follow-up usually with a TB nurse. In some areas of the UK once a patient is stabilised on TB treatment they can see their GP for regular review and prescriptions.

If the patient is extremely unwell or does not have a GP (because they are homeless or a newly arrived immigrant who has not yet registered with a GP) it would be usual for the patient to present via the Accident and Emergency department of their local hospital. If necessary, patients are admitted to hospital and see the TB team while an in-patient or alternatively are sent home with a referral to the hospital TB service, in the same way as a patient referred via their GP.

As TB is an infectious disease, with the potential to infect others, part of the UK TB Service's role is to investigate and screen contacts of a TB case that may be at risk of becoming infected. This is known as 'contact tracing'. This role is often undertaken by TB nurses either at the hospital or in the community. In the UK there is no systematic TB screening programme.

The majority of patients with active/symptomatic TB (as with most diseases) will first make contact with primary care; primary care therefore has an important role in diagnosing TB early.

1.7 Control and Prevention of TB

The most important part of TB control and prevention is the identification and treatment of those who already have the disease so as to shorten their infectious period and to halt onward transmission.

TB can be controlled by:

- prompt diagnosis and treatment of people with active disease to achieve cure and so prevent transmission
- ensuring TB patients complete their treatment
- identifying and treating people with early infection / latent TB (i.e. through contact tracing or new entrant screening)
- prevention through BCG immunisation

Due to the changing epidemiology of TB in the UK, there has been increased concern and debate as to the effectiveness and adequacy of current TB control and prevention measures.[2,3,7, 74] A better understanding of TB from a primary care perspective would be useful so that improved control and prevention measures could be taken in this setting.

1.8 The GPRD as a Data Source for Research

The General Practice Research Database (GPRD) is the world's largest computerised primary care database.[75,76] Currently it contains over 20 years of longitudinal data (1987-2009) on approximately 13 million patients from approximately 500 practices covering 5.5% of the UK population.[77] Twelve percent of GPRD GP practices are based in London and a third in urban areas (personal communication – Mary Thompson, EPIC). The GPRD is considered broadly representative of the UK population in terms of age and sex structure[76] and has been widely used for epidemiological research.[78-83] Several validation studies have shown that there is a high degree of completeness and reliability in the data contained within the GPRD[76, 80, 84-88] and there is good agreement between the GPRD prescribing data and national data from the Prescription Pricing Authority.[76,85] Data quality is high as data from practices are routinely validated by internal checks,[85] and only data meeting a minimum preset standard are added to the database.[75]

Anonymised patient data are collected from participating practices. Practices use standard software to record patient and consultation information and periodically down-load anonymised patient information to a central computer system managed by the

UK's Medicines and Healthcare Products Regulatory Agency (MHRA).[77] All information for individual patients is handled using a unique identification number.[75] GPs record medical diagnoses and symptoms using Read and Oxford Medical Information System (OXMIS) codes; these codes are based on, and can be mapped to, the internationally recognised classification systems of Disease (ICD-10) and Surgical Operations and Procedures (OPCS-4). Data for each patient includes: age, sex, medical diagnosis and symptom records, details of each patient consultation, medication prescribed, a summary of specialists' clinical notes and hospital letters, results of laboratory tests and a free-text section.[75, 77]

The GPRD is a powerful tool for epidemiological and health service research from which a substantial number of studies in peer-reviewed literature have been generated.[75] It has been validated for use in respiratory epidemiological studies by Hansell[80] who used the GPRD to assess 11 respiratory diagnoses, TB being one of these. She found a TB rate of 1.4 patients /10,000 patient years plus good agreement between the GPRD and the 4th Morbidity Survey in General Practice.[80,89] The GPRD has been used for many high quality studies including those in respiratory disease.[78, 80, 83, 90-92]

The GPRD is an extremely useful tool for research purposes. Its strengths lie in its:

1. Large size (13 million patients)[77]
2. Continuous, longitudinal primary care data (now 20 years of data)[77]
3. Comprehensive coverage of UK population[76]
4. High-quality validated data - reliable and with a high degree of completeness[85, 93]
5. Substantial statistical power - so it is a unique tool for rare diseases or events
6. Instantly available as it is already collected and processed
7. Ability to link prescriptions to diagnoses[77]
8. Allows investigation into long-term morbidity and mortality[79, 80, 83, 92]
9. Use in epidemiological studies[80, 92]

It does however have some limitations, these are:

1. Primarily a clinical database of real-time patient notes so requires intensive data management and cleaning to make suitable for statistical analyses
2. Complexity of data recording and therefore analysis[93]
3. Consists of observational data
4. No requirement to enter diagnoses for minor consultations
5. Variation in GP recording
6. Limited secondary care data
7. Free-text fields not available unless requested
8. No ethnicity data or social class data during the analysis period
9. Many similar but different terms coded but not linked to over-arching diagnosis
10. Geographic distribution of GP practices

1.9 Overview of Thesis

This thesis consists of a number of related studies of TB in the UK using data from the GPRD. The studies use classical epidemiological methods (case-control and cohort methodology) and a more recently developed but well established method, self-controlled methodology, to assess TB from a primary care perspective analysing consultation behaviour prior to a diagnosis of TB and morbidity and mortality following TB.

Chapter 1 - Introduction: This sets the scene, reviewing TB as a disease, its epidemiology, its public health consequences and how TB Services are run in the UK. In addition, it gives an overview of the GPRD dataset, which was used to undertake the research for this thesis.

Chapter 2 - Literature Review: This chapter details the literature reviews on 'TB from a primary care perspective' and 'TB and diagnostic delay'.

Chapter 3 - Methods: This chapter details the study design and methods, the dataset, data cleaning and the statistical analysis undertaken.

Chapters 4 to 7 - Detail the different epidemiological studies used to assess TB from a primary care perspective: two describe TB epidemiology pre-TB, a third looks at morbidity post-TB and a fourth looks at mortality post-TB. All these studies try to increase our understanding of TB from a primary care perspective. The first two assess how TB patients present and consider how to reduce diagnostic delay and the latter two try to assess clinical outcomes.

Chapter 8 - Discussion: This aims to draw the findings presented in Chapters 4 to 7 together in a discussion, looking at what the implications of these research findings are for primary care and TB in the UK and then documents future steps.

In summary, the overall aims of this thesis are to investigate TB from a primary care perspective, to understand the role of primary care in TB case detection and management, review the problem of late diagnosis and assess the long-term outcomes for patients in terms of morbidity and mortality. On the basis of the evidence, this thesis will then set out proposals for improving TB prevention and control in the primary care setting.

Chapter 2

Literature Review: TB and Primary Care and TB and Diagnostic Delay

2.1 Introduction

Late diagnosis of TB is thought to be responsible for increased disease transmission,[17] poor patient outcomes and an increased risk of mortality.[15,16,18,68,94] Research on diagnostic delay (classified as a delay from symptom onset to diagnosis or start of treatment)[19,20] has been conducted and shows considerable variation in the length of such delay.[15-20,23,24] Early diagnosis & treatment of TB are essential for an effective TB control programme. Delays need to be minimised where ever possible to reduce the risk of ongoing transmission and to improve clinical outcomes. Most transmission occurs prior to treatment when the patient is symptomatic with cough. Research has shown patients remain infectious for as-long as treatment is delayed.[95]

2.2 Aims

The aims of this literature review were to assess:

1. the role of primary care in TB diagnosis
2. the literature on diagnostic delay for TB patients: to summarise what is known about diagnostic delay and to quantify this delay.

2.3 Method

2.3.1 Literature Retrieval

A systematic review of computerised databases from 1980 to 2009 was carried out using Medline, EMBASE, CINAHL, HMIC, the Cochrane Database of Systematic Reviews and the Cochrane Central Register of Controlled Trials. Standard MeSH headings were selected using the PubMed MeSH database supplied by the US National Institute of Health in its digital archive of bio-med journals. The following MeSH headings and keywords were used; tuberculosis, primary care (primary care, general practice, family practice, family physician, family practitioner, general practitioner), diagnostic delay (diagnos*, delay*, late, patient delays, health care delays). The search strategies are detailed in Appendix 1. All searches were limited to studies in humans and English language articles. Articles from low income countries were excluded and those from health systems very different to that of the UK. The reference lists of articles identified by the systematic review were also reviewed for additional references. The most recent editions of relevant journals (BMJ, Thorax, Lancet, NEJM, JAMA, American Journal of Respiratory and Critical Care Medicine and others) were hand searched at the time of the literature review (2009) to identify relevant articles not yet entered in the electronic databases. Original papers were retrieved and critically appraised using the Critical Appraisal Tools developed by the Critical Appraisal Skills Programme (CASP) based at the Public Health Research Unit in Oxford.
[http://www.phru.nhs.uk/casp/critical_appraisal_tools.htm]

A summary of the search strategies is shown in Table 2.1 and 2.2.

Table 2.1 Search Strategy for TB and primary care literature

Database	Search terms	Relevant Results
Medline: 1980 – Dec 2009	Tuberculosis AND primary care, general practice, family practice, family practitioner, general practitioner, family physician ALL SUBHEADINGS, LIMITED to English language & human	**185**
EMBASE: 1980 - Dec 2009	Tuberculosis AND primary care, general practice, family practice, family practitioner, general practitioner, family physician ALL SUBHEADINGS, LIMITED to English language & human	**271**
CINAHL: 1981 **–** Dec 2009	Tuberculosis AND primary care, general practice, family practice, family practitioner, general practitioner, family physician LIMITED to English language	**64**
HMIC (all years)	Tuberculosis AND primary care, general practice, family practice, family practitioner, general practitioner, family physician	**8**
Cochrane Database of Systematic Reviews (all years)	Tuberculosis AND primary care	**0**
Cochrane Central Register of Controlled Trials (all years)	Tuberculosis AND primary care	**2**

Table 2.2 Search Strategy for TB and diagnostic delay in the UK

Database	Search terms	Relevant Results
Medline & EMBASE 1980 – June 2009	Tuberculosis AND diagnos*, delay*, late ALL SUBHEADINGS, LIMITED to English language & human	**596**
HMIC (all years)	Tuberculosis AND diagnos*, delay*, late	**4**
CINAHL (1981 **–** June 2009)	Tuberculosis AND diagnostic delay, treatment delay LIMITED to English language	**1**
Cochrane Database of Systematic Reviews (all years)	Tuberculosis AND diagnostic delay, treatment delay	**0**
Cochrane Central Register of Controlled Trials (all years)	Tuberculosis AND diagnostic delay, treatment delay	**0**

2.3.2 Selection of Relevant Research Papers

To review the literature on 'TB and primary care' and 'TB and diagnostic delay' I used a structured approach to formulating my search questions,[96] a technique recommended by the Evidence Based Medicine Working Group.[97] This technique helped focus my research questions by linking a patient group with an intervention (exposure) and an outcome (see table 2.3) and enabled me to generate key search terms for use with bibliographic databases.

The specific questions were:

- What is the role of primary care / the GP in diagnosing TB?
- What is the delay from symptoms to diagnosis experienced by TB patients?
- What are the factors associated with, or causes of, diagnostic delay for TB patients?

Table 2.3 Structured search questions on diagnostic delay

Question	Patient group / population	Intervention	Outcome / Outcome measure
What is the delay for patients diagnosed with TB?	TB patients	• Diagnosis (missed diag, delayed diag, late diag) • Primary care • Secondary care	• Earlier diagnosis • Earlier referral • Less delay in diagnosis
What are the causes of diagnostic delay for patients with TB?	TB patients	• Diagnosis (missed diag, delayed diag, late diag)	• List of causes
What is the role of primary care / the GP in diagnosing TB?	TB patients	• Primary care	• Diagnosis • Referral
What is the role of secondary care in diagnosing TB?	TB patients	• Secondary care • Hospital	• Diagnosis • Referral

2.3.3 Study Exclusion and Inclusion Criteria

Studies were included if they:

1. Focused on TB
2. Included patients as individuals or groups
3. Mentioned primary care or primary care practitioners
4. And if they included my key search terms to answer my structured questions

Studies were excluded if they:

1. Did not involve humans
2. Were not published in English
3. Were based in countries dissimilar to the UK where TB prevalence is high and country income is low

2.3.4 Assessment of the Strength of the Evidence

I graded the evidence using a system developed by the Scottish Intercollegiate Guidelines Network (SIGN)[98,99] and used by the National Institute for Clinical Excellence (NICE) (see Table 2.4). I used a standardised data collection form to extract data including: year, author, country, study type, sample size, factors increasing and decreasing delay, quantification of delay and any comments.

Table 2.4 Criteria for grading evidence and recommendations [99]

Table 2 Hierarchy of evidence and recommendation classification

Levels of evidence		Classification of recommendations	
Level	Type of evidence	Class	Evidence
1++	High-quality meta-analysis (MA), systematic reviews (SR) of randomised controlled trials (RCTs), or RCTs with a very low risk of bias.	A	Level 1++ and directly applicable to the target population *or* level 1+ and directly applicable to the target population AND consistency of results. Evidence from NICE technology appraisal.
1+	Well-conducted MA, SR or RCTs, or RCTs with a low risk of bias.		
1–	MA, SR of RCTs, or RCTs with a high risk of bias.	Not used as a basis for making a recommendation.	
2++	High-quality SR of case-control or cohort studies. High-quality case-control or cohort studies with a very low risk of confounding, bias or chance and a high probability that the relationship is causal.	B	Level 2++, directly applicable to the target population and demonstrating overall consistency of results. *or* extrapolated evidence from 1++ or 1+.
2+	Well-conducted case-control or cohort studies with a low risk of confounding, bias or chance and a moderate probability that the relationship is causal.		
2–	Case-control or cohort studies with a high risk of confounding, bias or chance and a significant risk that the relationship is not causal	Not used as a basis for making a recommendation.	
3	Non-analytic studies (for example case reports, case series).	C	Level 2+, directly applicable to the target population and demonstrating overall consistency of results *or* extrapolated evidence from 2++.
4	Expert opinion, formal consensus.	D	Level 3 or 4 *or* extrapolated from 2+ *or* formal consensus *or* extrapolated from level 2 clinical evidence supplemented with health economic modelling.
		D (GPP)	A good practice point (GPP) is a recommendation based on the experience of the GDG.

Diagnostic study level of evidence and classification of recommendation was also included.

2.4 Results

2.4.1 The Role of Primary Care in the Diagnosis of Tuberculosis

2.4.1.1 Literature Yield

The key words used in my search strategy generated a pool of nearly 500,000 references. Individual key words were then combined to narrow the search to references most likely to be relevant, so reducing the number of papers to 548. The titles and abstracts of these papers were then reviewed and 30 papers met my broad inclusion criteria (see Section 2.3.3). Ten articles[8, 10-13, 20, 100-103] covered TB and primary care in a way relevant to my research question "what is known about the role of primary care in diagnosing symptomatic TB". These were reviewed and are discussed here. Six papers looked at TB in primary care from a new entrant / new registrant screening perspective rather than that of new symptomatic disease[63, 100, 104-107] these were excluded, as were a further twelve papers that covered TB and primary care in an international setting that was not directly relevant to the UK situation.

2.4.1.2 The Research Evidence

There is very little published on the diagnosis of TB in primary care. Of the ten articles that met the inclusion criteria, only one was a study specifically looking at the diagnosis of TB from a primary care angle,[8] one assessed diagnostic delay in primary care,[20] three were editorials,[11-13] three were letters[10,101,103] and the remainder just mentioned TB in a primary care setting. It is important to note that editorials and letters are opinions rather than objective evidence and as such would not normally be included in a literature review. However, due to the lack of published peer-reviewed research I felt it was appropriate to include them.

A paper by Metcalf et al. (2007)[8] was the only study in this systematic literature review that specifically explored TB from a primary care perspective in the UK. This small qualitative study assessed the process of diagnosing TB in primary care by using semi-structured interviews with 17 patients and 16 GPs and then analysed the data thematically. Metcalf highlights the huge variety of TB presentations seen by GPs and the difficulties they face in making a timely diagnosis. She found that where diagnosis is clear, GPs generally test specifically and refer appropriately. She identified barriers to

prompt diagnosis as: atypical presentations, low level of clinical suspicion of TB by GPs, lack of continuity of care, workload demands that limit time with patients and sub-optimal clinical / patient communication. Overall, this paper highlights the fact that early diagnosis is one of the most important contributions primary care can make to improved TB control and that raising awareness, through improved education among health professionals, is essential. In addition, as part of this study Metcalf searched the databases Medline, Pubmed, EMBASE from 1966 to 2005 using the terms "tuberculosis" and "diagnosis." She found no study explored the process of diagnosing TB in primary care in high income countries.

A study by Lewis et al. (2003),[20] assessed the delay in starting TB treatment in East London. It showed a median delay of 18 weeks (range 0-129) from first symptoms to starting treatment; with a median 'patient delay' (from first symptom to first GP visit) of 9 weeks and a median 'healthcare delay' (from first GP assessment to starting TB treatment) of 5 weeks. Patient delay was significantly longer than healthcare delay ($p=0.019$). This suggests there is a need to improve awareness of TB symptoms in both patients and GPs.

Three references were editorials, one by Griffiths (2007),[11] one by Davies (2003)[13] and another by Singh (2002).[12] "The new tuberculosis" by Griffiths[11] comments that raising awareness of TB in general practice is vital if we are to combat the increasing incidence of TB in the UK. "TB in primary care" by Singh[12] reflects on the responsibility of GPs to diagnose TB early, support and treat their patients and consider the wider community. Singh also states that "a low index of suspicion is emphatically required for early identification of TB", and that "GPs should remain alert and vigilant". Davies in "TB: current problems for primary care, a danger of complacency,"[13] confirms Singh's statements. Basically, these three editorials discuss TB and primary care from slightly different angles but all reach the same conclusion that raising awareness of TB in primary care practitioners leads to earlier diagnosis and therefore better TB control.

TB is traditionally seen as a disease managed in secondary care with respiratory consultants diagnosing and treating it. Primary care, however, could be, and should be, diagnosing and managing more TB. As Levin (1993)[108] in the US suggests, primary care physicians have an important role in screening, diagnosing, treating and following up TB cases with the goal of controlling TB. This sentiment is supported by others[8, 10,

[101, 103] (Class D evidence). Three references that specifically looked at TB and primary care were letters: one from MacRorie in 2007[10] stating that "GPs deserved better, clearer guidelines on diagnosis and treatment" of TB and commented that primary care has a major role to play in early diagnosis and ongoing TB patient care; one from Singh[101] reinforcing the role that GPs have to play in diagnosing and controlling TB; and another from MacRorie in 2002[103] arguing similarly to Singh.[101]

A study by White et al. (2002),[102] that looked at the management of TB and the utilisation of resources in a British inner-city population, found that just under half of all TB patients admitted to hospital had not presented via primary care (42%).(Class C evidence)

I have not reviewed in detail papers that my literature review found on "new entrant screening," as this was not the question I sought to answer. One paper by Griffiths[100] however, is worth reviewing as although it discusses screening for TB in people registering in primary care it also looks at educational outreach to primary care staff to promote screening. Griffiths undertook a cluster randomised controlled trial for educational outreach and showed that an educational intervention, promoting TB screening in people registering with primary care, "improved identification of active and latent TB and increased BCG coverage". This is supported by Margolis again with reference to TB,[104] but has also been shown in many areas of medicine, that raising awareness of a disease through a specific educational intervention increases diagnosis and action to combat the disease.

In summary, the role of primary care in the diagnosis of TB is little studied and therefore warrants greater investigation to understand how to diagnosis TB earlier and improve outcomes for patients.

2.4.2 TB and Diagnostic Delay

2.4.2.1 Literature Yield

Over three million references were generated by the key words in my search strategy: tuberculosis, diagnos*, delay*, late, 'diagnostic delay', 'treatment delay' (see Appendix 1 for details). Individual key words were then combined to narrow the search to relevant references. This reduced the number of papers to 596 from Medline and EMBASE, and 601 overall (see Table 2.2). The titles and abstracts of these papers were reviewed and 68 papers met my broad inclusion criteria (see Section 2.3.3). I reviewed each paper and from these and their reference lists found 28 papers that discussed TB and diagnostic delay in countries and health care systems similar to the UK.

2.4.2.2 The Research Evidence

68 of 601 studies identified met my inclusion criteria and were subjected to detailed review. 28 studies were finally included in the analysis, see Table 2.5 for details.[8, 19-25, 95, 109-127]

There is a relatively large body of work looking at TB and diagnostic delay in a variety of international settings. These papers use a variety of definitions for diagnostic delay and assess different risk factors; many are not relevant to the UK setting. I therefore carefully selected 28 studies that met my inclusion criteria and assessed health care systems similar to the UK. The majority of studies were retrospective cohort studies (16); of the remaining studies: seven were prospective studies, two were retrospective analyses of national surveillance data, one was a retrospective case series, one a systematic review and one a qualitative interview study. See Table 2.5 for details. Overall, there is a small amount of research looking at diagnostic delay and TB in the UK and other high income countries and this I review and discuss in more detail.

Diagnostic delay can be defined in a number of ways; the most common definitions use the patient pathway as the starting point. Diagnostic delay can therefore be broken down into a number of stages:

- **Disease Delay** - is defined as delay due to the disease itself i.e. the patient is infected but asymptomatic (has latent TB).

- **Patient Delay** (PD) - is defined as the time interval between the onset of first symptom and the first consultation with a health care provider.[20] It occurs when a patient is unaware of their likely diagnosis, when they fail to seek help early or are unable to access appropriate healthcare.
- **Health Service Delay** (HSD) - is defined as the time interval between the first contact with health services and the start of treatment,[20] sometimes known as 'doctor's delay'.
- **Total Diagnostic Delay** (TDD) is the patient delay plus health service delay (PD + HSD).[19]

Table 2.5 Diagnostic Delay in TB – summary details of the research evidence assessed

Ref No.	First Author	Year	Location	Study type	Grade of Evidence (SIGN)	Sample Size	Factors increasing total delay	Factors decreasing total delay	Comments
115	Altet-Gomez	2003	Spain	Prospective Study	B	287	-	-	• Patients were responsible for >50% of delay • Interview based study
110	Asch	1998	USA (Los Angeles)	Prospective study	B	248	**Factors increasing PD** • uncertainty of where to get care • belief in efficacy of self treatment • unemployment	-	• Only assessed PD • Only assessed symptomatic patients • Showed disease severity had little impact on delay
127	Diez	2004	Spain	Retrospective study	B	7037	**Factors increasing PD** • no resp. symptoms • age >14yrs • female sex if severe PD	**Factors decreasing PD** • cough • extra-pulmonary TB	• Only assessed PD
111	Diez	2005	Spain	Retrospective study	B	5,184	**Factors increasing HSD** • >44yrs • IVDU • cancer & TB • non-TB symptoms • sputum smear -ve TB • female sex if severe delay • diagnosis in primary care, but not statistically signific.	**Factors decreasing HSD** • homelessness • immigrants • HCW • hospital admission • History of contact with TB • cough • LN or CNS TB (i.e. extra-pulmonary TB	• Only assessed culture confirmed cases • Only assessed HSD • Study covered 2/3rds of Spanish TB cases 1996-1997
118	Farah	2006	Norway	Retrospective cohort study	B	83	• extra-pulmonary TB led to approx 2x delay • born in Norway HSD 2x PD	• age>60 shorter PD (but sample size = 8) • immigrants shorter HSD	• HSD > PD for those born in Norway • clearly defines patient & health service delay • National TB registry data combined with clinical case notes from hospital & primary care providers
117	Franco	1996	Spain	Retrospective cohort study	B	109	**Factors increasing PD** • non-specific symptoms	• sputum smear +ve TB • cavities on CXR	• Small local study • Results mostly as %
119	Gagliotti	2006	Italy	Retrospective cohort study	B	271	**Factors increasing PD** • recent migrants **Factors increasing HSD** • 1st consult non-hospital • cough of recent onset or no cough	-	• Low incidence area • 52/48 split in Italian *vs* migrant cases **Conclusions:** • Need to improve migrants access to health services • Increase HCW suspicion of TB

Ref No.	First Author	Year	Location	Study type	Grade of Evidence (SIGN)	Sample Size	Factors increasing total delay	Factors decreasing total delay	Comments
114	Golub	2005	USA	Prospective cohort study	B	158	**TDD** • 1st seen by private physician not A&E or public health **Factors increasing PD** • non-white • less educated • no resp. symptoms **Factors increasing HSD** • age >50yrs • non-English speaking • History of asthma • antibiotic Rx • 1st seen by private physician	**Factors decreasing HSD** • cough • longer PD • sputum smear +ve TB • Investigation at 1st visit (TST/CXR/sputum)	• Assessed only pulmonary TB cases with sputum +ve culture • Did not include those who died at diagnosis
113	Golub	2005	USA	Nested case-control study within prospective study	B	158	• Alternative diagnosis and prescription of antibiotics	• CXR	• Study suggests need more wide spread use of CXR rather than immediate antibiotic for community acquired pneumonia - ?cost effective
95	Golub	2006	USA	Prospective cohort study	B	124	-	-	• Did not assess factors leading to DD but risk factors for TB transmission after >90days delay
122	Greenaway	2002	Canada	Retrospective study	B	429	• older age • HIV infection • non-cavitatory CXR • no cough • sputum smear -ve TB		• Only looked at hospitalised patients with active TB • Did not assess PD and HSD • Clearly showed hospitals with lower rates of TB had longer delays
120	Gulbaran	1996	France (Paris)	Retrospective cohort study	B	52	-	-	• Study conducted in 1 hospital over 18mths • Results given as mean delay not median
125	Ho M.-J.	2004	USA (New York)	Retrospective cohort study	B	60	• traditional health care provider if unaware of national TB programme	-	• Study assessed who symptomatic Chinese immigrants see first • Concludes education of traditional Chinese HC providers needed
124	Kam	2007	Canada	Retrospective cohort study	B	23	• extra-pulmonary TB	-	• Very small study, 23 adolescents over a 6 year period
20	Lewis	2003	UK (London)	Retrospective Case series	C	93	• extra-pulmonary TB	• Pulmonary TB	• large recent immigrant population • TDD= symptom to treatment • PD > HSD
8	Metcalf	2007	UK	Qualitative inductive study	C	17	-	-	• interview based study with thematic analysis • looked at barriers to prompt diagnosis from GP and patient viewpoint

Ref No.	First Author	Year	Location	Study type	Grade of Evidence (SIGN)	Sample Size	Factors increasing total delay	Factors decreasing total delay	Comments
126	Moudgil	1994	UK (Lothian)	Retrospective cohort study	B	87	• for lymph node TB - ethnic group (non-white > white), 26 cf. 9 weeks • site of extra-pulmonary TB	-	• Study period 1980-89 – i.e. historic • lymph node TB was found to be much more common in NC • PD / HSD differentiation not possible due to incomplete data
116	Ohmori	2005	Japan	Retrospective analysis of surveillance data	B	160,000	• PD shown to have greater impact on TDD than HSD • Labourer • Low income / on benefits	-	• study used national surveillance data • assessed trends in delay over 15yrs • assessed only smear +ve cases
109	Paynter	2004	UK (London)	Retrospective cohort study	B	71	• age >33yrs • born in low risk country • TB service capacity • organisational factors • 1st presentation to GP	• Born in high risk country • First presentation to A&E • Younger age • Sputum smear +ve TB	• Patient and HS delays contribute to delays in treatment • HSDs longer in patients who present first to GP
21	Pirkis	1996	Australia	Retrospective cohort study	B	109	• extra-pulmonary TB • 1st presentation to GP	-	• Retrospective record review at one hospital • assessed delay for pulmonary TB cases separately
123	Rao	1998	USA (St Louis)	Retrospective cohort study	B	203 over 8yrs	• inexperience of health care professionals	-	• Study only looked at in-patient delays • Surprisingly found delays in treatment initiation • 1/3 delays > 10 days
19	Rodger	2003	UK (London)	Retrospective analysis of surveillance data	C	853	• sex (F>M) • ethnic group (white>ethnic)	-	• Used surveillance data • No HIV data • TDD = symptom to diagnosis or treatment
22	Rodrigo	2001	Spain (Barcelona)	Retrospective study	B	11,485	-	-	• Over 13yr period TDD little changed • Factors for delay not assessed
121	Sanz	2007	Spain (Madrid)	Prospective cohort study	B	296	**TDD** • 1st presentation to GP • cough • extra-pulmonary TB **Factors increasing HSD** • sputum smear –ve TB • 1st presentation to GP • employment	**Factors decreasing PD** • fever • sputum AFB +ve	• Study done to assess immigrant TB • Found that PD & HSD of immigrants lower than non-immigrants

Ref No.	First Author	Year	Location	Study type	Grade of Evidence (SIGN)	Sample Size	Factors increasing total delay	Factors decreasing total delay	Comments
23	Sherman	1999	USA (New York)	Prospective study	B	184	**Factors increasing PD** • older age (55-64) • 2nd language English **Factors increasing HSD** • no cough • sputum smear -ve TB • Homelessness • no CXR at 1st visit	• cough • sputum smear +ve TB	Concluded – • patient education needed to encourage early seeking of care • High index of suspicion needed by HCWs
112	Storla	2008	Worldwide	Systematic review of observational studies	B	58 studies	• HIV • sputum smear-ve TB • Extra-pulmonary TB • Rural residence • Poor HC access • Older age • Poverty • Female gender • Alcoholism • Hx of immigration • Low educational level • Stigma		• Reviewed 58 papers on diagnostic delay and TB treatment • Analysed together high and low income countries plus high and low endemic countries • Concluded risk factors exist but heterogeneous and vicious cycle of repeat visits pre-diagnosis is key problem
25	Ward	2001	Australia (Queensland)	Retrospective study	B	782	**Factors increasing HSD** • >45yrs • female • failure to inv. cough • misdiagnosis of CXR	**PD & HSD** • born in high risk country • indigenous Australians	• complete record of all cases over 13yrs • very low prevalence area • Used surveillance data
24	Wares	1999	UK	Retrospective cohort study (letter)	B	43	-	-	• small study at one hospital • pulmonary TB only • retrospective 5yr record review • found no ethnic difference in delay • did not formally assess causes of delay

Definitions

PD = patient delay (time from 1st symptom to 1st contact with a health services)

HSD = health service delay (time from1st contact with health services to the start of treatment)

TDD = total delay (PD + HSD)

Details of the Key Studies Included in This Review

Rodger et al. (2003)[19] analysed diagnostic delay in London using national and local surveillance data over a 3 year period from 1998 to 2001. They showed for 853 patients that the median diagnostic delay (symptom onset to diagnosis or treatment) was 49 days (14-103) and that diagnostic delay was more common in white people (adjusted OR:1.67) and women (adjusted OR:1.42). This study found that nearly a third of patients had missing data, making it impossible to determine the relative contribution of patient or health service delay from the data available. Rodger's concluded that although TB is more common in ethnic minority groups, campaigns to raise awareness of TB should not just be directed at ethnic minority groups but also white patients who comprise a third of TB cases in London.

Lewis et al. (2003)[20] in a retrospective analysis of diagnostic delay reviewed 93 patients and looked in detail at the contribution of patient and healthcare delays to total diagnostic delay. Their study based in East London revealed substantial delays in the presentation of TB patients and found a median total delay of 126 days. This total delay is much longer than that found in other studies. They also found that "patient delay was significantly longer than healthcare system delay ($p=0.019$)". In addition, Lewis comments that many of the study patients were recent immigrants and that this group may have had particular difficulty accessing primary care. On multiple regression, the only factor associated with increased diagnostic delay was extra-pulmonary disease ($p=0.035$). Lewis concluded that there is a need to improve TB awareness both of patients and of health professionals in addition to improving patient access to health care.

In a small retrospective cohort study, Paynter et al. (2004)[109] looked at diagnostic delay in 2001/02 in 71 patients attending a North London hospital. Data was collected from hospital notes, notification forms and GP's via a postal questionnaire. Diagnostic delay was shorter for those born in high prevalence countries, those presenting first to hospital casualty departments rather than to a GP, younger patients and those sputum smear positive. These factors all contributed to shorter health service delay. This study supports the finding of Rodger et al.[19] that individuals from ethnic minority groups have a shorter total diagnostic delay. Paynter comments that the difference in delay between patients attending A&E and a GP could be explained by the fact that patients presenting

to A&E may be more obviously ill and may be investigated more thoroughly as a consequence; and that patients presenting to their GP may need to wait for the results of investigations and chest clinic appointments. Theoretically it should be possible to reduce these GP service delays.

Golub et al. conducted a number of studies to look at diagnostic delay.[95, 113, 114] Their primary study (2005) looked at diagnostic delay in 158 TB patients in Maryland, USA.[114] In this study, Golub showed that median patient delay (32 days) was increased in patients who were non-white, less educated and those who had no respiratory symptoms; and that median health service delay (26 days) was increased in patients aged over 50, who spoke no-English, who initially presented to a private physician, had a negative sputum smear and had not been investigated at their first visit with a skin test, chest x-ray or sputum smear. Sex, race, country of origin and drug use did not influence health care delay. Golub also noted that patients presenting with cough or with long patient delays had shorter health service delays, and those with a history of asthma had longer health service delays. Once again the conclusion of this paper is that "education of the patient population about TB symptoms" and "increased physician awareness of the current epidemiology of TB and better use of available diagnostic tools" are needed to reduce diagnostic delay and in turn reduce TB transmission.

Golub et al. (2006)[95] in another study took a different approach to diagnostic delay; and assessed TB transmission following a diagnostic delay of over 90 days in 124 close contacts of patients with pulmonary TB. They showed that a longer time to diagnosis (>90 days) led to an increase in TB transmission and that for contacts of US-born cases risk factors for transmission included ethnicity (black greater than white (adjusted OR:3.03)), sputum smear positivity (adjusted OR:3.29) and cavitation on x-ray (adjusted OR:3.11). Interestingly no association was seen between foreign born patients and the risk of transmission among their contacts. Overall, contacts exposed for longer due to a delay in diagnosis experienced higher rates of TB transmission.

Farah (2006)[118] in a study conducted in Norway, found disease site to be a significant predictor of patient, health service and total delay. He showed extra-pulmonary TB, as compared to pulmonary TB, led to a relative total delay of 1.81 (95% CI:1.10-2.98) and that patients born abroad, as compared to patients born in Norway, had a significantly shorter relative health service delay of 0.32 (95% CI:0.11-0.94). Farah also showed that

health service delays were longer than patient delays; this is at odds with most of the other studies. Farah suggests that "the main reason for the long total delay was that healthcare providers did not initiate specific TB examination despite symptoms such as cough, weight loss and night sweats", this is possibly due to the low incidence of TB in this population. Farah conducted a small sub-study looking at half the patients with a diagnostic delay greater than two months: 50% had received several courses of antibiotics and for 25% there had been difficulties isolating M.TB.

Pirkis et al. (1996)[21] assessed diagnostic delay in 109 patients in Australia from 1991-1993 and found that extra-pulmonary TB and first presentation to a GP increased total diagnostic delay (TDD). Pirkis tried to define an acceptable standard interval between onset of symptoms and treatment and suggested 0-30 days. Once again this paper concluded that raising patient and healthcare worker awareness of TB is essential to reduce diagnostic delay.

Sherman (1999)[23] in his study of 184 patients in New York, showed that patient delay (PD) was increased in those aged 55 to 64 (adjusted OR:10.6) and in patients whose first language was not English (adjusted OR:2.5). He also showed that patient delay did not vary by ethnicity, site of disease or presence of another medical condition. Statistically significant health service delay (HSD) ($p<0.001$) was shown for patients with no cough (adjusted OR:2.9), sputum smear negative TB (adjusted OR:10.2) and the homeless (adjusted OR:7.1).

Diez (2005)[111] conducted a large retrospective study of over 5,000 culture-confirmed TB cases in Spain to assess health service delays. This study covered two thirds of all Spanish TB cases presenting during 1996 and 1997. Diez found that migration, homelessness and extra-pulmonary TB lead to less health service delay. The findings, that homelessness and extra-pulmonary TB reduce delay, are opposite to the findings of Sherman.[23] Factors increasing health service delay included age over 44, IVDU, cancer, sputum smear negative diseases (a negative sputum smear increased the risk 4 times (OR:4.5, $p<0.0001$)). Diagnosis in primary care (OR 1.51) appeared to increase health service delay but this was statistically non-significant $p=0.072$. Only with extreme delay did gender have an impact: extreme HSD was associated with female gender. Diez using the same dataset assessed factors increasing patient delay and showed that age

greater than 14 years and non-respiratory symptoms increased patient delay and that cough and extra-pulmonary TB decreased patient delay.[127]

Sanz et al. (2007), conducted a prospective cohort study in Spain to specifically assess diagnostic delay in immigrants.[121] For 296 TB cases the median total diagnostic delay was 40 days, with a median patient delay of 15 days and health service delay of 5 days. Sanz found that both patient delay and health service delay in immigrants was shorter than that for the Spanish population. Possible explanations given for the shorter patient delay included: firstly, that immigrants may only receive healthcare from the hospital emergency services where investigations are immediately available (unlike in primary care) and secondly, immigrants may be more aware of the symptoms of TB and so present earlier. Possible explanations given for a shorter health service delay for immigrants included: firstly, greater suspicion of TB by healthcare professionals in this population; secondly, greater use of hospital emergency services leading to less delay in the diagnostic process; and thirdly, the younger age of the immigrant population.

Asch (1998)[110] in a prospective study of 248 patients in the US showed that disease severity had little impact on delay but access to care seemed to be a major problem. Asch concluded that "improving the availability of services for high risk groups may substantially reduce TB patients' delay in obtaining care and thus limit the spread of the disease." In addition Asch suggested that public education should be targeted at at-risk groups to raise awareness of symptoms and how to seek care.

Storla et al. (2008)[112] undertook a systematic review of 58 studies from around the world that looked at diagnostic delay and treatment of TB. Storla's analysis revealed that defining diagnostic delay is quite complex. He showed a mean total diagnostic delay of 72 days +/- 28 days (SD) with no consistent pattern of delay due to patient or healthcare worker or between high and low income countries. He found risk factors for delay to be heterogeneous: "a risk factor for increased delay in some studies was a risk factor for decreased delay in other studies." But this may have been due to the fact that he compared studies from high and low income countries with very different health care settings. Storla et al. concluded that the main problem leading to diagnostic delay was the "vicious cycle of repeat visits" to the same health care providers before a diagnosis was eventually reached.[112] This was particularly the case in high endemic, low income countries with poorly trained government health personnel and private practitioners.

Total diagnostic delay seemed to be two-fold either due to a lack of effective diagnostic tools or a lack of suspicion of TB. Storla also showed, however, that once a TB diagnosis was reached, treatment was initiated within a reasonable time.

The Meaning of Diagnostic Delay

The majority of studies reviewed used the definitions for patient delay, health service delay, and total diagnostic delay as defined at the start of Section 2.4.2.2. However, the endpoint for health service delay or total diagnostic delay was defined by Wares[24] and Sanz[121] as being from first symptom to diagnosis rather than to treatment, and by Rodger[19] as being from first symptom to diagnosis or treatment. Although not specifically stated, all of these studies in high income countries probably considered treatment to start on the day of diagnosis.

The use of the 'symptom onset date' as a starting point from which to measure diagnostic delay is particularly problematic as the onset of TB can be insidious and so people find it hard to remember a specific date. This may therefore affect the calculation of the length of diagnostic delay.

Quantification of Diagnostic Delay

Most studies defined diagnostic delay as a specific number of days. Of the 28 studies analysed 22 quantified diagnostic delay: 12 looked at patient delay, 13 at health service delay and 16 at total diagnostic delay (not mutually exclusive groupings). Table 2.6 summarises these studies. The median patient delay varied from 7 to 74 days, with most studies reporting patient delays of about one month (mean 33 days). The median health service delay varied from 6 to 39 days, with a mean of 26 days. Where studies assessed both patient delay and health service delay separately (nine studies)[20, 23, 25, 109, 114, 115, 118, 119, 121] seven studies showed that patient delay was longer than health service delay. For studies assessing total diagnostic delay (TDD), the median TDD varied from 36 to 126 days with a mean of 65 days. Some studies did not report median total diagnostic delays but mean TDDs; Gulbaran[120] reported a mean TDD of 14 days in just under a third of patients and over 28 days for 50%, Kam[124] found a mean TDD of 145 days and Storla[112] a mean TDD of 72 days.

Some studies assess delay specifically for pulmonary TB.[21,24,114] The potential public health consequences are higher for a case of pulmonary TB as the risk of transmission is

greater. Total delays (TDD) of more than one month were found in 72% of smear positive cases in Blackburn by Wares,[24] 69% of pulmonary TB patients in Australia by Pirkis[21] and 46% of pulmonary cases in Maryland USA.[114]

Table 2.6 Quantification of diagnostic delay in days

Study	Location	Sample Size	PD Median (range)	HSD Median (range)	TDD Median (range)
Altet-Gomez[115] 2003	Spain	287	43	38	81
Asch[110] 1998	USA (Los Angeles)	248	74 (20% >60 days)	-	-
Diez[127] 2004	Spain	7037	22 (1-160)	-	-
Diez[111] 2005	Spain	5,184	-	6 (0-94)	-
Farah[118] 2006	Norway	83	28 (2-217)	33 (1-378)	63 (4-434)
Franco[117] 1996	Spain	109	>1mth in 30%	>1mth in 71%	>1mth in 90%
Gagliotti[119] 2006	Italy	271	7	36	65
Golub[114] 2005	USA	158	32 (0-539)	26 (0-519)	89
Golub[113] 2005	USA	158	-	39 (2-519) on antibio. *Vs* 15 (0-191) no antibio.	-
Gulbaran[120] 1996	France (Paris)	52	-	-	**mean TDD** 14 for 27% 21 for 19% >28 for 54%
Kam[124] 2007	Canada	23	-	-	**mean TDD** 145
Lewis[20] 2003	UK (London)	93	56 (0-728)	35 (0-1470)	126 (0->1000)
Paynter[109] 2004	UK (London)	71	34	29	78 (39-159)

Study	Location	Sample Size	PD Median (range)	HSD Median (range)	TDD Median (range)
Pirkis[21] 1996	Australia	109	-	-	52
Rao[123] 1998	USA (St Louis)	203 over 8yrs	-	6 days from hospital admission	-
Rodger[19] 2003	UK (London)	853	-	-	49 (14-103)
Rodrigo[22] 2001	Spain (Barcelona)	11,485	-	-	36
Sanz[121] 2007	Spain (Madrid)	296	15 (IQR:6.5-30)	5 (IQR:0-30)	40 (IQR:16-90)
Sherman[23] 1999	USA (New York)	184	25 (0-731)	15 (0-430)	57 (4-764)
Storla[112] 2008	Worldwide	58 studies			**mean TDD** 72 days +/- 28 days (SD)
Ward[25] 2001	Australia (Queensland)	782	29	22	-
Wares[24] 1999	UK	43	-	-	49 (IQR: 28-91)

Definitions
PD = patient delay (time from 1st symptom to 1st contact with health services)
HSD = health service delay (time from1st contact with health services to the start of treatment)
TDD = total delay (PD + HSD)

Risk Factors Associated with Diagnostic Delay

This literature review, of studies done in high income, low incidence countries, shows that there are a wide range of factors that lead to diagnostic delay. The main risk factors influencing delay are shown in Table 2.7.

Clinical factors that clearly increase diagnostic delay include: sputum smear negative TB,[23,109,111,112,114,117,121,122] extra-pulmonary TB,[20,21,112,118,121,124,126] no respiratory symptoms,[111,114,117,127] a history of asthma,[114] no chest x-ray at first visit,[23,113,114] an alternative diagnosis for which an antibiotic is prescribed,[113,114] and HIV infection.[112,122] A few clinical factors appeared in some studies to increase delay and in others reduce

delay. For example a cough, Sanz[121] found this to increase delay where as Sherman, Greenaway, Diez and Golub showed it reduced delay.[23,111,114,122,127]

Other factors that were shown to reduce diagnostic delay were cavities on chest x-ray,[117,122] hospital admission[111] and history of contact with TB.[111]

Demographic factors increasing delay include: being white,[19] older age[23,25,109, 111, 112, 114, 122,127] and being female.[19,25,111,112] Immigrants from high-risk countries appeared in some studies to have increased diagnostic delay[112,119] and in others reduced delay.[25,109, 111,118,121]

Other factors noted in some studies to increase diagnostic delay included: first presentation to a GP,[21,109,111,119,121] inexperience of health care professionals,[123] poor health care access,[112] poverty,[112,116] alcoholism,[112] non-English speaking people[23,114] and low educational attainment.[112,114]

Diagnostic delay is a problem for TB patients and health services. The literature reviewed suggests that to lessen this delay:

1. Patient education is needed about symptoms to encourage earlier presentation to health services[20, 21, 23, 114]
2. Increasing awareness of TB in health professionals is needed to raise the index of suspicion for TB[20, 21, 23, 114, 119]
3. Better access to health services is needed for recent migrants[119, 121]

Table 2.7 Risk Factors for Diagnostic Delay

Risk Factors	**Evidence supporting increased diagnostic delay** Study Reference No.	**Evidence supporting decreased diagnostic delay** Study Reference No.
Clinical factors		
sputum smear -ve TB	122, 112, 117, 109, 23, 121, 114, 111, 121, 23	
extra-pulmonary TB	124, 20, 126, 21, 118, 121, 112	127, 111
no respiratory symptoms	117, 114, 127, 111	
cough	121	23, 122, 127, 111, 114
Antibiotic prescription 1st	113 ,114	

Risk Factors	**Evidence supporting increased diagnostic delay** Study Reference No.	**Evidence supporting decreased diagnostic delay** Study Reference No.
History of asthma	114	
no CXR at 1st visit	23, 114, 113	
HIV infection	122, 112	
cavities on CXR		122, 117
Demographic factors		
Immigrant	112 ,119	109, 25, 121, 111, 118
Ethnic group (white>ethnic)	19	114
Older age	122, 112, 109, 111, 114, 25, 23, 127	
Gender (Female > Male)	19, 112, 25, 111	
Other factors		
1st seen by GP	109, 21, 121, 119, 111	1stseen in A&E 109, 114
1st seen by private physician	114	
Inexperience of health care professionals	123	
Poor HC access	112	
Poverty	112, 116	
unemployed	110	
less educated	114, 112	
uncertain where to get care	110	
IVDU	111	
cancer & TB	111	
non-English speaking	23, 114	
Alcoholism	112	
Homelessness	23	111

2.4.2.3 Discussion of Diagnostic Delay

Understanding diagnostic delay in TB is important as diagnostic delay can lead to an increased risk of TB transmission,[95] worsening morbidity and potentially increased mortality. To reduce TB incidence both theory and data suggest that shortening diagnostic and treatment delay would be beneficial.

The prevalence and epidemiology of TB in high prevalence / low income countries is very different to the UK. I therefore analysed only the research literature on diagnostic delay from countries similar to the UK, i.e. countries with high income but low TB prevalence. Many of the studies reviewed were "retrospective studies" of the delay experienced by TB patients, so subject to recall bias. Patients who died without a diagnosis of TB or who were diagnosed at post-mortem may not be included in the analyses and so results could therefore be biased.

The Meaning of Diagnostic Delay

While assessing the meaning of diagnostic delay I found that most of the studies I reviewed used the same definitions for patient delay, health service delay, and total diagnostic delay. However, Storla et al.[112] noted in their research that defining diagnostic delay was complex and that different studies used different definitions. This difference is partly explained by the fact that Storla looked at studies from around the world in very different populations: both low and high income countries and low and high prevalence areas.

Patient delay occurs when a patient is unaware of their likely diagnosis, when they fail to seek help early or are unable to access appropriate healthcare. Health service delays occur when doctors fail to make the diagnosis, perhaps through inexperience, or failure to consider the diagnosis in patients perceived as being at low risk. The symptoms of TB are often non-specific and this makes early diagnosis difficult as symptoms such as cough can be explained by many different diseases. Difficulty in recognizing symptoms as due to TB is likely to account for much of the diagnostic delay seen in TB.

Quantification of Diagnostic Delay

From the studies reviewed a mean total diagnostic delay of 65 days was calculated. As the studies selected, were from healthcare settings similar to the UK it is interesting that

the mean total diagnostic delay is similar to that found by Storla et al. (72 days) in their systematic review of high and low income countries, and high and low prevalence countries.[112]

Patient and health care system delays contribute to the overall delay experienced by patients from onset of TB symptoms to diagnosis and treatment. In this review the majority of studies (7 out of 9) showed that patient delay was longer than health service delay. It should be noted however, that patient delay can be difficult to quantify precisely. The date of symptom onset and first medical presentation for TB may not be accurately known as the symptoms and signs of TB are often nonspecific and insidious.

Diagnostic delay leads to increased disease transmission. The public health response to TB in the UK tends to focus on the case once diagnosed. This response includes treating the case and identifying and treating close contacts. However TB transmission is likely to have occurred before the patient presents to health care services; therefore reducing patient delay should reduce the chance of transmission and therefore the burden of TB in the UK. Delays of more than two months have been shown to increase the risk of transmission to domestic contacts.[95]

Risk Factors Associated with Diagnostic Delay

Understanding the risk factors that lead to a delay in TB diagnosis is crucial to ensure earlier diagnosis and treatment and so minimise the risk of transmission and reduce poor outcomes for patients.

This literature review suggests that there are many factors that lengthen diagnostic delay. Clear-cut determinants leading to increased diagnostic delay include: sputum smear negative TB, extra-pulmonary TB, no respiratory symptoms, a history of asthma, HIV infection, no chest x-ray at the first visit, being white, older and female and an alternative diagnosis for which an antibiotic is prescribed. From the research presented here, patients with extra-pulmonary TB experience longer delays prior to diagnosis. These patients are less likely to be infectious and are therefore considered less of a public health problem. However, diagnostic delay can lead to worsening disease and increased mortality.[94, 126] Diagnostic delay increases with age and this may be due to diagnostic confusion in older people, where inter-current illness or co-existing disease may initially be diagnosed rather than TB.[23, 25, 109, 111, 112, 114, 122, 127]

For the majority of factors associated with diagnostic delay the evidence in high income, low prevalence countries shows the risk factors to be the same. Only in a few areas is the research contradictory. The main area where some studies show less delay and others show more is TB in immigrants. In some studies immigrants from high risk countries have less diagnostic delay;[25, 109, 111, 118, 121] Other studies, however, show the opposite: that migrants experience more diagnostic delay.[112, 119] Possible explanations for a shorter delay include: migrants may not be registered with a GP and so use hospital emergency services more where investigations are immediately available; migrants may be more aware of the symptoms of TB and so they present earlier; and in one UK study less delay is attributed to a greater suspicion of TB in migrants among healthcare workers.[109] For studies that show migrants experience longer diagnostic delay, it is possible that access to health care and language barriers may have an influence. Generic research assessing immigrants access to health care has shown that recent immigrants are more likely to have difficulty accessing healthcare.[128] Diagnostic delay for non-immigrants may be due to the fact that TB is less likely to be suspected.[119] In the UK, health care professionals need to be informed that ethnic minorities generally have a higher risk of TB regardless of how recently they arrived in the UK but also that a high index of suspicion for TB should be held for any person whose symptoms and signs suggest TB.

Another area where the research evidence is contradictory is for cough. One study found cough to increase delay[121] while others showed it reduced delay.[23, 111, 114, 122, 127] Cough may increase a delay in diagnosing TB where it is attributed to other diseases such as chronic airways disease, asthma or pneumonia and treatment is given for these first and only later TB is suspected. In general however cough reduces delay, particularly when linked to other typical TB symptoms.

A number of studies[113,118,129,130] have documented diagnostic delay due to the prescription of antibiotics for a community-acquired pneumonia, which was subsequently diagnosed to be TB. Golub showed a doubling of diagnostic delay for patient's that received antibiotics prior to a TB diagnosis.[113] He also showed this delay was regardless of the antibiotic class prescribed: fluoroquinolones are no different to any other antibiotic class. This is contrary to what was previously thought.[129, 131, 132] In his conclusion Golub suggests more widespread use of chest x-rays before prescribing

antibiotics presumptively.[113] This is not, however, a cost effective way of managing patients. Clearer diagnostic algorithms may help to reduce these problems.

Although much work suggests that ethnic minorities and immigrants are at higher risk of TB, it is important not to forget that there is still a sizeable proportion of TB in white patients. In London up to a third of cases are white[19] and many of these come from hard to reach groups such as the homeless, drug users and those with a history of imprisonment.[133] Education of health care workers, not only to improve their knowledge of TB but also to raise awareness of the risk factors for TB and those groups most affected, is key to reducing diagnostic delay. Many of the studies reviewed here come to a similar conclusion: the need for improved TB education for healthcare workers.

A number of studies comment that diagnostic delay is increased if a patient presents first to a GP[21, 109, 111, 119, 121] and shortened if they present initially to A&E.[109,114] This may be due to the availability of chest x-rays in A&E departments. However, increasing knowledge of TB in primary care professionals is also crucial if diagnostic delay is to be reduced.

In low incidence countries factors leading to diagnostic delay are mostly similar and non-controversial. However Storla,[112] in his review of high and low income countries and high and low prevalence countries found that risk factors for diagnostic delay exist but that they are heterogeneous (some studies showing one effect and others showing the opposite) and that a "vicious cycle of repeat visits" pre-diagnosis is a key problem.

In summary, in low incidence countries a delay in TB diagnosis is still common and the risk factors for this are well documented.

2.5 Conclusions

This literature review assesses the literature on TB from a primary-care perspective and summarises current research on TB and diagnostic delay. By reviewing the literature on diagnostic delay, quantifying it and understanding the factors that increase it, conclusions can be drawn on how diagnostic delay might be reduced in a primary care setting.

This literature review shows that the role of primary care in the diagnosis of TB is little studied and therefore warrants greater investigation. One small study by Metcalf et al. is the most relevant. This paper highlights the fact that early diagnosis is one of the most important contributions primary care can make to improved TB control and that raising awareness through improved education among health professionals is essential. In addition this paper identifies barriers to prompt diagnosis: atypical presentations, low level of clinical suspicion of TB by GPs, lack of continuity of care, workload demands that limit time with patients, and sub-optimal clinical / patient communication. All of these barriers are born out in my literature review that specifically looks at diagnostic delay and the factors that lead to this. My literature review reveals diagnostic delay for TB patients in low incidence countries is still common: a median total diagnostic delay of 2 months is not unusual. Diagnostic delay is a problem for TB patients and the health service. To lessen this delay, patient and health care worker education is urgently needed.

In conclusion, this literature review shows that TB patients often experience diagnostic delay and that the role of primary care in this and other areas of TB care is little studied. Further research is therefore needed to understand the role of primary care in the diagnosis of TB and investigate ways to reduce diagnostic delay in this setting. Early diagnosis & treatment of TB are essential for effective TB control.

Chapter 3

Methods

3.1 Introduction

This epidemiological research was carried out using the General Practice Research Database (GPRD) the world's largest computerised primary care database. At the time of this research it contained 15 years of anonymised longitudinal data (1987-2002), for 8 million patients from 743 GP practices and covered 6% of the UK population with a median of 3.7 years of data per patient. My research to investigate the epidemiology of TB from a primary care perspective consisted of four studies using GPRD data:

Study 1: Consultation Behaviour in Primary Care Prior to a Diagnosis of TB: A Case-Control Study

The aims and objectives of this study were:

- To assess the extent of missed opportunities for a diagnosis of TB in primary care by measuring excess GP consultations for TB cases as compared to controls in the 6 months prior to diagnosis
- To describe the nature of these consultations by examining consultation behaviour and consultation type
- To investigate risk factors for TB

Study 2: Consultation Behaviour in Primary Care Prior to a Diagnosis of TB: A Self-Controlled Analysis

The aims and objectives of this study were:

- To assess the extent of missed opportunities for a diagnosis of TB in primary care by measuring excess GP consultations in the 6 months prior to diagnosis in TB cases as compared to 12 to 18 months prior to diagnosis in the same individuals (i.e. a self-controlled study which eliminates the potential effect of relatively static unmeasured confounders such as ethnicity and social status)
- To describe the nature of these consultations by examining consultations and consultation behaviour
- To compare the results of this self-controlled study with the case-control study (Study 1)

Study 3: Morbidity Associated with TB: A Self-Controlled Study

The objectives of this study were:

- To investigate the long-term outcomes (morbidity) of patients with TB
- To describe and quantify morbidity from TB
- To assess / examine if patients are more likely to suffer chronic respiratory problems following an episode of pulmonary tuberculosis

Study 4: Mortality Associated with TB: A Cohort Study

The objectives of this study were:

- To describe and quantify mortality from TB
- To calculate the mortality rate in patients with TB
- To examine if TB indirectly increases the risk of death from non-TB causes
- To describe risk factors for death in TB patients

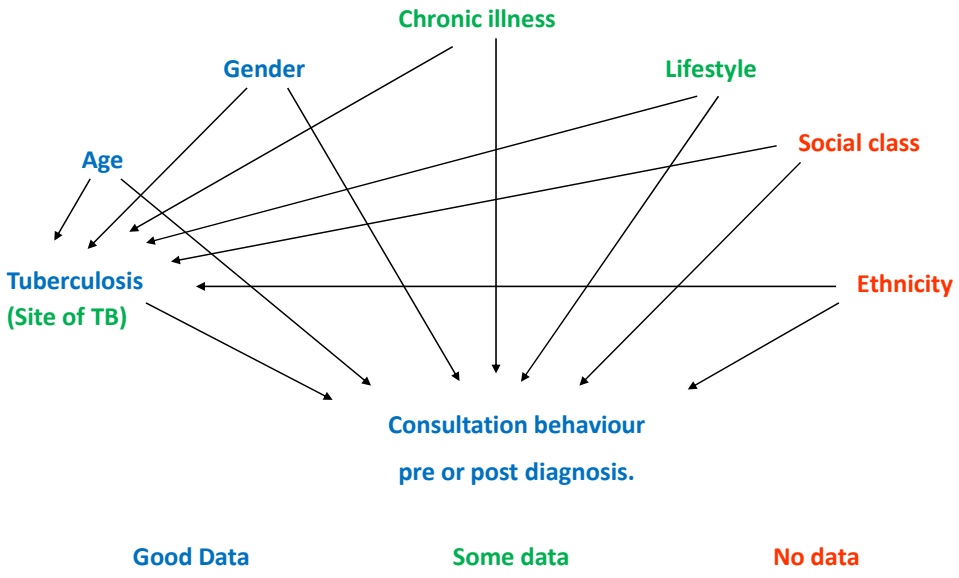
In this chapter I will discuss the generic methods leaving the detailed methods to be described in the relevant study Chapters 4-7.

3.2 Study Design

To investigate TB epidemiology from a primary care perspective I needed to design a number of studies that would enable GPRD data to give me a window on TB pre- and post- diagnosis. By using GPRD data I was able to look at data not visible to conventional, often hospital based TB studies, i.e. data on general consultations pre-diagnosis and also long-term outcomes. To this end I started with a conventional case-control study to look at data pre-TB. In addition I completed another study pre-TB using self-controlled analysis to reduce confounding due to unmeasured confounders such as ethnicity and socioeconomic status. To understand TB epidemiology and outcomes post-diagnosis I continued using the self-controlled methodology to look at longer-term morbidity and undertook a cohort study to look at mortality.

Consultation behaviour and TB disease are both influenced by age, sex, chronic disease, lifestyle, social class and ethnicity. Data is variably recorded / coded in the GPRD for these factors and so consideration needed to be given to these factors in the study design (see Figure 3.1).

Figure 3.1 Study design considerations



3.3 Setting

I examined data from UK general practices that provided data to the General Practice Research Database (GPRD) during the time period April 1990 to April 2002. I used only data from practices that met a pre-set quality standard for data recording.

3.4 Data Source - The GPRD

The General Practice Research Database (GPRD) is the world's largest computerised primary care database with over 15 years of longitudinal data at the time of this study. (see Chapter 1, Section 8 for more details on the GPRD)[75,76] The GPRD contains computerised clinical information on individual patients. It is entered in a standardised manner by General Practitioners (GPs). At the time of this study 743 GPs, covering a population exceeding 8 million, systematically provided data files anonymously to the Office of National Statistics (ONS) and allowed their information to be used for research purposes.[134] ONS and subsequently the MHRA organises this information and performs a series of regular quality checks.[75,76] The information contained within the GPRD includes demographic details of each patient, details of all patient visits, medication prescribed, a summary of specialists' clinical notes and hospital letters, results of laboratory tests and a free-text section. The free-text section is where details of the consultation are recorded, such as symptoms and relevant patient history. In addition, the GPRD records patient deaths either entered by the GP or by ONS from death certification.

The GPRD contains four types of record for each patient:

The Patient Record

A record of basic demographic data: date of birth, sex, family number (a number shared by patients living at the same address), registration data and death date.

The Medical Record

A record of every GP consultation a patient has had; this includes information on diagnoses, symptoms and signs with associated dates in addition to where the patient was seen or the information came from and an outcome.

The Therapy Record

A record of GP prescriptions including: drug names, drug codes, doses, formulations and dates of prescribing.

The Prevention Record

A record containing additional information such as: occupation, vaccinations, smoking status, alcohol consumption and patient characteristics e.g. height and weight.

In addition to the four data tables described above there are a number of ***lookup tables***. These contain lists of GPRD codes and definitions, e.g. drug codes and their description. These lookup tables ensure that the data tables are easier to manipulate as they contain numeric codes rather than long text strings.

A unique number links each patient record and is known as the "PatientId".

3.4.1 Clinical Data Entry

To undertake research using the GPRD an understanding of how GPs document an 'event' or diagnosis in the medical record is key. When a patient consults their GP a 'diagnosis' is entered into the patient's medical record. In the GPRD dataset this diagnosis appears as text and as an abbreviated code. Two types of codes are used, OXMIS and Read Codes.

OXMIS (Oxford Medical Information System) codes are problem codes based on two widely used coding systems: the 10th version of the International Classification of Diseases (ICD-10) and the Office of Population Censuses and Surveys Classification of Surgical Operations and Procedures (4th Edition) (OPCS).[135]

Read Codes are a hierarchical coded thesaurus of clinical terms whose purpose is to create a clinical record that can be easily manipulated by computer systems. The codes were developed in 1982 for use in General Practice by Dr James Read;[136] in 1990 the NHS bought them, expanded them and now maintains and develops them. They are used throughout the NHS as standardised clinical codes and there usage is now mandatory in General Practice. Read Codes can be mapped to the 10th version of the International Classification of Diseases (ICD-10) and the 4th revision of the Classification of Surgical Operations and Procedures (OPCS-4). Read codes are not

based on the ICD or the OPCS codes unlike OXMIS and so contain many more clinically descriptive / non-mapped terms. OXMIS codes can be mapped to Read Codes.

The codes used by GPs depend on the version of practice software used by the GP submitting data to the GPRD. Older versions of practice software up to VAMP 5 are likely to have used OXMIS codes, however ViSion software (and VAMP Millennium complaint software) use Read codes. There are four times as many Read codes as OXMIS codes, allowing GPs greater flexibility when recording patient consultations.

3.4.2 Definition of a Consultation

To analyse GP consultations a clear definition of a 'consultation' was needed. For the purpose of this research a 'consultation' or 'visit' refers to the date on which an 'event' or event(s) took place. An 'event' refers to any diagnosis or symptom recorded by the GP during a consultation: in any one GP consultation there may be a number of 'events' recorded. Statistical analysis of the events and consultations pre-diagnosis were similar, so as a 'consultation' seems a more understandable concept I have used this rather than 'events' to support my arguments and analysis.

3.4.3 Up-To-Standard Data

To ensure that data entered into the GPRD database is of sufficient quality for the purposes of research the GPRD team at the MHRA check and validate the data against a set of 'standards' that all GP practices submitting data to the GPRD try to adhere to. These standards ensure that the data contained within the GPRD is as complete as possible to minimise missing values. When data from a GP practice meets these standards it is considered to be 'up-to-standard' (UTS), the practice is granted UTS status and a UTS date is generated for the practice. Any patient registering with a practice after this date will automatically have UTS data. If however a patient has registered before the practice is given UTS status, the date on which the practice receives UTS status is the date on which the patients UTS period begins. MHRA recommends that research using the GPRD should only be conducted on patients' UTS data. In addition, to a practice meeting an agreed minimum standard to submit data for inclusion into the GPRD, ongoing quality checks occur on the data regularly extracted. If a practice fails to meet the quality standards they are removed from the GPRD.[75]

3.5 Study Participants

Patients were selected from the GPRD between 1990 and 2002; they were grouped into cases and controls by outcome for the case-control study and TB cases and non-TB cases for the cohort study. In all studies cases were all patients with a diagnosis of active-TB.

3.5.1 Criteria for the Selection of TB Cases

To extract TB cases from the GPRD dataset I used the following criteria:

Inclusion Criteria for Cases:

- Patients with a medical code for active-TB
- Patients with at least 18 months of up-to-standard (UTS) patient data prior to a diagnosis of TB

Exclusion Criteria for Cases:

- Patients with less than 18 months of UTS data prior to a diagnosis of TB

To select cases with active-TB I created a search strategy to generate a TB-code list, i.e. medical codes, compatible with a diagnosis of active-TB (see Section 3.5.2) and then combined this with up-to-standard GPRD data (see Section 3.4.3).

3.5.2 Search Strategy to Generate Medical Codes Compatible with TB

This systematic search strategy was used to generate a list of GPRD diagnoses (medcodes) that were compatible with an overall diagnosis of TB.

Step 1 - Search NHS Clinical Terminology Browser for Read Codes Compatible with a Diagnosis of "Tuberculosis"

The NHS Clinical Terminology Browser (version 1.0), available from the NHS Information Authority, is a tool that helps to identify Read codes compatible with a descriptive medical diagnosis. I chose "tuberculosis" from within the infectious disease section of the NHS Clinical Terminology Browser and used this to search for all relevant Read codes and their descriptors which are associated with TB. [http://www.nhsia.nhs.uk/terms/pages/default.asp]. n = 164 codes

Step 2 - Link NHS Browser Read Codes to a "5 Digit GPRD Medcode"

To ensure that I had a complete code list for TB diagnoses I exported the NHS clinical terminology codes for TB into ACCESS and then mapped these to the GPRD Read codes using a "5 digit medcode" common to both sets.

Step 3 – Conduct a 'Word' Search for All GPRD Codes Related to Tuberculosis

Using the GPRD "Look-up Med Code" table, that lists over 100,000 Read and OXMIS codes, I undertook a 'word' search for all Read and OXMIS codes relating to tuberculosis using the term "*tuberc*". I truncated the word tuberculosis at both ends using *…..* to ensure all codes related to TB were found. n = 314 codes

Step 4 - Map "5 Digit GPRD Medcode" to Equivalent PCPS Codes

Using the complete list of TB codes generated in Step 3 I then mapped these to a PCPS code. A PCPS code is a unique code generated by UCL's Department of Primary Care and Population Sciences (PCPS) which uniquely identifies each diagnostic code within the GPRD whether Read or OXMIS.

Step 5 - Remove Medcode Duplicates Using the Statistical Package STATA

n = 392 codes

Step 6 - Search for Extra Codes Specific for TB in Prevention Records

This step was used as a cross-check to ensure that no codes for TB had been missed by using "*tuberc*" in the GPRD Medcode look-up table or by using "tuberculosis" in the NHS Clinical Terminology browser. The GPRD prevention record files were scanned for TB related investigations and vaccinations such as BCG. These records were then linked back to the medical diagnoses of individual patients - no additional information was found in the prevention records.

Step 7 - Search for Codes Specific for TB in Therapy Records

This step was used as a cross-check to ensure that no codes for TB had been missed using steps 1-6 above. Using the British National Formulary section 5.1.9: "TB drugs" as a guide [http://bnf.org/bnf/] I searched the GPRD therapy records using the main drugs used to treat TB: pyrazinamide, ethambutol and isoniazid with rifampicin. No additional information was found in the therapy records when linked back to the medical diagnoses of individual patients.

Step 8 - Remove Inappropriate Codes

Finally the complete list of diagnostic codes for TB was reviewed by hand and inappropriate codes: those not diagnostic of active TB were removed. Examples of codes removed include: 'TB contact', 'contact screening' and 'family history of TB'. A total of 88 codes were removed

304 medical codes compatible with active-TB were found and subsequently used to select patients from the GPRD full dataset to generate the cases on which this research is based (see Appendix 2 for a full list of TB codes used in this research).

3.5.3 Details of TB Cases Selected

Of 5,161 cases with active-TB in the GPRD database, 3,032 had at least 18 months of up-to-standard data pre-diagnosis.

3.5.4 Criteria for the Selection of Controls

Five controls per case were randomly extracted from the GPRD dataset using the following criteria:

Inclusion Criteria for Controls:

- Age matched to age-band of linked TB case
- Registered at the same GP practice as linked TB case (linked using practice id number)
- At least 18 months UTS data preceding a control index date (a random date within the controls UTS period)

Exclusion Criteria for Controls:

- A medical code for active-TB
- Taking TB drugs (see Appendix 3 for a TB drug code list)

The random sample was computer generated by an independent data manager not involved in the study.

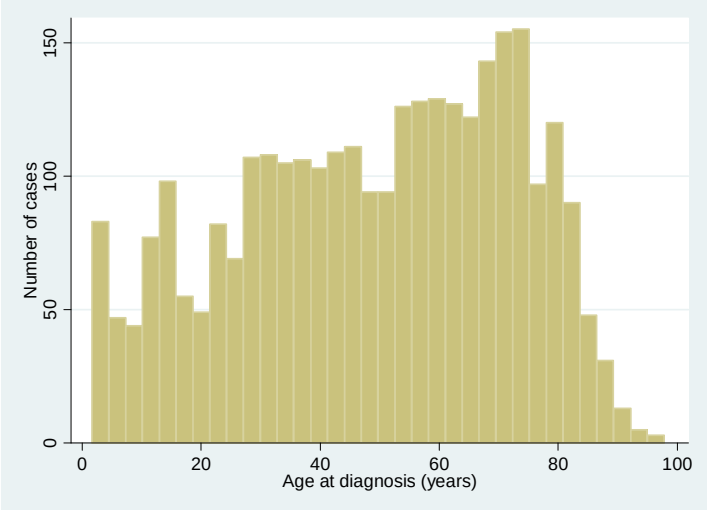
Age matching was used because the age distribution of cases differed from the age distribution of randomly selected 'test' controls. Age bands of 20 years were used to sub-divide the cases, and then five times the number of cases were selected for each age band (see following for details).

Cases and controls were matched on practice to minimise the confounding effect of social class and ethnicity (for which data is not available), with the assumption that patients registered with the same GP are more likely to be in similar socio-economic and ethnic groups; although there is of course considerable heterogeneity within a practice.

3.5.5 Age Matching of Cases and Controls

To decide how to select the controls for this study I reviewed the age distribution of TB cases and that of a group of 'test' controls. Figure 3.2 shows the age distribution of TB cases; the mean age was 49 years and cases were not 'normally' distributed but slightly skewed to the left (a case's age was taken to be that at diagnosis of TB).

Figure 3.2 Distribution of age at diagnosis for 3032 TB cases



A randomly selected 'test' group of 5 controls per case was extracted from the GPRD and the age distribution reviewed. The 'test' controls were not normally distributed but skewed to the right therefore having a different age distribution to the cases. The mean age of 'test' controls was 12 years younger than cases at 36 years.

As the age distribution of cases and 'test' controls was different, matching by age was appropriate for the control group. The criteria used to select 'controls' from the GPRD dataset included the restriction 'match by age'. To 'match by age,' the 3032 TB cases were divided into five age bands, each of 20 years, and five controls per case were selected from within each age band (table 3.5.3). A random date was selected on which each control had to be in the same age-band as its linked case.

Table 3.1 Cases and controls by age band used

Age Band	Cases	Controls
<=20	427	2135
>20 - <=40	651	3255
>40 - <=60	803	4015
>60 - <=80	929	4645
>80	222	1110
Total	**3032**	**15160**

3.6 Methodological Approaches and Statistical Analysis

The detail of individual study methods and statistical analysis can be found in the relevant chapter (Chapters 4 - 7). Here I discuss in generic terms the approaches I used and why.

3.6.1 Study 1 - A Case-Control Study

I used a case-control study to examine consultation behaviour prior to a diagnosis of TB and to explore whether differences exist between cases and controls in the six months prior to diagnosis. A case-control study was conducted as it offered an efficient study design and allowed me to take multiple controls. A six month period of analysis was chosen because previous research suggests that most TB patients present within 6 months of first symptoms, within one study an average of 3 months was found.[20] Five controls per case were randomly extracted from the GPRD to maximise statistical power. Research shows that taking up to 4 controls per case increases the power for each additional control taken; above 4 the benefits rapidly decrease.[137-139] I selected 5

controls per case so that I would have at least 4 controls even if some controls were lost due to insufficient longitudinal data when analysed.

In this study I took cases with TB and controls without-TB and looked back in time to see how frequently cases consulted with the symptoms of TB as compared to controls. For the first part of my study I was therefore not using the design of a case-control study in the traditional way to identify risk factors, but rather to gain an understanding of the natural history of TB prior to diagnosis and also to understand if opportunities exist to diagnose TB earlier. However, the case-control design also afforded the opportunity to address a secondary objective of identifying risk factors for TB.

To assess consultation behaviour pre-TB, I analysed consultation rates for cases and controls, the frequency of consultations and the odds of TB for the classical symptoms of TB and diagnosed respiratory conditions. Odds ratios (OR) were calculated using conditional logistic regression and compared the odds of symptoms, or odds of respiratory diseases, in cases in the six months prior to a diagnosis of TB with those in the six months prior to an index-date in controls. Conditional logistic regression is appropriate for the analysis of individually matched data and for case-control studies. The strength of the association between classical TB symptoms and tuberculosis was assessed using odds ratios comparing cases to controls and similarly the association between TB and respiratory diagnoses (for which TB might be mistaken). Respiratory diagnoses were divided into 3 specific diagnostic groups: ‘all respiratory diseases’ which included all acute and chronic respiratory conditions other than TB, ‘respiratory infections other than TB’ and ‘lower respiratory infections other than TB’ (LRTI).

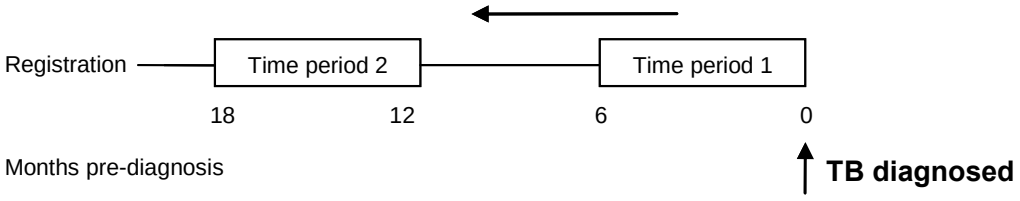
To investigate potential risk factors for TB I used single and multivariable analysis to calculate odds ratios for TB. Multivariable conditional logistic regression models were constructed and all variables that showed statistical significance in the prediction of TB in single variable analysis ($p<0.05$) were retained in the final multivariable model. Analyses were controlled for gender, smoking and pre-existing morbidity (e.g. coronary heart disease, chronic renal failure, chronic respiratory disease, diabetes). Results of the conditional logistic regression analyses were reported as adjusted odds ratios with 95% confidence intervals. The multivariable analysis allowed me to control for several variables / confounders at the same time and to understand the dependency of certain variables on several other explanatory variables.

The results of a case-control study will be incorrect if confounding variables are not accounted for in the analysis. At the time of analysis the GPRD dataset did not contain any ethnicity data, a major confounder for TB (over two thirds of UK TB cases occur in people from ethnic minorities)[47] So although the dataset was large, confounding may suggest a difference between cases and controls that does not truly exist. To try to minimise confounding I selected controls from the same GP practice and in a similar age band to TB cases: a matched design. I chose controls from the same GP practice, with the assumption that patients registered with the same GP were more likely to be from a similar social background and ethnic group; although realising that considerable confounding is still likely to exist.

3.6.2 Study 2 - A Self-Controlled Study

I used a self-controlled (case-only) study as a second method to examine consultation behaviour prior to a diagnosis of TB. I used this method so that self matching of cases would eliminate control selection bias, remove confounding by permanently fixed factors e.g. gender, ethnicity, genetics and minimise confounding by factors that vary little over the comparison period e.g. chronic illness, age, social class, lifestyle.[140-142]

Due to confounding, a case-control study may not show the true association between an exposure and a disease or in this study between consultations pre-diagnosis and disease. Potentially important confounding variables, such as ethnicity and social status are not recorded in the GPRD or are poorly recorded e.g. smoking; and so it is important to control for these variables. One method to evaluate the findings of a case-control study and simultaneously reduce confounding is to use a 'self-controlled' study design. A self-controlled study, alternatively described as self-matching of cases or case-only analysis,[140-143] allows a case to act as its own control during a disease-free period either prior to or after the event under study. A self-controlled study can be defined as repeated observations on an individual in two different time periods (as shown in the diagram below). In this study, time periods 1 and 2 are observation periods 0-6 and 12-18 months prior to a diagnosis of TB:



Similar studies using the GPRD and this methodology have been conducted and have been to shown to be robust.[83]

In a self-controlled study the analysis is based only on data from cases, who act as their own controls, "*The incidence of events during one predefined time period is compared with the incidence of events in the same individuals in a different predefined base-line period*".[142]

By comparing 12-18 months pre-TB with 0-6 months pre-TB I reduced the impact of seasonality on disease. Research shows that TB exhibits seasonal variation,[144-146] in addition to diagnoses such as LRTI which might be confused for TB. It also ensured that I was comparing a TB-free period (18-12 months pre-diagnosis) with a period in which TB was likely to be developing (6-0 months pre-diagnosis). Since the self-controlled analysis included only cases with at least 18 months of data prior to a TB diagnosis and only one control per case (that is itself) it was less powered than the case-control study.

This self-controlled study used the same TB cases as the case-control study presented in Chapter 4, i.e. TB cases with at least 18 months of data pre-diagnosis, and compared consultation behaviour in two time periods 12-18 months pre-TB and 0-6 months pre-TB.

The Generalised Estimating Equation (GEE) model was used for statistical analysis.[22,24,25,147] The GEE model compares total counts for each individual and then calculates a summary measure for all individuals, the incidence rate ratio (IRR) with 95% confidence intervals. In methodological terms the GEE model calculates an incidence rate ratio (IRR); in practical terms, however, in this study it calculates the ratio of consultations in two different time periods. I will therefore refer to it as calculating a rate ratio as it is not calculating a true incidence as an individual may consult twice with, for example, a persistent cough which would not necessarily be a new incidence of cough.

The GEE model was considered the most appropriate analytic tool for a number of reasons: firstly as repeated measurements were taken on the same individual,[22,25,148] secondly because the outcome measures were total number of GP consultations during specific time intervals, thirdly because the distribution of cases was not a 'normal' distribution, fourthly it allows for correlated observations without explicitly defining a model and fifthly focuses on estimating an average response over a population.

Consideration was given to other statistical methods / models that might be used to analyse a self-controlled study. Conditional logistic regression would not be suitable as it needs independent individuals for its analysis. In this study comparisons were made in different time periods for the same individual, so the individuals were not independent. The Wilcoxon Signed Rank Test, only gives a *p* value and a distribution type for the data, it is not a decision-making tool and so gives no relative risk or comparator of cases to controls and a Paired T-test would not be appropriate as it compares means and we have a skewed distribution. A review was made of the 'case-series' method for statistical analysis but this is primarily for a time varying exposure and its association with an outcome event.[140, 142, 149-151] A 'case-series' method was not appropriate for the analysis in this pre-TB study.

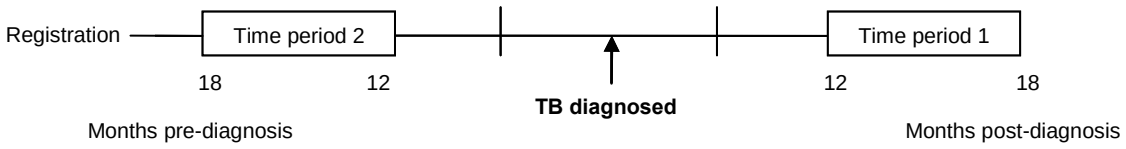
The downside of the self-controlled analysis is the fact that only cases with at least 18 months of data prior to TB and only one control per case can be used so reducing the power of the study. However, using cases as their own self-matched controls guarantees no difference in fixed confounders between cases and their self controls, so long as the matching is preserved in the analysis, and minimises differences in confounders that vary little over the time period in question.

Study 1 and 2 therefore look at the same relationship between consultation behaviour prior to diagnosis and disease using different methodologies (case-control and self-controlled) in an attempt to understand and minimise the impact of possible confounders such as ethnicity and social-status.

3.6.3 Study 3 - A Self-Controlled Study to Assess Morbidity Post TB

To test my null hypothesis that once TB patients have received treatment they return to full health I used a self-controlled study. Following a diagnosis of TB, statistical analysis of patient outcomes is possible with longitudinal data and the use of a cohort

study. But it is complex due to the loss of both cases and controls to follow-up in addition to the ever present problem of controlling for confounders. A self-controlled study using the self-controlled 'case series' methodology provides a good solution to minimise the effects of confounding variables.[142,149,150,151,152,153] The use of a self-controlled study design for the post-TB period is similar to that described for the pre-TB self-controlled study earlier; however, it uses data from the post-TB, 12-18 month period, as the control period instead of 12-18 months pre-TB.



In this analysis the consultation behaviour of cases in the 12-18 months prior to diagnosis was compared with their own consultation behaviour in the 12-18 month period following diagnosis. Total consultation rates, respiratory consultation rates and antibiotic and inhaler prescribing rates were compared in these two time periods. A similar approach is not possible for mortality. Standard TB treatment lasts for six months, occasionally patients will take treatment for over 6 months. By using data 12 months after a TB diagnosis I can test my null hypothesis that post-treatment TB patients have no increase in morbidity: they consult no differently than when they themselves are well 12-18 months pre-diagnosis.

To analyse this study I used the Generalised Estimating Equation (GEE) model to generate an incidence rate ratio (IRR) with 95% confidence interval.[22,24,25,147] This method is the same as described for the pre-TB self-controlled study (study 2) in Chapter 5.

3.6.4 Study 4 - A Cohort Study to Assess Mortality Post-TB

I conducted a conventional cohort study to examine the survival of patients diagnosed with TB. To determine whether TB patients have a higher death rate not directly attributable to TB I compared TB cases and cases without TB using survival analysis. This was done using death from any cause (excluding TB coded deaths) and death from a respiratory cause (excluding TB coded deaths). Analyses were controlled for age,

gender, smoking and pre-existing disease such as coronary heart disease, chronic renal failure, chronic respiratory disease and diabetes.

To identify risk factors for death in patients with TB a further survival analysis was performed and was restricted to cases of TB. Factors that were explored as predictors of death/confounders included sex, site of TB, delay in diagnosis, major chronic diseases and smoking status.

Statistical analysis was first conducted using a single variable analysis. Categorical variables were compared using the Chi Squared test. Multivariable Cox proportional hazards regression analysis was used to evaluate risk factors for death. All variables with $p<=0.05$ were retained in the final multivariable model. All p values <0.05 were considered statistically significant.

A Cox proportional hazards model was used to compare the survival of patients with TB to those without.[154] This model was used to help account for patients who did not provide complete data throughout the follow-up period (censored observations) and was adjusted for sex and chronic disease. Hazard ratios and their 95% confidence intervals were calculated and a Cox Model survival curve plotted.

All data in the four studies discussed above were analysed using STATA version 9.0 (Stata Corp, College Station, TX, USA)[155] and ACCESS software. Further details of the statistical methodology for each study are presented in the relevant chapter.

3.7 Data Preparation

3.7.1 Development of 'Code' Lists to Analyse Patient Consultations

I assessed the NHS Clinical terminology browser codes and GPRD code groups developed by researchers from the Office of National Statistics, the Communicable Disease Surveillance Centre, now the HPA Centre for Infections, and the UCL Centre for Infectious Disease Epidemiology. I present here the final search strategies I generated using STATA to give diagnostic code lists to assess patient consultations in my research.

To compare patient consultations pre and post TB I selected the following criteria:

- Symptoms and signs compatible with TB
- Diagnoses similar to TB which might be diagnosed instead of TB
- TB drugs
- Antibiotic prescriptions
- Inhaler prescriptions
- Chronic diseases and smoking status
- HIV and cancer

To analyse the data in the GPRD I needed to generated code groupings compatible with these criteria and this is what I present here.

3.7.2 Search Strategy to Identify Codes for the Symptoms and Signs of TB

Step 1 – Determine the Classic Symptoms and Signs of TB

Using the paper by Lewis[20] as reference I selected the following symptoms and signs as classic for TB:

- Cough
- Fever
- Sweats
- Lymphadenopathy
- Haemoptysis
- Weight loss

Step 2 – Search GPRD for All Infection Related Codes

Using the GPRD "Look-up Med Code" table, which lists all Read and OXMIS codes and covers all the potential diagnoses a GP might use, I selected all codes of an 'infectious nature' i.e. where infection was the likely cause. = 4292 codes

Step 3 – Search GPRD Subset for All Respiratory Infection Codes

From within the infectious code list I selected all the infectious respiratory codes.

= 242 codes

Step 4 – Search GPRD Respiratory Infection Subset for All Codes Related to the Classic Symptoms and Signs of TB

From within the infectious respiratory codes I selected all codes relating to the classic symptoms and signs of TB:

- Cough - 59
- Fever, - 30
- Sweats - 12
- Lymphadenopathy - 62
- Haemoptysis - 6
- Weight loss - 10 = 179 codes

Step 5 - Link Symptom and Sign OXMIS and Read Medcodes to Equivalent PCPS Codes

Using the list of codes generated in step 4 I then linked each code to a PCPS code, the final code to be used to search the GPRD medical records database (see Section 3.5.1 for definitions).

3.7.3 Search Strategy to Identify Codes for Medical Conditions that Might Be Mistakenly Diagnosed Instead of TB

To generate codes compatible with conditions that might be mistakenly diagnosed instead of TB I selected 'all respiratory infection' codes (excluding TB) as for step 3 in 3.7.2 above and combined these with all GPRD codes for asthma and COPD to create a code group 'all common respiratory diseases' (378 codes). In addition I separated out codes specifically for lower respiratory tract infections (LRTI) (100 codes).

3.7.4 Search Strategy to Identify Therapy Code Groups

To search the therapy records held in the GPRD, appropriate drugs were selected from the British National Formulary (BNF) and converted into BNF drug codes. These BNF codes were then linked via the GPRD drug dictionary with GPRD recognised drug codes and used to search the GPRD therapy records. The BNF is a publication produced by the British Medical Association and Royal Pharmaceutical Society that details key information on the prescribing, dispensing and administration of drugs in the UK.

TB Drugs

Step 1 - Search for Codes Specific for TB Drugs in the BNF

Using the BNF, I selected chapter 5 – Respiratory Diseases and from within this chapter section 5.1.9 – drugs to treat TB. A BNF section number can then be converted into a BNF code e.g. 05.01.09.00 which can then be used to search the drugs dictionary of the GPRD to find all relevant drugs and their GPRD drug code, the multilex code.

Step 2 - Search GPRD Drug Dictionary for TB Drugs Using the BNF Codes

Using the BNF code for TB drugs 05.01.09.00, I searched the GPRD drugs dictionary for the main drugs used to treat TB: pyrazinamide, ethambutol, isoniazid and rifampicin. = 80 codes

Respiratory Antibiotics

Similar steps as those taken for TB drugs were used to search for respiratory antibiotics. Using chapter 5 of the BNF – Infections, I selected section 5.1 Antibacterial drugs and from this the subsections on – penicillins, cephalosporins, tetracyclines, macrolides, cotrimoxazole and quinolones (5.1.1, 5.1.2, 5.1.3, 5.1.5, 5.1.8 and 5.1.12). I then used the BNF codes I had generated to search the GPRD drugs dictionary for all antibacterial drugs used to treat chest infections. = 1721 codes

Respiratory Inhalers

Similar steps were taken to search for respiratory inhalers. Using chapter 3 of the BNF – Respiratory Systems, I selected sections 3.1 Bronchodilators, 3.2 Corticosteroids and 3.3 chromoglycates, and used the BNF codes 03.01.00.00, 03.02.00.00 and 03.03.00.00 to search the GPRD drugs dictionary for relevant drugs and then restricted the final code list to only those available as an inhaler. = 298 codes

3.7.5 Search Strategy to Identify Chronic Disease Codes

Code lists for coronary heart disease, chronic renal failure, diabetes and chronic respiratory disease (which included asthma, COPD, bronchiectasis and rarer diseases such as bird fanciers lung) were generated using a similar methodology to that detailed in Section 3.7.2.

3.7.6 Search Strategy to Identify HIV Codes

Using STATA I generated a code list for HIV by searching the GPRD medcode list using the key words 'HIV', 'human immunodeficiency virus' and 'AIDS'. This generated 57 codes for HIV. = 57 codes

3.7.7 Search Strategy to Identify Cancer Codes

I generated a code list for all cancers using the following method:

1. I searched the GPRD medcode list using key words: cancer, carcinoma, neoplas*, sarcoma, leukaemia, lymphoma, hodgkins, malignan*, melanoma, mesothelioma, myeloma and tumour;
2. I then did a 'stem' search query looking for all GPRD medcodes starting with 'B*' as this represents all READ codes related to cancer;
3. I then trawled the combined list of codes generated by steps 1 and 2 to exclude and delete irrelevant codes such as 'benign', 'FH', 'family history', 'screen*', 'suspected' and 'non malignant';
4. Finally I removed all duplicate codes

This method generated a list of 5,862 cancer codes. = 5,862 codes

3.7.8 Strategy to Determine Cause of Death

To determine a patient's cause of death I took all deaths and grouped them into one of six causes; i.e. death due to TB, cardiovascular disease, respiratory disease (excluding TB), cancer, other diseases (including diabetes and renal failure) and unknown. I then cleaned the death codes per patient (varying between 1 and 6 per patient) to one code, the most likely cause of death, and used the following assumptions:

1. If death occurred on the same day that TB was diagnosed I assumed TB was the cause of death
2. If death occurred within 1 month of a TB diagnosis I assumed the death was likely to be due to TB unless another obvious cause such as CVA was recorded on the day of death as well as TB

3. All other deaths I assumed were not TB and allocated them to the most likely generic group listed above.

3.8 Data Cleaning

Medical records for 3032 TB cases were extracted from the full GPRD dataset. Over 196,000 individual contacts with primary care were recorded for these patients. To clean the data of consultations in which the patient had not formally seen their GP I reviewed all consultations and separated the "true consultations" from the "non-GP" consultations. To do this I analysed how often a type of consultation occurred and where it occurred with the view to removing all "non-GP" consultations such as consultations where the patient did not see their GP, that is, "did not attend" or "had a test taken."

3.8.1 Identification of True Consultations with a GP

To identify 'true' GP consultations I analysed all the patient medical records by frequency of contact with primary care. Over twelve thousand different contact descriptions and therefore analytic codes were found. The first 40 of these are tabulated in Table 3.2.

Table 3.2 Top 40 reasons for TB patients to be in contact with primary care

GPRD consultation diagnosis	Frequency
Chest infection	2884
INFECTION CHEST	2440
Medication requested	2435
Tuberculosis	2336
TUBERCULOSIS	2291
Cough	2264
Did not attend - no reason	2050*
URTI (UPPER RESPIRATORY TRACT INFECTION)	1945
Blood sample -> Lab NOS	1916*
COUGH	1899

GPRD consultation diagnosis	Frequency
Upper respiratory infection NOS	1798
Reviewed at hospital	1727*
BLOOD TEST	1587*
Had a chat to patient	1517
Asthma	1405
Injection given	1363*
MEDICAL CERTIFICATE	1360
REVIEWED AT HOSPITAL	1235*
Chest pain	1228
Influenza vaccination	1190*
Incoming mail NOS	1183*
MED3 - doctor's statement	1179
ASTHMA	1178
Respiratory tract infection	1103
Screening - general	1068
DNA (DID NOT ARRIVE)	1041*
Depressive disorder NEC	1027
Patient's condition improved	962
Repeated prescription	926
Urinary tract infection, site not specified NOS	922
Pain	914
PAIN BACK	903
UTI (URINARY TRACT INFECTION)	887
X-RAY CHEST	886*
PAIN	853
FULL BLOOD COUNT	844*
Discharged from hospital	835*
Chronic obstructive airways disease	831
Backache	821
Pulmonary tuberculosis	809

Listed by decreasing frequency of occurrence

** = non-consultation*

Some of these 12,000+ descriptors of contact with primary care are not in fact consultations with a GP but represent codes for other things; for example; 2050 records were for patients who did not attend, 1916 were for blood samples and 1727 records referred to a hospital attendance. To meet my study objectives it was crucial to look only at 'true' contacts with a GP where a diagnosis of TB might be possible. The data therefore needed to be cleaned of "non-GP" consultation and the following steps were used:

1. All 12,000+ descriptors of contact with primary care were ordered by frequency, with the commonest first.
2. The consultation types were then reviewed manually and a preliminary selection of non-GP consultations was made.
3. All consults were then re-ordered alphabetically to show up similar non-GP consultations which might have been missed visually the first time.
4. A final selection was made and discussed with an epidemiologist and a GP.
5. A consensus was reached and over 50,000 non-GP consultations were removed (see Appendix 4 for the top 40 "non-GP" consultations removed).

This led to over 18% of "non-GP" consultations being removed. The top four of these removed were: did not attend; blood test; reviewed at hospital; and injection given.

3.8.2 Analysis of Where the Contact with a Patient Occurred

Within each medical record there is a descriptor called 'locate,' short for location. 89% of patient contacts within primary care had a location recorded. Of 20 possible locations, five were clearly not taking place in primary care and did not involve a consultation with the patients GP: letter from outpatients, discharge details, A&E attendance, mail from patient, administration (Did not attend, consultant letters…)

I considered removing these five locations from my analysis. However, they only represented 2.5% of all patient contacts and most were removed by the restrictions applied during identification of "true" GP consultations cleaning process described above, and so I decided I did not need to remove these five location consultations specifically.

3.9 Approval for this Research

A protocol summarising this research was prepared and submitted to the University College London (UCL) GPRD steering group. After acceptance by UCL the protocol was submitted to the independent Scientific and Ethical Advisory Group (SEAG) of the GPRD. SEAG governs the research use of the GPRD and considers the scientific and ethical soundness of research that will use the GPRD database. SEAG gave its approval for this research to be carried out and issued protocol number 608. I did not have access to named patient data for patients with TB and did not perform any analyses that could compromise patient confidentiality.

Chapter 4

Consultation Behaviour in Primary Care Prior to a Diagnosis of TB: A Case-Control Study

4.1 Introduction

TB is on the increase but little is published on its natural history or epidemiology prior to diagnosis or from a primary care perspective. To understand these aspects of TB epidemiology, I undertook a large study to examine consultation behaviour and risk factors prior to a diagnosis of TB using primary care data from the General Practice Research Database (GPRD).

4.2 Aims & Objectives

4.2.1 Aims

The aims of this study were to explore how TB patients present in primary care, to assess the extent of missed opportunities for a diagnosis of TB in primary care and assess ways to diagnose TB earlier.

4.2.2 Objectives

- To measure excess GP consultations for TB cases as compared to controls in the 6 months prior to diagnosis
- To examine consultation behaviour and consultation type
- To investigate risk factors for TB

4.3 Methods

Study Design

A case-control study was conducted to compare consultation rates and behaviour in the 6 months prior to a diagnosis of TB for cases with the 6 months prior to a randomly selected date in age and practice matched controls (patients without TB). The same case/control participants were also used to assess risk factors for TB.

Setting

I examined data from 743 general practices in the UK who provided data to the General Practice Research Database (GPRD) during the time period April 1990 to April 2002. I used only data from practices that met a pre-set quality standard.

Data Source

The General Practice Research Database (GPRD) is the world's largest computerised primary care database with over 20 years of longitudinal data (1987-2009). Currently it contains data on approx 13 million patients from approximately 500 practices covering 5.5% of the UK population.[77] It is considered broadly representative of the UK population,[76] has been widely used for epidemiological research[78-83] and data quality has been shown to be high in independently validated studies.[76, 80, 84-86] Anonymised patient data is collected from participating practices and all information for individual patients is handled using a unique identification number.[75] GPs record medical diagnoses and symptoms using Read and Oxford Medical Information System (OXMIS) codes; these codes are based on, and can be mapped to, the internationally recognised classification systems of Disease (ICD-10) and Surgical Operations and Procedures (OPCS-4). Data for each patient includes age, sex, medical diagnosis and symptom records, details of patient consultations, mediation prescribed, a summary of specialists' clinical notes and hospital letters, results of laboratory tests and a free-text section.

Participants

Patients were selected from the GPRD between 1990 and 2002; they were grouped into cases and controls by outcome.

Cases: were GPRD patients with a medical code for active TB (see Appendix 2)
Controls: were a random sample of five people without TB for every person with TB; matched on age-band and general practice but not year of TB diagnosis. The random sample was computer generated by an independent data manager not involved in the study. For analytic purposes I chose a random index date for each control and considered this equivalent to the date of TB diagnosis for cases.
'Exclusion' criteria: I excluded patients who had less than 18 months of pre-diagnosis data (see Chapter 3, Section 3.5).

Measurements

To identify all patients with a diagnosis of TB I developed Read and OXMIS code lists (see Appendix 2 and Chapter 3, Section 3.7 for more details). I compared consultation behaviour for TB cases in the six months before diagnosis with that of controls in the six months before a randomly selected index-date. To assess consultation behaviour, I analysed consultation rates for cases and controls and the frequency of consultations and the odds of TB for the classical symptoms of TB and diagnosed respiratory conditions. To investigate potential risk factors for TB I used single and multivariable analysis to calculate odds ratios for the association between consultations and TB.

Bias

I tried to reduce ethnicity and social status bias by matching patients on GP practice.

Statistical Analysis

All the available cases in the GPRD that met my inclusion criteria were used to maximise the power of the study. Five matched controls were used per case; more controls than this are unnecessary as research suggests they only marginally increase power.[137-139]

Statistical analysis was performed using STATA version 9.0 (Stata Corp, College Station, TX, USA).[155] Consultation rates for cases and controls were calculated and

compared. GP consultations for specific symptoms and respiratory diagnoses were analysed using single variable models. Odds ratios (OR) and their 95% confidence intervals (CI) were calculated using conditional logistic regression with TB as the outcome. A *p* value of 0.05 or less was considered statistically significant. Single variable and multivariable analysis was undertaken to study associations between several different variables: sex, smoking status, pre-existing morbidity (e.g. coronary heart disease, chronic renal failure, chronic respiratory disease and diabetes) and the odds of TB. Results of the multivariable model are reported as adjusted odds ratios with 95% CIs. For further details of the statistical method see Chapter 3, Section 3.6.

4.4 Results

A total of 3,032 TB cases and 15,160 age-matched controls (5 per case) were identified from the GPRD. 196,518 consultations were analysed for cases and 610,539 consultations for controls.

4.4.1 Characteristics of Study Participants

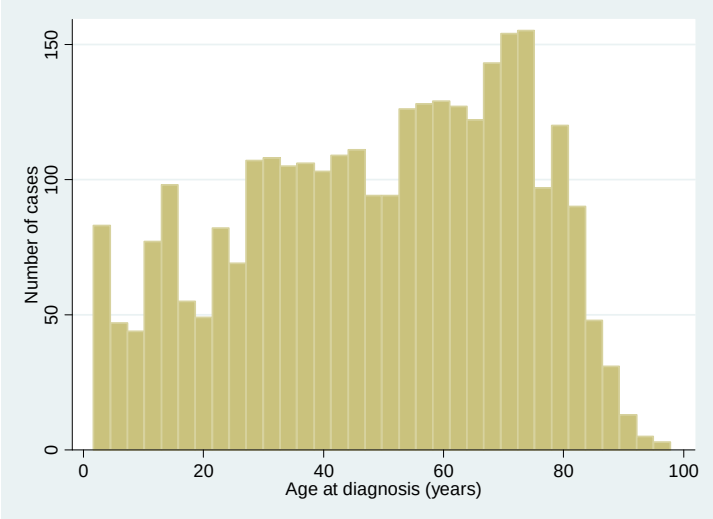
Table 4.1 details the baseline characteristics of cases and controls. Similar numbers of males and females were diagnosed during the study period. The age distribution of cases and controls was equivalent as controls were matched to cases on pre-selected age bands. The mean age of cases was 49 years and the age distribution (which is not 'normally' distributed) is shown in Figure 4.1. (S.D = 23.09, range = 9 months to 98 years).

Table 4.1 Characteristics of study participants

Characteristic	CASES (n=3,032) No. (%)	CONTROLS (n=15,160) No. (%)
Sex		
Female	1490 (49%)	7833 (52%)
Male	1542 (51%)	7327 (48%)
Age		
<=20	427 (14%)	2135 (14%)
>20 - <=40	651 (22%)	3255 (22%)
>40 - <=60	803 (26%)	4015 (26%)
>60 - <=80	929 (31%)	4645 (31%)
>80	222 (7%)	1110 (7%)
Chronic Disease		
Any	1061 (35%)	2729 (18%)
Respiratory	861 (28%)	1837 (12%)
Renal	31 (1%)	42 (0.3%)
Cardiac	213 (7%)	708 (5%)
Diabetes	162 (5%)	475 (3%)
Ever smoked		
Yes	439 (15%)	1732 (11%)

At diagnosis 35% (n=1061) of cases and only 18% (n=2729) of controls had one or more of the four main groups of chronic disease. The largest difference was seen for chronic respiratory disease where 28% of cases and 12% of controls had a chronic respiratory disease. Only six people, 0.20% of cases, were recorded by their GP to be HIV positive as compared to two people, 0.01%, of controls. Fifty four percent of cases appeared to have been prescribed TB treatment in primary care.

Figure 4.1 Distribution of age at diagnosis for 3032 TB cases

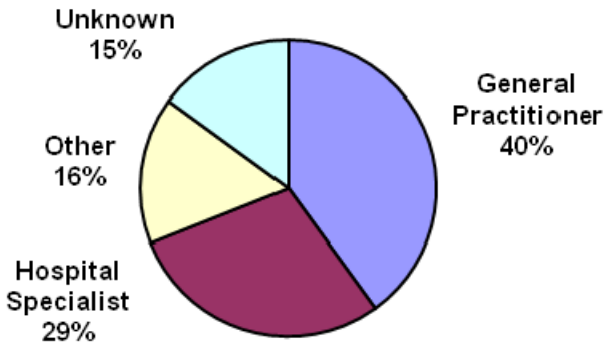


TB can be divided in to a number of sub-types depending on its pathological site. Within this dataset 25% (n=759) were coded as pulmonary TB, 16% (n=483) as extra-pulmonary TB, 58% (n=1774) as 'unclassified TB' and 1% 'TB at an unknown site'.

4.4.2 Place of TB Diagnosis

TB was diagnosed by GPs in 40% of cases (n=1,208), by a hospital specialist in 29% (n=865), by another but un-named source in 16% (n=492) and the place of diagnosis was unknown for 15% (n=407).

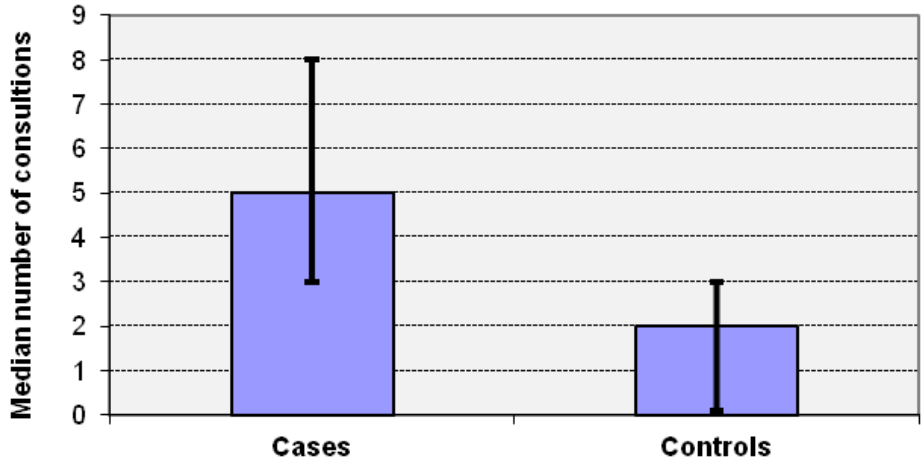
Figure 4.2 Place of TB diagnosis



4.4.3 Consultation Data – General

In the 6 months prior to a diagnosis of TB, 3032 TB patients had 15,461 consultations recorded by their GP (as defined in Chapter 3, Section 3.4.2). In the same time period 15,160 controls had 28,789 consultations. The median number of consultations, per TB case was 5 with an inter-quartile range of 3 to 8; and for controls was 2 with an inter-quartile range of 0 to 3 (see Figure 4.3).

Figure 4.3 Median number of consultations for cases and controls



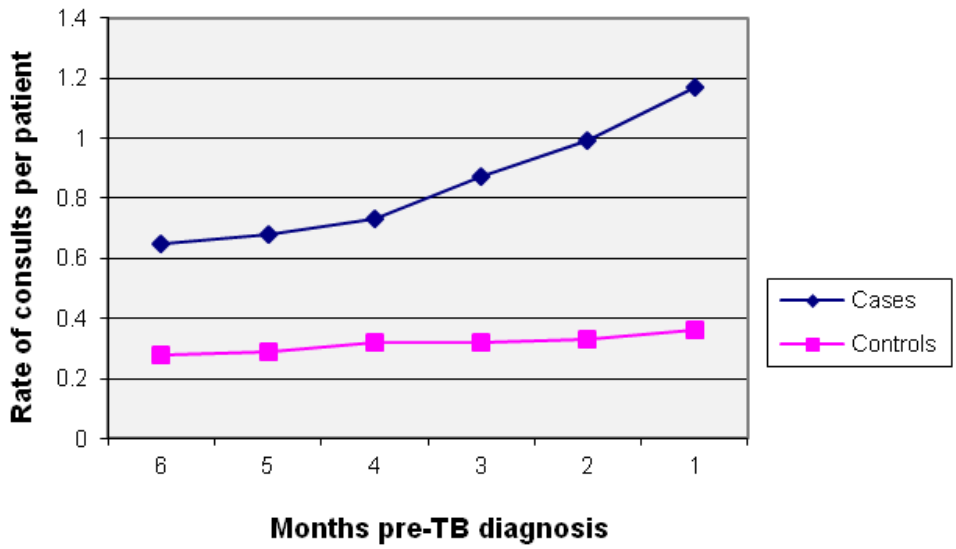
Overall, cases consulted 2.68 times more often than controls in the 6 months prior to a diagnosis of TB. Table 4.2 shows a comparison of consultation rates in the six months pre-TB for cases and prior to a fixed index-date for controls. The number of GP consultations for cases was 2.26 times higher than for controls at the start of the study period (RR 2.26), consultations then rose steadily for TB cases from 4 months pre-diagnosis (RR 2.70 at 3 months) with the rise most marked 1 month prior to diagnosis when TB patients consult 3.30 times more than controls (see Table 4.2 & Figure 4.4).

Table 4.2 Comparison of GP consultations for TB cases and controls

	CASES 0-6 months pre-TB (n=3,032)		CONTROLS 0-6 months pre-index date (n=15,160)		
CONSULTATIONS	Number	Consults per case	Number	Consults per control	**Consultation Rate Ratio** (95% CI)
Total consultations 0-6 months pre-TB	15,461*	5.10	28,789	1.90	**2.68** (2.63-2.74)
Consultations per month					
0-1 months	3560*	1.17	5383	0.36	**3.31** (3.17-3.45)
1-2 months pre-TB	3013	0.99	4993	0.33	**3.02** (2.88-3.16)
2-3 months pre-TB	2636	0.87	4880	0.32	**2.70** (2.57-2.83)
3-4 months pre-TB	2225	0.73	4795	0.32	**2.32** (2.20-2.44)
4-5 months pre-TB	2071	0.68	4425	0.29	**2.34** (2.22-2.47)
5-6 months pre-TB	1956	0.65	4313	0.28	**2.26** (2.15-2.39)

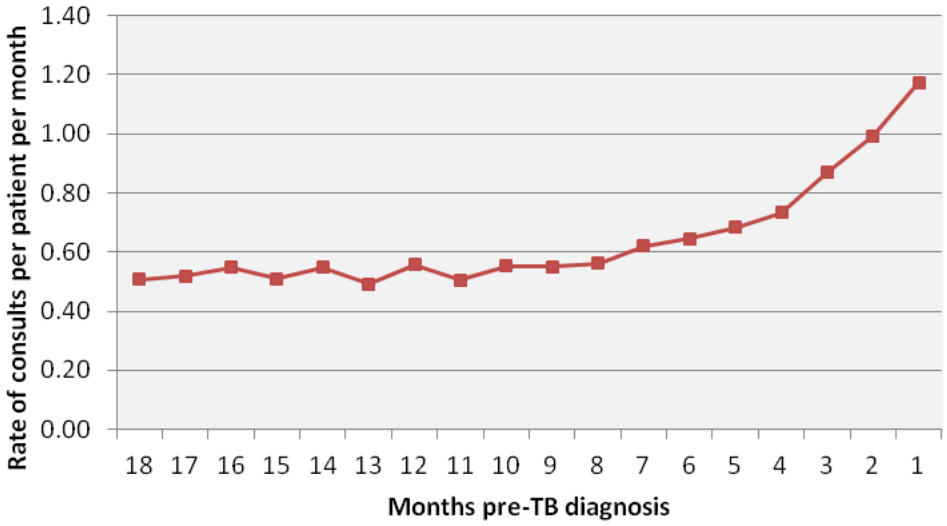
***** *includes only consultations occurring prior to a TB diagnosis*

Figure 4.4 Rate of GP consultations in the 6 months pre-diagnosis



Consultation rates per case, per month pre-TB were reviewed from 18 months pre-diagnosis (see figure 4.5). This showed that from about 7 months pre-TB consultation rates started to rise for patients who were subsequently diagnosed with TB. The background consultation rate for cases was therefore always higher than that of controls.

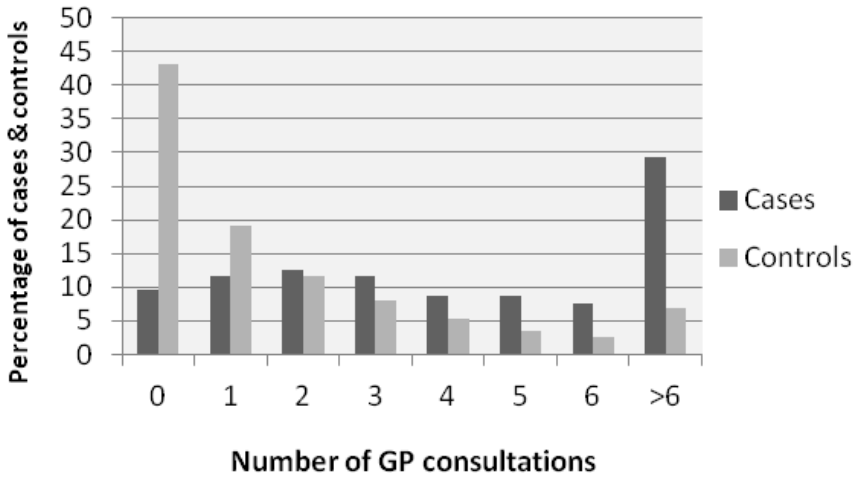
Figure 4.5 Rate of consultations per case per month pre-TB



4.4.4 Analysis of GP Consultations per Patient

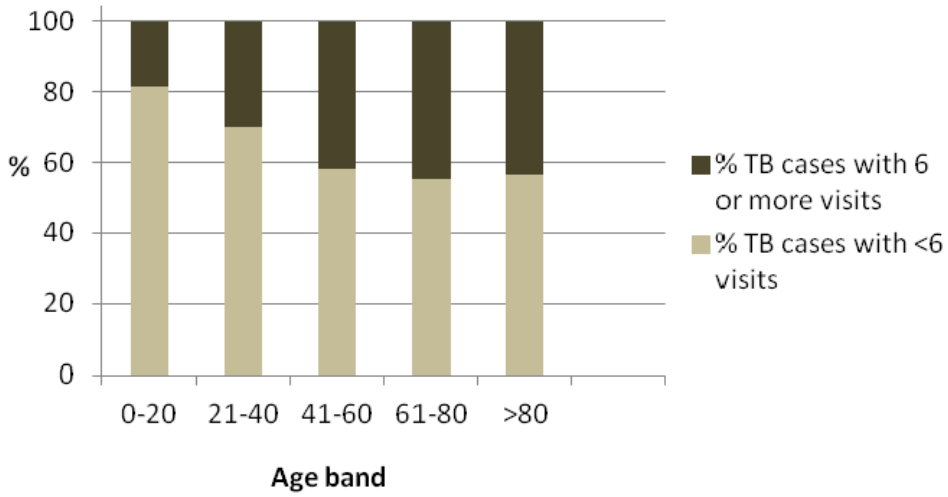
Cases consulted more often than controls in the six months prior to a diagnosis of TB. The median number of consultations was 5 with a range of 0-41. In the six months prior to a TB diagnosis, a third of TB patients (n=1121) had 6 or more consultations compared to 9% (n=1419) of controls. Half (54%) of cases had more than 3 consultations and 12% saw their GP only once. 10% of cases had no GP consultations compared to 43% of controls (see Figure 4.6).

Figure 4.6 Distribution of visits for cases and controls in the 6 months pre-TB



In order to explore variations in the consultation rates of TB patients, a subgroup analysis of the characteristics of high consulters (6 or more consults) compared to lower consulters was performed. TB patients with 6 or more consultations were more likely to be older (see Figure 4.7). Of the people who consulted 6 or more times, 26% had one or more chronic diseases of the four main groups assessed, compared to 35% for all TB cases.

Figure 4.7 Percentage of TB cases by age band comparing those with six or more consults to those with less than six



4.4.5 Analysis of Consultations for the Classic Symptoms of TB

GP consultations for the classic symptoms of TB (see Chapter 3: Section 3.6.2, for further details) were compared for cases and controls in the 6 months pre-TB or index-date. Nearly a third of TB patients (n=832) had symptoms suggestive of TB recorded by their GP prior to diagnosis as compared to 4.5% (n=682) of controls. Consultations for each of the classic symptoms of TB showed that in the 6 months prior to a TB diagnosis, cough was the most common symptom recorded by GPs (16%) (see Table 4.3). Four percent of controls presented with cough and only very rarely with any of the other classic symptoms of TB.

Table 4.3 Consultations for TB symptoms in cases and controls in the 6 months prior to a diagnosis of TB including odds ratios for TB

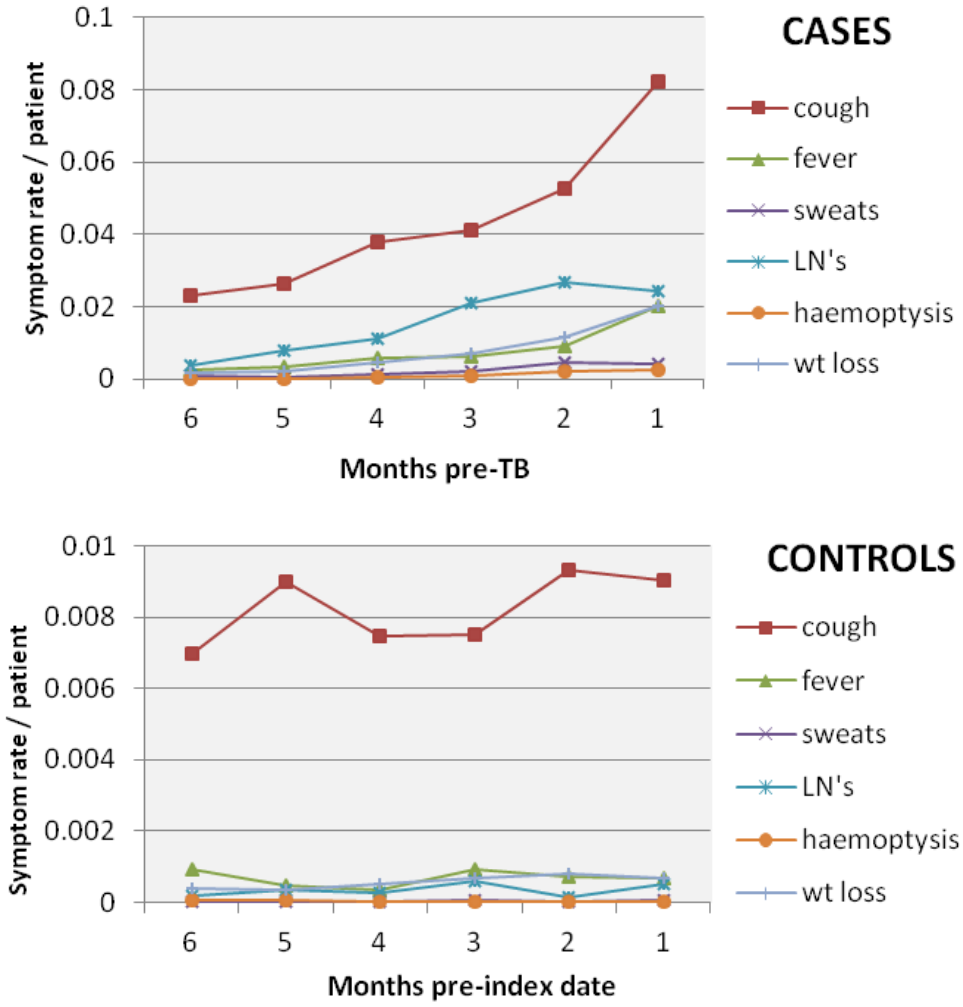
Symptoms	CASES (n=3,032) No. (%)	CONTROLS (n=15,160) No. (%)	ODDS RATIO for TB (95% CI)	*p* value
Cough	497 (16.4%)	570 (3.8%)	**5.20** (4.56 – 5.93)	<0.001
Fever	103 (3.4%)	57 (0.4%)	**9.66** (6.92 – 13.50)	<0.001
Sweating	28 (0.9%)	2 (0.01%)	**70.00** (16.68 – 293.82)	<0.001
Lymphadenopathy	167 (5.5%)	30 (0.2%)	**32.90** (21.61 – 50.11)	<0.001
Haemoptysis	13 (0.4%)	2 (0.01%)	**32.50** (7.33 – 144.02)	<0.001
Weight loss	114 (3.8%)	43 (0.3%)	**14.05** (9.80 – 20.15)	<0.001
Any TB symptom	832 pats (27.4%)	682 pats (4.5%)	**8.37** (7.43 – 9.41)	<0.001

Of note: Table 4.3 shows for each TB symptom, how many cases or controls exhibit this symptom at least once during the 6 month study period. Any one individual may have a number of different symptoms recorded and so may be counted in more than one symptom group.

The strength of association between the classic symptoms of TB and a later diagnosis of TB are shown by the odds ratios in Table 4.3. In the 6 months prior to diagnosis the odds of TB was 8.37 times greater for those with any of the classic symptoms of TB than those without and the odds of TB was 5.2 times greater for those with a cough than those without. The odds ratios for TB and its cardinal symptoms show that fever (OR:9.7), weight loss (OR:14.05) and lymph nodes (OR:32.9) are strongly and individually more predictive than cough (OR:5.2) for patients who subsequently go on to develop TB. Odds ratios for sweating and haemoptysis are large but are associated with very wide confidence intervals due to small numbers. Symptom counts for controls are small.

A review of GP consultations month by month in the six months pre-TB revealed that TB cases have a gradual increase in consultations for 'cough' and lymphadenopathy from 4 months pre-diagnosis; with these consultations increasing most rapidly 2 months prior to diagnosis. Consultations for fever and weight loss showed a significant increase only in the last month and sweats and haemoptysis were uncommon through-out the six months period (see Figure 4.8). For controls, GP consultations for the classic symptoms of TB remained little changed throughout the 6 month period prior to a fixed index-date.

Figure 4.8 Classic TB symptom rates by month for cases and controls



Further analysis of different combinations of classic TB symptoms was unhelpful. Very few GPs recorded more than one TB symptom at any one consultation: only 76 of 15,461 consultations (0.5%) have two TB symptoms recorded at the same consultation.

4.4.6 Analysis of Respiratory Consultations

Nearly half, 42% (n=1268) of TB cases were diagnosed with a respiratory problem (respiratory infection or chronic respiratory disease) in the 6 months pre-TB as compared to 10% (n=1570) of controls. One in three TB patients (n=1177) presented with a respiratory infection and of these over half were lower respiratory tract infections (see Table 4.4).

On single variable analysis the odds ratio for TB for patients consulting with a 'respiratory disease' was nearly seven times that of those without TB (OR:6.81) and the odds of TB for those with a 'respiratory infection or cough' was greater than seven times that of those without TB (OR: 7.31) (see Table 4.4).

Table 4.4 Consultations for respiratory disease in cases and controls in the 6 months prior to a diagnosis of TB including odds ratios for TB

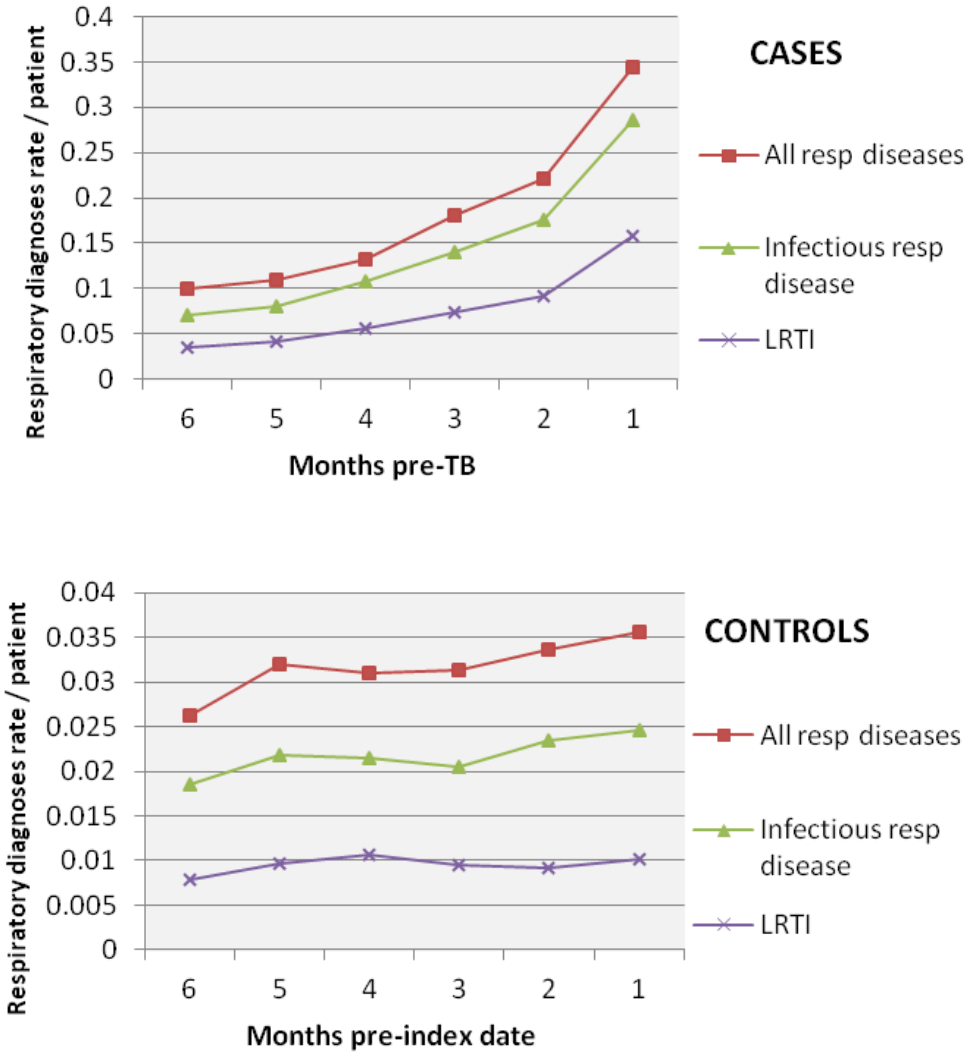
Respiratory Disease*	CASES (n=3,032) No. (%)	CONTROLS (n=15,160) No. (%)	ODDS RATIO for TB (95% CI)	*p* value
LRTI	751 (25%)	564 (4%)	**7.81** (6.92 – 8.82)	<0.001
All respiratory infections (incl. cough)	1177 (39%)	1312 (9%)	**7.31** (6.61 – 8.09)	<0.001
All respiratory diseases (incl. resp. infection)	1268 (42%)	1570 (10%)	**6.81** (6.19-7.51)	<0.001

Of note - each 'respiratory disease' group listed is a sub-group of the next
**Excluding TB*

Analysis of the frequency with which cases presented with respiratory symptoms showed that 29% (n=347) of TB patients with respiratory infections had 3 or more respiratory infections diagnosed during the six months pre-TB. Twenty percent of all TB patients had two or more respiratory infections diagnosed in the six months pre-TB and 11%, three or more.

For cases, a gradual increase in respiratory consultations is seen in the 6 months prior to a diagnosis of TB. However for controls the consultation rate is more or less constant. The rate of increase in respiratory consultations is significantly higher for cases one month before diagnosis. This is mostly explained by the diagnosis of respiratory infections of which about half are specifically lower respiratory tract infections (see Figure 4.9).

Figure 4.9 Respiratory diagnoses by month for cases and controls

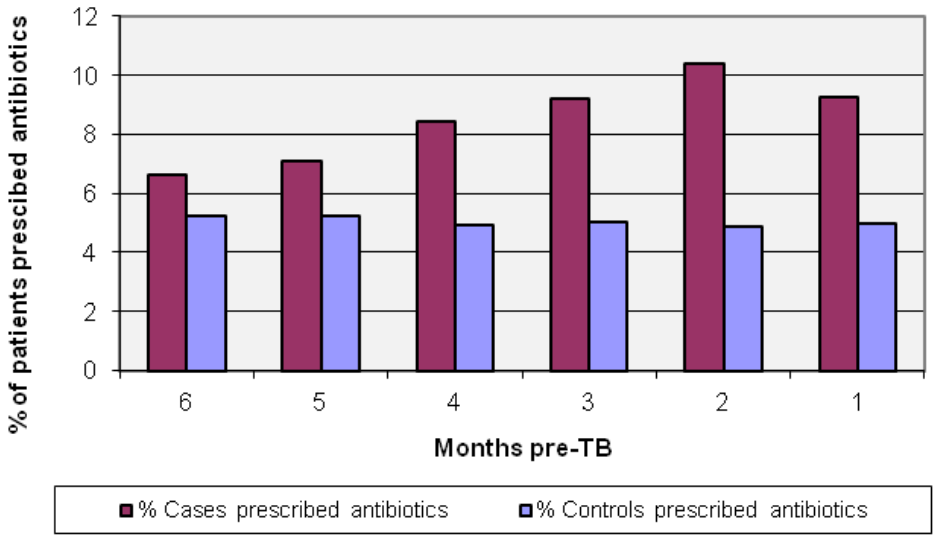


4.4.7 Analysis of Prescribing Data for Cases and Controls

Prescriptions for antibiotics were analysed in the six months prior to TB for cases and prior to a fixed index date for controls. 8.6% of cases and 5.1% of controls received antibiotics during the 6 month study period (for further details of the method see Chapter 3; 3.6.4), this represents a significant difference between case and control prescribing ($p<0.001$).

Analysis of antibiotic prescribing by month in the six months prior to TB showed that for cases there was a doubling of antibiotics prescribed over the 6 month period while for controls antibiotic prescriptions remained constant (see figure 4.10).

Figure 4.10 Percentage of cases and controls prescribed antibiotic in the 6 months pre-TB



4.4.8 Analysis of Risk Factors for TB

4.4.8.1 Single Variable Analysis of Risk Factors for TB

Variables (i.e. risk factors for TB) that might confound the findings of this case-control study were assessed individually using conditional logistic regression. The results are shown in Table 4.5.

Table 4.5 Results of single variable analysis of risk factors for TB

Variable / risk factor	CASES Indiv. with TB (n=3,032) No. (%)	CONTROLS no TB (n=15,160) No. (%)	Crude Odds Ratio (95% CI)	*p* value
Sex				
Male	1490 (49%)	7833 (52%)	**1.11** (1.02 – 1.20)	0.011
(*vs* female)	1542 (51%)	7327 (48%)	**1**	
Ever smoked (*vs* never smoked)	439 (15%)	1732 (11%)	**1.35** (1.20 – 1.52)	<0.001
Any chronic disease (*vs* no chronic disease)	1061 (35%)	2729 (18%)	**2.72** (2.48 – 2.97)	<0.001
Chronic respiratory disease (*vs* no CRD)	861 (28%)	1837 (12%)	**2.98** (2.71 – 3.28)	<0.001
Diabetes (*vs* no diabetes)	162 (5%)	475 (3%)	**1.78** (1.48 - 2.15)	<0.001
Chronic renal disease (*vs* no CRD)	31 (1%)	42 (0.3%)	**3.69** (2.32 – 5.87)	<0.001
Chronic heart disease (*vs* no CHD)	213 (7%)	708 (5%)	**1.58** (1.34 - 1.87)	<0.001

The odds of TB were found to increase with male gender, smoking, any chronic disease and each of four chronic disease groups analysed separately. The odds of TB was nearly three times higher for patients with a chronic disease, or pre-existing morbidity, than those without (OR:2.72); and similar for patients with 'chronic respiratory disease' (see Table 4.5).

4.4.8.2 Multivariable Analysis of Risk Factors for TB

Following single variable analysis a multivariable analysis was conducted using a conditional logistic regression model. All variables assessed in single variable analysis were included in the multivariable model as all were statistically significant at $p<0.05$. Table 4.6 shows the adjusted odds ratios for the variables: gender, smoking and 'any chronic disease'; and Table 4.7 shows the adjusted odds ratios for the variables: gender, smoking and each of the four chronic diseases.

Table 4.6 Results of first multivariable analysis of risk factors for TB

Variable / risk factor	Crude Odds Ratio (95% CI)	Adjusted Odds Ratio* (95% CI)	*p* value
Gender (male *vs* female)	1.11 (1.02 – 1.20)	**1.07** (1.00 – 1.17)	0.06
Ever smoked (*vs* never smoked)	1.35 (1.20 – 1.52)	**1.29** (1.14 – 1.45)	<0.001
Any chronic disease (*vs* no chronic disease)	2.72 (2.48 – 2.97)	**2.69** (2.46 – 2.95)	<0.001

** adjusted for all other variables in the model (gender, smoking and any chronic disease)*

Multivariable analysis showed that smoking (OR:1.29) and chronic disease (OR:2.69) were independent risk factors for TB. The risk of TB was found to be higher in men but this was not statistically significant (adjusted OR:1.07). The odds of TB was nearly three times higher in patients with a 'chronic disease' than in those without (adjusted OR:2.69). The crude odds ratios were reduced only very slightly by the multivariable analysis. Further analysis of the significance of 'any chronic disease' was undertaken and each individual chronic disease (respiratory, renal, cardiac and diabetes) was assessed separately against other variables (see Table 4.7).

Table 4.7. Results of second multivariable analysis of risk factors for TB

Variables	Crude Odds Ratio (95% CI)	Adjusted Odds Ratio* (95% CI)	*p* value
Gender (male *vs* female)	1.11 (1.02 – 1.20)	**1.09** (1.00 – 1.18)	0.038
Ever smoked	1.35 (1.20 – 1.52)	**1.28** (1.13 – 1.45)	<0.001
Chronic respiratory disease	2.98 (2.71 – 3.28)	**2.91** (2.64 – 3.20)	<0.001
Chronic renal disease	3.69 (2.32 – 5.87)	**3.20** (1.96 – 5.23)	<0.001
Chronic cardiac disease	1.58 (1.34 – 1.87)	**1.30** (1.10 – 1.55)	0.002
Diabetes Mellitus	1.78 (1.48 -2.15)	**1.54** (1.27 – 1.88)	<0.001

** adjusted for all other variables in the model (gender, smoking and each of four chronic diseases)*

The multivariable model adjusting for gender, smoking status and separate chronic diseases showed that chronic respiratory, renal and cardiac disease in addition to diabetes, male gender and smoking were all independent risk factors for TB. Adjusted odds ratios were very similar to the crude odds ratios for TB. The odds of TB was nearly three times greater for those with 'chronic respiratory disease' than those without (OR:2.91). The odds of TB was more than three times greater for those with 'chronic renal disease' however this had a wide confidence interval probably due to small numbers. This analysis shows that 'chronic disease' is strongly associated with TB and this is due in the main to 'chronic respiratory disease.'

4.5 Discussion

4.5.1 Key Findings

TB cases consulted their GP three times more often than controls in the six months before diagnosis. Consultation rates increased from four months pre-TB suggesting a diagnostic delay of up to 4 months but also that opportunities exist to diagnose TB earlier. When comparing cases and controls, cases have higher rates of consultation for TB symptoms and respiratory infections and twice the amount of chronic disease particularly chronic respiratory disease. This study shows the strength of association between TB and its classic symptoms and TB and common alternative diagnoses.

A separate analysis, using the same cases and controls, looked at risk factors for TB and showed that male gender, smoking and chronic disease (particularly chronic respiratory disease) were independent risk factors for TB. This is consistent with the research literature.

4.5.2 Consultations – General

Analysis of patient consultations showed that TB cases consulted their GP more often than controls. In the six months pre-diagnosis 80% of TB patients consulted their GP two or more times and a third consulted six or more times (four times more than controls). The increase in case consultation was seen from 4 months prior to diagnosis and was most marked 1 month pre-TB when cases consulted over 3 times more often

than controls. This data suggest that there was a delay in TB diagnosis of up to 4 months; this is consistent with research on diagnostic delay conducted by others.[8, 19, 20, 24, 109, 112]

There is therefore a window of opportunity in which GPs could intervene and diagnose TB earlier; particularly if diagnostic criteria are clear and GPs knowledge and awareness of TB is high. A delay in diagnosis is an issue for public health and the prevention and control of TB.

A possible explanation for repeat visits to a GP prior to a diagnosis of TB might be for investigation and result giving. However, this is an unlikely explanation in the context of this study because the rise in GP consultations is seen from about 4 months pre-TB, much earlier than would be explained by consultations for result giving. Furthermore, visits for investigations were removed during data cleaning as I considered these to be "non-GP consultations".

Further analysis of the rise in case consultation suggests that the rise occurred from about 7 months pre-TB (see Figure 4.5) but that it was always higher than that of controls. Possible explanations for the different consulting rate of cases and controls includes: firstly, TB cases have twice as many chronic diseases as controls; secondly, patients who consult more may be consulting due to a suppressed immune system leading to recurrent infections and may therefore be more prone to TB either through recrudescence or from new infection, and thirdly, it may represent the insidious onset of TB causing a non-specific illness that only gradually becomes increasingly symptomatic. This latter point is particularly evident in Figure 4.5.

In this study 40% of TB cases were diagnosed in primary care and possibly 56% if all those diagnosed by an unnamed source were also in primary care. This is similar to a study by White et al. that found 58% of patients in an inner city population had presented via primary care.[102] However this study finds less diagnosis occurring in primary care than a much smaller study by Paynter[109] where 80% of TB cases were diagnosed by GPs. This disparity may represent a difference in the 'recording' of where the diagnosis was made rather than the 'place' the diagnosis was actually made.

4.5.3 Symptom Consultations

Consultation rates for each of the classical symptoms of TB showed that in the 6 months prior to diagnosis, cough is the most commonly recorded symptom by GPs at 16%. If a patient is diagnosed with one or more TB symptoms, their odds of TB is eight times greater than those without these symptoms (i.e. those without TB).

Diagnosing TB in primary care can be challenging; the symptoms and signs are often nonspecific. Cough is a common problem in primary care and for the majority of people it will not represent TB but rather an acute respiratory infection or chronic respiratory disease. It is important therefore that GPs specifically ask about, and pick-up, the other signs and symptoms that help diagnose TB. This study shows that particular signs and symptoms; fever (OR:9.7) weight loss (OR:14.05) and lymphadenopathy (OR:32.9) are more strongly associated with TB than cough (OR:5.2). Although these symptoms have higher odds ratios than cough, they are much less common. Raising GP awareness of TB symptoms and the strength of association between these and a diagnosis of TB could help GPs diagnose TB earlier. Haemoptysis and sweating are much rarer diagnoses and if found, are more highly associated with TB with odds ratios of 32 and 70 respectively; the number of consultations, however, was small and confidence intervals wide. As this is a case-control study it is not possible to measure the positive predictive value of different symptoms directly. To assess this further a clinical cohort study, using a network of GP practices, would be needed to show what proportion of those consulting with a particular symptom or symptom combination are eventually diagnosed with TB.

Although symptom combinations are likely to increase the chance of diagnosing TB, this research found that very few GPs (0.5%) recorded two or more symptoms in a searchable form on the same day.

Interestingly, GPs record ‘symptoms’ alone much less often than ‘diagnoses’. Only a minority 27% of TB cases, had a TB symptom recorded in the 6 months pre-diagnosis as compared to 42% who had a ‘respiratory disease’ recorded (many of these patients would have presented with symptoms similar to those of TB). This tendency to code ‘diagnoses’ rather than ‘symptoms’ will have reduced the size of associations between classic symptoms and TB. But more importantly, the fact that only a minority of cases had symptoms recorded in a searchable form could mean that raising GP awareness of TB symptoms alone might be helpful for only a minority of cases but not the majority.

It is likely that further information was recorded in the free-text field, but this cannot easily be searched. This work suggests that raising awareness of TB symptoms and their strength of association is important but needs to be part of a package of measures to raise awareness among GPs of the local epidemiology of TB, risk factors for TB, other disease associations and the process of local referral for diagnostic tests.

Although this study did not explicitly examine the positive predictive value of symptoms (which would have required a different design) it is clear that for the cardinal symptoms the positive predictive value is very low (especially for the more common symptoms such as cough). In the UK the annual incidence of TB is around 15 per 100,000 population[47] whereas the proportion of the population presenting with any of the symptoms (based on our control data) is many times higher.

4.5.4 Respiratory Disease Consultations

This study showed that pre-diagnosis, TB cases have a much higher consultation rate for respiratory disease and respiratory infections than controls. In the six months pre-TB, 1 in 3 cases had a respiratory infection and the odds of TB in this group was seven times higher than in controls. As consultations for cough and respiratory infections may be different ways of coding the same thing, I combined these consultations into one variable 'respiratory infection or cough' for analytic purposes.

A third of TB cases with respiratory infections, and 11% of all TB cases, had three or more respiratory infections in the six months pre-TB. This suggests that repeat respiratory infections, particularly over a short period of time, and that do not respond to standard treatment, could be used to increase a GPs diagnostic suspicion for TB and could act as an alert or a 'red flag' for TB. On the other hand, GPs should also be mindful of the fact that patients with chronic disease are at an increased risk of TB and should not ignore TB symptoms in those with chronic disease.

High rates of respiratory infection are supported by the fact that cases, as compared to controls, are prescribed nearly twice as many antibiotics in the 6 months pre-diagnosis and that the receipt of these prescriptions rises gradually towards diagnosis. It is likely that patients are being misdiagnosed / misclassified with respiratory infections rather than TB and this may lead to a delay in TB diagnosis and an increase in antibiotic prescribing. This research supports that of others who have found that receiving

antibiotics prior to a diagnosis of TB delayed the start of TB treatment.[113,118,129-132,156] GPs record much more frequently that TB cases present with respiratory disease (42%) than with any of the classic TB symptoms such as cough (16%). This may be an anomaly of GP recording, in that GPs are more likely to code a 'diagnosis' e.g. 'LRTI' in preference to a symptom such as 'cough' which may just appear in the unanalyzed free-text field. This suggests that GPs may be misdiagnosing TB as a simple respiratory infection (LRTI), prescribing antibiotics and possibly delaying the diagnosis of TB.

One third of patients prior to a diagnosis of TB consult their GP more than 6 times in 6 months. These patients are increasingly older than those who consult less often, interestingly though, these patients also had less co-morbidity than all TB patients whatever their age. With increasing age one would expect an increase in co-morbidities rather than a decrease. The association between TB and chronic disease warrants further investigation, this I have attempted to do in the self-controlled study discussed in Chapter 5.

4.5.5 Reducing Diagnostic Delay

In my literature review, (see Chapter 2) I reviewed and discussed research that looked at diagnostic delay and TB in the UK and other high income countries. This case-control study indirectly reviews 'health service delay', defined as the time interval between the first contact with health services and the start of treatment;[20] but does not assess 'disease' or 'patient delay' as this information is not recorded in the GPRD.

My literature review quantified 'diagnostic delay' using 26 studies relevant to the UK setting and found that the median 'health service delay' varied from 6 to 39 days, with a mean of 26 days. In this study I did not specifically quantify diagnostic delay as it is difficult to know when 'health service delay' truly started from GPRD data and whether increased consultations pre-TB are specifically due to TB. However, one could extrapolate that diagnostic delay exists because patients consulted their GP nearly 3 times more often and with symptoms typical of TB prior to their diagnosis. The data suggests a delay of at least 4 months is not uncommon. Many of the studies reviewed in my literature review were retrospective cohort studies using data collected from cases specifically for TB research. This is very different to the data used in this study which

was data on all consultations for all diseases collected at the time of presentation in primary care.

A considerable amount of research has assessed risk factors for diagnostic delay in TB but rarely from a primary care angle. Clear cut determinants leading to increased diagnostic delay include: sputum smear negative TB,[111,121, 122] extra-pulmonary TB,[20,121] no respiratory symptoms,[111,114] a history of asthma,[114] HIV infection,[122] no chest x-ray at the first visit,[23,114] ethnic group (white>ethnic),[19] older age,[25,111] female gender[19,25, 111] and an alternative diagnosis for which an antibiotic is prescribed.[114,129] Understanding the risk factors that lead to a delay in TB diagnosis is crucial to ensure earlier diagnosis and treatment and so minimise transmission and reduce poor outcomes.

TB remains a diagnostic challenge in primary care because of its variable and non-specific features. Difficulty in recognizing symptoms as due to TB is likely to account for much of the diagnostic delay seen in this study. Health service delays occur when doctors fail to make the diagnosis, perhaps through inexperience, or failure to consider the diagnosis in patients perceived as being at low risk. Symptoms such as cough can also be explained by many diseases and are only rarely attributable to TB.

As this study and others show diagnostic delay is an issue in primary care. Ways to reduce diagnostic delay could include:

1. Increase awareness of TB symptoms and risk factors in health professionals to raise their index of suspicion for TB.[20, 21, 23, 114, 119]
2. Patient education to encourage earlier presentation to health services.[20, 21, 23, 114]
3. Improving access to health services for migrants and other at-risk groups.[119, 121, 133]
4. Encourage earlier GP referral for screening using standard sputum and chest x-ray methods.

Increasing awareness of TB in patients would help earlier presentation and therefore earlier diagnosis. This is not the sole responsibility of primary care, it is also that of public health. To this end *TB Alert*, a UK TB charity, has recently been asked by the Department of Health to produce literature on TB specifically for use in primary care.

4.5.6 Additional Comments

Only 6 cases, 0.2% of the TB cases, were known by GPs to be HIV positive. This is a large under-representation of the known dually infected population in the UK, estimated in 1998 to be 3.3% by Rose,[66] 5.7% by Ahmed in 2003[67] and 26% in London TB patients in 2003 by Rodger.[157] It is likely that, as the majority of data for this study was collected during the 1990s, patients may have been unaware of their HIV status as HIV testing was not routinely offered to subjects with active TB, that the GPRD TB subjects were less likely to be young non-UK born and so at lower risk perhaps of HIV co-infection, and finally that disclosure of HIV status to general practitioners was, over the time of this study period, not much performed either by secondary care HIV providers or HIV co-infected individuals.

The finding that 54% of cases appear to have been prescribed TB treatment in primary care according to GPRD data (BNF code group 05.01.09.00) is higher than expected. It is unlikely that TB treatment was initiated in primary care. Most drugs for TB are started in secondary care and continued by secondary care via the TB nurse to ensure compliance and free prescription. However, TB prescriptions may appear in the GPRD drug record not because they are prescribed by GPs but rather to remind GPs of possible drug interactions when they prescribe additional medication. GPs 'flag' prescriptions electronically in different ways and much of the information as to whether prescriptions had been truly prescribed is in the free text fields which I did not have access to.

4.5.7 Risk Factors for TB

This study is one of the first to examine risk factors for TB in the UK on such a large scale. Multivariable analysis of case and control variables showed that male gender, smoking and individual chronic diseases were independent risk factors for TB. These results are important as most health care workers think of 'ethnicity' alone as a risk factor for TB but rarely the others.

The observed association of TB with male gender is in agreement with that reported by other studies where TB notification rates are higher for men than for women.[47,158,159]

This research supports that of others that pre-existing morbidity and a number of chronic diseases increase the risk of TB. There is strong evidence of an association

between TB and diabetes,[160,161] TB and malignancy,[162-164] and TB and chronic respiratory disease.[165,166] Although chronic renal disease has a stronger effect (OR 3.2), chronic respiratory disease (OR 2.91) is more common and so the population impact of this is greater. Multivariable analysis therefore suggests that 'chronic disease' is strongly associated with TB, and that this is mostly due to 'chronic respiratory disease.'

At diagnosis a third of TB cases have at least one chronic disease; this was twice that of controls, with chronic respiratory disease being the most common for cases and controls. This association between chronic respiratory disease and TB is confounded by smoking but may be explained in two ways. Firstly, that patients with chronic respiratory disease have a higher likelihood of developing TB due to the chronically damaged nature of their lungs[165,167,168] and secondly, by GP misdiagnosis in the six months prior to a TB diagnosis: the more a patient consults with a cough the greater the likelihood a GP will diagnose a chronic respiratory disease when in fact the case is presenting with the early symptoms of TB.

This study confirms that smoking is an independent risk factor for TB and supports the work of others.[167,169,170] Smoking status (whether smoker or non-smoker) is known to be a poorly recorded variable in the GPRD dataset; Osborn et al. showed this but also that there was no systematic difference between cases and controls.[82] In this study fewer than 40% of individuals had a smoking status recorded; an association, however, between smoking and TB is still seen. Fifteen percent (n=439) of TB cases were recorded to have smoked as compared to eleven percent (n=1732) of controls. The GPRD appears to under-record smokers as exemplified by the fact that in 2007 21% of the UK population were considered 'current smokers' (ONS online 09.10.09 - http://www.statistics.gov.uk/cci/nugget.asp?id=313-). Recording bias of TB case smoking status may represent a limitation in this study. I have shown that TB cases are likely to present to their GP with a respiratory complaint and therefore GPs are likely to ask about their smoking status and so may record smoking status more often for cases than controls. The data used in this research is pre-2002. In recent years the recording of smoking status by GPs has greatly improved as quality improvement initiatives have taken hold in primary care.

These findings, along with those of others, that smoking is a risk factor for TB support the hypothesis of an increased vulnerability of smokers to TB infection and the

development of clinical disease.[167,168] Furthermore, research by Pasipanodya shows that survivors of symptomatic TB have long term lung damage and are 5.4 times more likely to have abnormal lung function tests than those with latent TB.[166] Smoking is one of the most modifiable of all the risk factors for TB. Primary care teams, by encouraging and supporting TB patients to stop smoking, could therefore have a key role in improving the outcome of patients that smoke.

So in summary, this study shows that male gender, smoking and individual chronic diseases were independent risk factors for TB in the UK and that raising awareness of these risk factors (along with those found by others[133]) could help primary care teams diagnose TB earlier and reduce poor outcomes.

4.5.8 The Role of Primary Care in the Diagnosis of TB

I reviewed the role of primary care in TB in Chapter 2 and concluded that this area warranted greater investigation due to the lack of published research, the need to reduce diagnostic delay and the major role primary care could play in this.[8,10,101,103,108] This study was designed to meet these needs and to investigate TB, its diagnosis and its risk factors, from a primary care perspective.

A small study by Metcalf et al. specifically looked at the diagnosis of TB from a primary-care perspective.[8] This study supports Metcalf's finding that GPs see a huge variety of TB presentations and that making a timely diagnosis is difficult. Metcalf identified barriers to prompt diagnosis as: "atypical presentations, low level of clinical suspicion of TB by GPs, lack of continuity of care, workload demands that limit time with patients and sub-optimal clinical-patient communication." This study confirms some of these barriers but also goes further and provides GPs with information on the strength of association between TB its classic symptoms and alternative diagnoses plus it elucidates risk factors for TB. Metcalf et al. suggest "a better understanding of the processes of diagnosis could highlight existing good practice and opportunities for enhanced care."[8] This study tries to do this.

Early diagnosis is one of the most important contributions primary care can make to improved TB control.[8,10,101,103,108] To assist GPs in making an earlier diagnosis, awareness of TB needs to be raised. This study suggests that awareness of the following could be used to raise a GPs level of suspicion for TB:

- Recurrent chest infections over a few months at a level higher than expected for this patient, or more than three chest infections in a 6 month period
- Respiratory symptoms not responding to antibiotics
- Recurrent cough (to trigger assessment for the other classic symptoms of TB, a chest x-ray and sputum specimen)
- Any of the rarer TB symptoms i.e. weight loss, haemoptysis or lymphadenopathy
- Repeat, unexplained visits to the GP for non-specific symptoms
- Repeat, unexplained visits to the GP in which no diagnosis is made
- Multiple visits over a few months at a level higher than expected for this patient
- In GP practices, or geographic areas, with a higher proportion of ethnic minorities, TB should be high on any list of differentials for patients with an unexplained or unresolved cough or any of the classic symptoms of TB.

Combining the knowledge of these 'red-flags' with the other features highlighted by this research such as the risk factors for TB (male gender, older age, suffering a chronic disease), the strength of association between clinical presentations and TB and raising awareness of newer investigations and the local referral process may help GPs make an earlier diagnosis. Practically the recurrent nature of non-specific, non-resolving symptoms or recurrent respiratory infections, particularly in patients with a known chronic disease, would be something that GP computer systems could flag and in so doing raise a GPs level of suspicion for TB.

Raising awareness of TB in general practice is vital if we are to combat the increasing incidence of TB in the UK. This is supported by the limited existing literature on TB in primary care.[11-13,74] There are numerous ways of raising awareness: online educational packages / articles in peer reviewed journals / courses / locally run TB awareness training etc. One example in the literature specific to TB and primary care showed that an educational intervention promoting information about TB and TB screening improved identification of active and latent TB in new GP registrants and increased BCG coverage.[100]

4.5.9 Strengths and Limitations of this Study

A huge strength of this study is its very large size with 3,032 TB patients and 15,160 controls. The findings are therefore extremely unlikely to have occurred by chance.

Bias is a potential limitation. To limit ***selection bias*** I randomly selected five controls per case and these were matched on age and GP practice with the assumption that patients registered with the same GP were more likely to be from a similar socioeconomic background and ethnic group than patients selected from other practices. Given the fact that most practice populations are ethnically and socially diverse the matching at practice level will at best only partially account for these factors. The lack of ethnicity data in the GPRD dataset is a major issue for this study as TB in the UK is much more common in ethnic minority groups and particularly those born abroad in countries with a high incidence of TB.[47, 60, 171] Not knowing the ethnicity of cases meant controls could not be matched on this and selection bias may have resulted. As this study excludes patients not registered with a GP and those registered for less than 18months, it is likely to exclude homeless people and ethnic minority new-entrants both of whom are at greater risk of TB than the general population.[47,59,133] The results of this study may therefore underestimate the true picture of TB in the community and prior to diagnosis in part due to selection bias.

GP recording bias and misclassification – the recording of consultations by the GP may have led to recording bias with the under-recording of symptoms. This study analysed both symptoms and signs as recorded by GPs. GPs record these electronically under a specific 'coded' heading from a drop-down list. In the 6 months pre-TB, GPs may record a patient's consultation using a 'symptom' such as 'cough' or a 'specific diagnosis' such as 'pneumonia'. The number of symptoms coded in the GPRD is much fewer than the number of respiratory infections diagnosed. For example 497 patients were recorded to have consulted with 'cough' compared to 1177 with a respiratory infection, most of which are likely to have presented with a cough. This probably reflects the fact that GPs more readily record a collection of symptoms as a "coded diagnosis"; for example "LRTI" and then record any symptoms in the GPRD free-text field rather than separately. For this reason symptom analysis probably provides an underestimate of the true occurrence of each symptom reported to a GP due to under recording. Only 27% of patients had TB symptoms recorded rather than a specific

diagnosis. This does not mean that only 27% of TB patients had symptoms pre-diagnosis, it reflects how GPs record their consultations. To reduce the impact of consultation recording bias, I conducted some additional analyses that combined all respiratory consultations regardless of whether they were for 'cough' or a specific 'respiratory disease'. However, a recommendation for future work would include further analysis of the GPRD free-text fields looking for data on specific TB symptoms in patients subsequently diagnosed with TB.

Recording bias of the outcome is also a significant problem as over 50% of cases were not classified as 'pulmonary' or 'extra-pulmonary' TB, thereby limiting the opportunity to explore how consultation patterns differ for these very different clinical manifestations of TB.

The GPRD with its national coverage of representative GPs has been shown to be broadly representative of the UK population;[76] and Hansell et al.[80] showed similarities between the rate of TB diagnosis in the GPRD and the fourth Morbidity Survey in General Practice (MSGP4).[89] Both these studies suggest that the results of this case-control study should be generalisable within the UK setting. However, the findings may not be generalisable to countries with different primary care structures. A possible limitation of this study is that TB cases may be under-represented in the GPRD dataset. Historically there have been only a few GP practices that enter data in to the GPRD in London.[76] As London has the highest TB incidence in the UK this might lead to an under-representation of TB cases. However, unless the diagnostic delays and risk factors for disease are very different in London compared to elsewhere, this will not greatly affect the generalisability of these findings.

Finally, a case-control study design allows estimates of risk to be made but not of absolute risk – therefore it is of limited value for assessing the positive predictive value of any particular symptom-set where a cohort study would be preferable.

So in summary there are a number of limitations to this study. To reduce control selection bias and the effects of ethnicity and socioeconomic factors, I conducted a second similar study (see Chapter 5) using self-controlled methodology, where a case becomes its own control in a different time period. The results of the self-controlled

methodology can then be compared to the case-control study to assess whether the limitations are a serious concern.

4.5.10 The Use of the GPRD as a Research Tool

The General Practice Research Database is a powerful tool for epidemiological and primary care research. It has been validated for use in respiratory epidemiological studies by Hansell[80] so is eminently suitable for this study. It has been used for many high quality studies including some respiratory studies.[78, 80, 83, 90-92] To further validate the use of the GPRD for this study, control patient consultation rates were compared to the average number of GP consultations per year in a national primary care study. In the Fourth National Study of Morbidity Statistics from General Practice (MSGP4), 1991-1992, 78% of patients consulted at least once during the year and each consulting patient had an average of 3.8 consultations per year.[89] In my case-control study, control patients consulted a median of 2 times in the 6 months before a fixed index-date, which extrapolates to 4 consultations per year. So consultation rates were found to be similar in the Fourth National Study of 'Morbidity Statistics from General Practice' and in this study.

The strengths of the GPRD for this study include its size, its comprehensive coverage of the UK population, the fact that it contains continuous, longitudinal primary care data which is high-quality and validated and that it has been previously used in many epidemiological studies, particularly respiratory studies. However, its major limitations for a study looking at TB are its lack of ethnicity and socioeconomic data, HIV data and variations in recording.

4.6 Conclusion

TB cases consult more often than controls, with this increasingly evident from 4 months pre-TB. They have higher rates of consultation for TB symptoms and respiratory infections and twice the amount of chronic disease. Diagnosing TB in primary care is a challenge due to its variable and non-specific nature but a window of opportunity exists from 4 months. GPs need support to make a TB diagnosis earlier. To this end, this study provides information on the strength of association between symptoms, respiratory

diseases and TB and suggests that patients who consult on multiple occasions, particularly with cough or recurrent respiratory infections, and without explanatory long-term chronic disease should be carefully assessed for the symptoms & signs of TB.

Awareness needs to be raised among GP's of indicators that could increase their level of suspicion for TB such as key symptoms, signs and diagnoses associated with TB to aid earlier diagnosis and so reduce diagnostic delay. But, as only a minority 27% of cases have TB symptoms recorded, raising GP awareness of TB symptoms alone may only be helpful for a minority of cases. Additional work will therefore be needed to improve GPs knowledge and awareness of TB, the risk factors for TB, its local epidemiology and management.

People with TB tend to be older, male and have more chronic disease than those without TB. This is consistent with the research literature.

To confirm the significance of the consultation rates and the other findings in this study and to check that chronic disease, ethnicity and socioeconomic status were not confounding the results I went on to conduct a self-controlled case study (see Chapter 5).

Chapter 5

Consultation Behaviour in Primary Care Prior to a Diagnosis of TB: A Self Controlled Study

5.1 Introduction

TB in the UK is much more common in ethnic minority groups[47,60] and those on a low income.[59] The General Practice Research Database (GPRD) can be used as a tool to study TB epidemiology from a primary care perspective. Certain key variables, however, such as ethnicity and socioeconomic status are missing. To strengthen the findings of my case control study and minimise confounding by known and unknown variables such as ethnicity, socioeconomic status, prior chronic disease and smoking I conducted a self controlled (case-only) study.

5.2 Aims & Objectives

5.2.1 Aims

The aims of this study were to assess the extent of missed opportunities for a diagnosis of TB in primary care using a self-controlled study and to describe the nature of GP consultations.

5.2.2 Objectives

- To measure excess GP consultations in the 6 months prior to diagnosis in TB cases as compared to 12 to 18 months prior to diagnosis in the same individuals
- To examine consultation behaviour and consultation type
- To compare the results of this self controlled study with the case-control study of Chapter 4

5.3 Methods

Study Design

A self controlled (case-only) study based on case-crossover methodology[83,140-142,143] was conducted to investigate consultation behaviour prior to a diagnosis of TB (see Chapter 3, Section 3.6.2 for more detail).

Setting

I examined data, as previously described in Chapter 4, Section 4.3, from 743 general practices in the UK who provided data to the General Practice Research Database (GPRD) during the time period April 1990 to April 2002.

Data Source

I used the GPRD (see Chapter 4, Section 4.3).

Participants

Patients were selected from the GPRD between 1990 and 2002;

Cases: were GPRD patients with a medical code for active TB i.e. the same TB cases as were used in the case-control study presented in Chapter 4 (n=3023).
Controls: were cases acting as their own control in a time period 12-18 months prior to their diagnosis of TB i.e. self controls.
'Exclusion' criteria: I excluded patients who had less than 18 months of pre-diagnosis data (see Chapter 3, Section 3.4.3).

Measurements

I compared the consultation behaviour of TB cases in two different time periods one 12-18 months pre-TB and the other 0-6 months pre-TB. I analysed 'all' GP consultations and GP consultations for specific TB symptoms and respiratory diagnoses (see chapter 3 for more details). By comparing consultations for the classic symptoms of TB and diagnosed respiratory conditions in two different time periods, I generated a summary measure the incidence rate ratio (IRR). The IRR represents statistically the comparison between consultations occurring 12-18 months pre-TB (when the patient is TB-free) and 0-6 months pre-TB (when the patient may be developing TB). It gives an indication of which symptoms and respiratory diseases are most associated with TB.

Bias

I used a self-controlled study to eliminate control selection bias, remove confounding by permanently fixed factors e.g. gender, ethnicity, genetics; and minimise confounding by factors that vary little over the comparison period e.g. chronic illness, age, social class, life-style.

Statistical Analysis

Statistical analysis was performed using STATA version 9.0 (Stata Corp, College Station, TX, USA).[155] Consultation rates for cases in two different time periods were calculated and compared, generating rate ratios. GP consultations for specific symptoms and respiratory diagnoses were analysed using the Generalised Estimating Equation (GEE) Model.[22,24,25,147] The GEE model compares total counts for each individual and then calculates a summary measure, the incidence rate ratio (IRR) for all individuals with a 95% confidence interval. A *p* value of 0.05 or less was considered statistically significant. In this study, in practical terms the GEE actually calculates a rate ratio: the ratio of consultations in two different time periods not a true incidence or IRR as consultations may or may not in fact be incident. For further details of the method see Chapter 3, Section 3.6.2.

5.4 Results

5.4.1 Characteristics of Study Patients

3032 TB cases with over 196,000 consultations were analysed for the time period 1990-2002 (see Chapter 4, Section 4.5.1).

5.4.2 Consultation Data – General

In the 6 months prior to a diagnosis of TB, 3032 TB patients had 18,493 consultations recorded by their GP. In the 12-18 months pre-TB TB patients had 9,469 visits recorded by their GP. The median number of consultations per patient 0-6 months pre-TB was 5 with an inter-quartile range of 3 to 8; and for 12-18 months pre-TB the median number of consultations per patient was 3 with an inter-quartile range of 2 to 5 (see Figure 5.1).

10% of cases did not consult at all in the six months prior to diagnosis and 12% only once.

Figure 5.1 Median number of consultations per patient 12-18months & 0-6months pre-TB

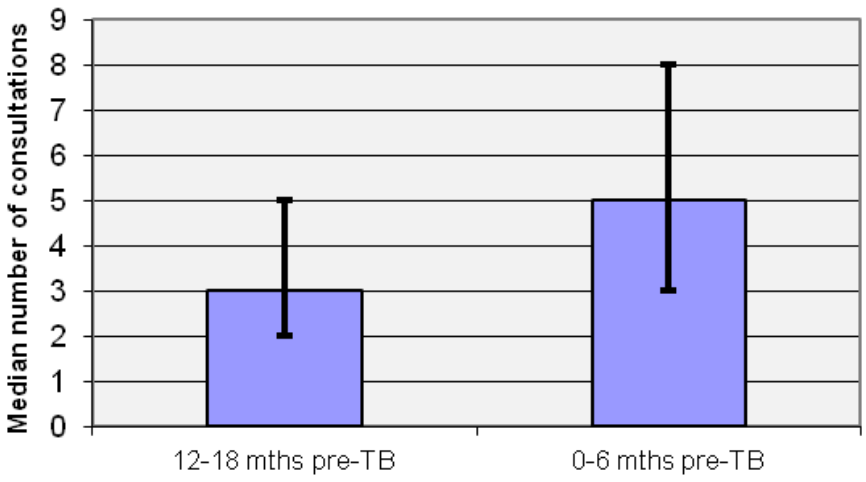


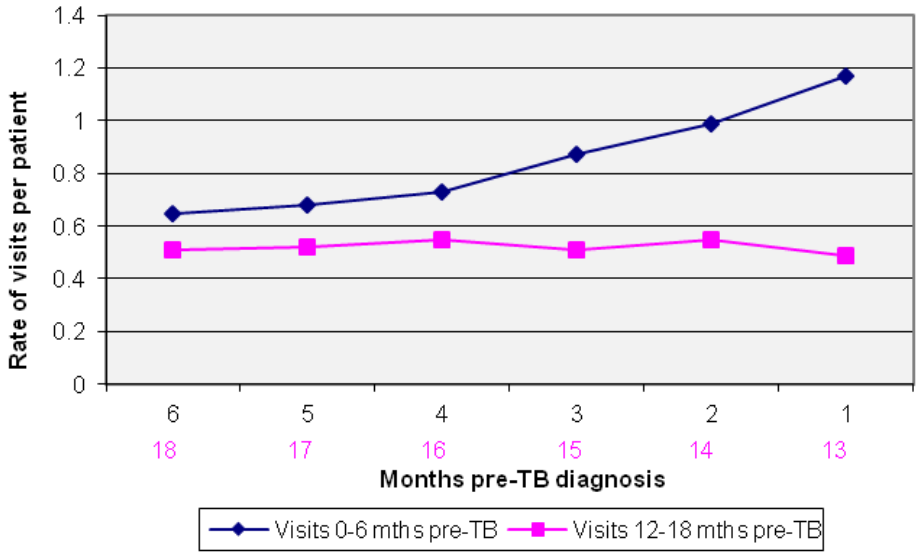
Table 5.1 shows a comparison of consultation rates 12-18 months and 0-6 months pre-TB by month and the total for each six month period. The number of GP consultations increased pre-TB; at 6 months pre-diagnosis it was a little higher (27% higher) than baseline levels, but the increase was most evident from 3 months (RR 1.7) and most marked 1 month prior to a diagnosis of TB when patients consult more than twice as often (RR 2.39) (see Figure 5.2). Overall a 63% increase in consultations was seen between the two time periods (RR 1.63).

Table 5.1 **Comparison of GP consultation rates for the same patient in two time periods pre-TB, one year apart expressed as a rate ratio**

	CASES 12-18 months pre-TB		CASES 0-6 months pre-TB		
CONSULTATIONS	Number	Consults per case	Number	Consults per case	**Consultation Rate Ratio** (95% CI)
Total consultations 12-18 and 0-6 months pre-TB	9,469*	3.1	15,461	5.10	**1.63** (1.59-1.68)
Consultations per month					
13mths preTB cf. 1mth preTB	1491*	0.49	3560	1.17	**2.39** (2.25-2.54)
14mths preTB cf. 2mths preTB	1661	0.55	3013	0.99	**1.81** (1.71-1.93)
15mths preTB cf. 3mths preTB	1548	0.51	2636	0.87	**1.70** (1.60-1.81)
16mths preTB cf. 4mths preTB	1661	0.55	2225	0.73	**1.34** (1.26-1.43)
17mths preTB cf. 5mths preTB	1571	0.52	2071	0.68	**1.32** (1.23-1.41)
18mths preTB cf. 6mths preTB	1537	0.51	1956	0.65	**1.27** (1.19-1.36)

***** *includes only consultations occurring prior to a TB diagnosis*

Figure 5.2 **Rate of GP consultations per patient per month 0-6 & 12-18 months pre-TB**



5.4.3 Analysis of Consultations for the Classic Symptoms of TB

GP consultations for the classic symptoms of TB were compared 0-6 months before diagnosis with those during a TB-free period 12-18 months before diagnosis. Nearly a third (27%) of TB patients had symptoms suggestive of TB 0-6 months pre-diagnosis as compared to 9% in the TB-free period. Table 5.2 shows the total number of individuals consulting their GP with the classic symptoms of TB 12-18 months and 0-6 months pre-diagnosis. The relative importance of the individual symptoms and a later diagnosis of TB are shown by the incidence rate ratio (IRR).

Analysis using the Generalised Estimating Equation (GEE) Model shows that in the 6 months pre-TB, patients had a 3.5 fold increase in consultation rate for 'any' of the classic symptoms of TB as compared to the TB-free period (IRR 3.48 95% CI:3.03-4.00) and that patients had a 2.5 fold increase in consultation rate for cough (IRR 2.5). Consultation rates for the individual cardinal symptoms of TB were significantly higher in the 6 months pre-TB than the TB-free period with sweats (IRR 13.0), lymph nodes (IRR 7.2) and weight loss (IRR 8.9) showing a particularly strong association with a diagnosis of TB. These symptoms were more strongly associated with a diagnosis of TB than fever and cough, as shown by the incidence rate ratios in Table 5.2, but are less often recorded by GPs.

Table 5.2 A comparison of TB symptom consultations for cases 12-18 and 0-6 months prior to a diagnosis of TB expressed as an IRR

Symptoms	**Individuals with symptoms 12-18 mths pre-TB** n=3032 No. (%)	**Individuals with symptoms 0-6 mths pre-TB** n=3032 No. (%)	**Incidence Rate Ratio** (95% CI)	***p* value**
Cough	233 (7.7%)	497 (16.4%)	**2.51** (2.15 - 2.93)	<0.001
Fever	25 (0.8%)	103 (3.4%)	**4.71** (2.99 - 7.42)	<0.001
Sweating	3 (0.1%)	28 (0.9%)	**13.00** (3.86 - 43.80)	<0.001
Lymphadenopathy	23 (0.8%)	167 (5.5%)	**7.15** (4.16 - 12.29)	<0.001
Haemoptysis	0	13 (0.4%)	**-**	-
Weight loss	15 (0.5%)	114 (3.8%)	**8.93** (5.19 - 15.38)	<0.001
Any TB symptom	284 patients* (9.4%)	832 patients * (27.4%)	**3.48** (3.03 – 4.00)	<0.001

** some patients had more than one symptom recorded*

Analysis of recurrent GP consultations pre-diagnosis revealed that patients present more often with recurrent cough and lymphadenopathy. Patients present rarely more than once with fever, sweats or haemoptysis (see Chapter 4).

In the 6 months pre-diagnosis patients were twice as likely to have a single classic symptom of TB and were 4.5 times more likely to have more than one TB symptom than in a TB-free period, see Table 5.3.

Table 5.3 Rate Ratios for multiple TB symptoms

	12-18 months pre-diagnosis No. (%)	**0-6 months pre-diagnosis** No. (%)	**RR**
Patients with one TB symptom	212 (7%)	485 (16%)	2.29
Patients with more than one TB symptom	72 (2%)	326 (11%)	4.53

5.4.4 Analysis of Respiratory Consultations

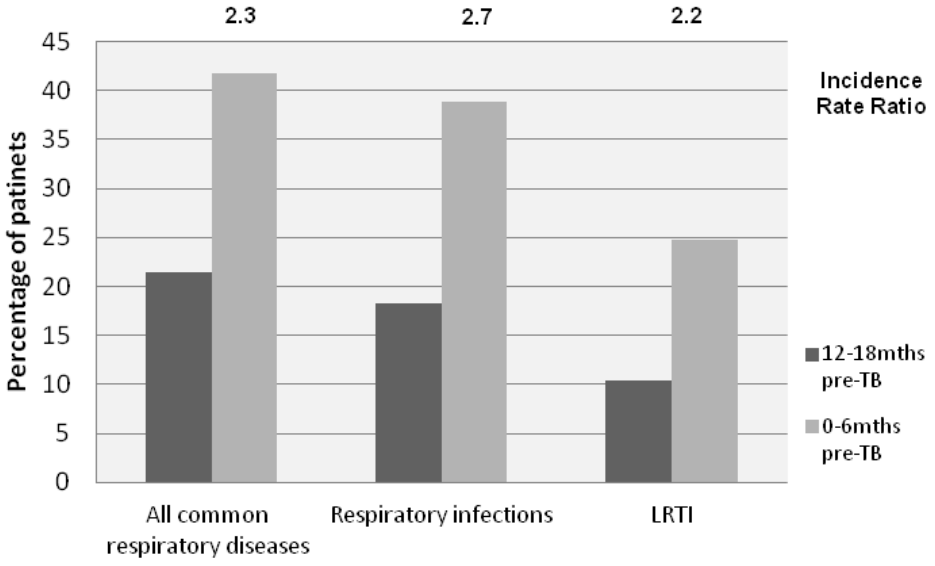
Nearly half (42%) of TB patients were diagnosed with a respiratory problem in the 6 months pre-TB. One in three patients (39%) presented with a respiratory infection and of these over half were lower respiratory tract infections (see Table 5.4). Analysis using the Generalised Estimating Equation Model showed that in the 6 months pre-TB patients had a two fold increase in consultation rate for 'any respiratory disease' as compared to the TB-free period (IRR 2.28; 95% CI:2.1-2.5); and a nearly 3 fold increase in consultation rate for 'respiratory infections & cough' (IRR 2.73; 95% CI:2.5-3.0).

Table 5.4 **A comparison of respiratory disease consultations for cases 12-18 and 0-6 months prior to a diagnosis of TB expressed as an IRR** (n=3032)

Respiratory disease (excluding TB)	**Individuals with symptoms 12-18 mths pre-TB** No.* (%)	**Individuals with symptoms 0-6 mths pre-TB** No. (%)	**Incidence Rate Ratio** (95% CI)	***p* value**
All respiratory diseases (incl. infections)	650 (21%)	1268 (42%)	**2.28** (2.07 – 2.50)	<0.001
All respiratory infections (incl. cough)	553 (18%)	1177 (39%)	**2.73** (2.47 – 3.01)	<0.001
LRTI	315 (10%)	751 (25%)	**2.20** (1.95 – 2.49)	<0.001

** Consults by individual = Number of individuals consulting at least once with condition*
LRTI = lower respiratory tract infection

Figure 5.3 **A comparison of respiratory diagnoses for patients 12-18 and 0-6 months pre-TB (showing IRRs)**



5.4.5 Stratification of Patients by TB Type

Data for the 3032 TB cases was re-analysed according to TB type. There are 305 different diagnostic codes for TB in the GPRD (see Appendix 2), I reclassified these codes under three headings – pulmonary TB, extra-pulmonary TB and 'unclassified' TB (i.e. site not otherwise specified) to try to understand how much each type of TB contributed to the overall symptoms and diagnoses seen in the 6 months pre-diagnosis. Table 5.5 shows a summary of the characteristics of TB cases by TB type.

Table 5.5 Characteristics of TB patients by TB type

	All TB n=3032 No. (%)	**Unclassified TB** n=1774 (58%) No. (%)	**Extra-pulmonary TB** n=483 (16%) No. (%)	**Pulmonary TB** n=759 (25%) No. (%)
Sex				
Female	1490 (49%)	820 (46%)	301 (62%)	359 (47%)
Male	1542 (51%)	954 (53%)	182 (38%)	400 (53%)
Age				
<=20	427 (14%)	316 (18%)	57 (12%)	48 (6%)
>20 - <=40	651 (22%)	397 (22%)	132 (27%)	121 (16%)
>40 - <=60	803 (26%)	447 (25%)	141 (29%)	211 (28%)
>60 - <=80	929 (31%)	515 (29%)	114 (24%)	297 (39%)
>80	222 (7%)	100 (6%)	39 (8%)	82 (11%)
Chronic Disease				
Any	35%	610 (34%)	130 (27%)	328 (43%)
Respiratory	28%	498 (28%)	88 (18%)	268 (35%)
Renal	1%	20 (15%)	7 (1%)	4 (<1%)
Cardiac	7%	109 (6%)	34 (7%)	69 (9%)
Diabetes	5%	91 (5%)	27 (6%)	44 (6%)
Ever smoked	439 (15%)	248 (14%)	70 (14%)	118 (16%)

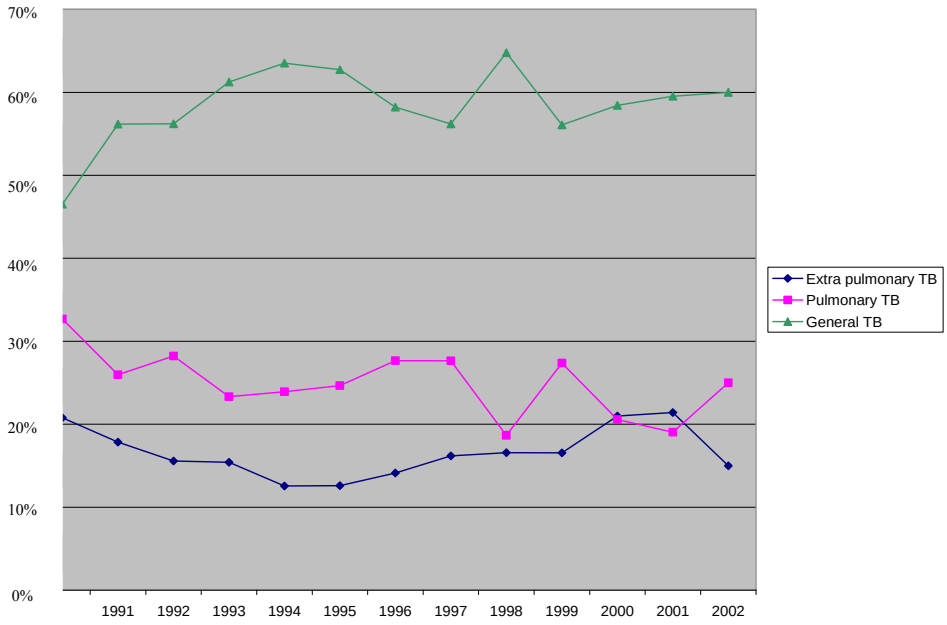
** 16 patients had a tuberculoma of unknown site and are not included*

Table 5.5 shows that GPs tend to record most TB patients as having 'unclassified TB' (58%). GPs specify the site of disease as either pulmonary or extra-pulmonary in less than half their patients. For those patients recorded as having extra-pulmonary TB there is a marked difference in gender: 62% are female and 38% male as compared to 47% female and 53% male in pulmonary TB; and age: extra-pulmonary TB is more common in the younger age groups, 39% under 40 compared to 22% for pulmonary TB. In terms of chronic disease, patients with extra-pulmonary TB have less chronic disease than

patients with pulmonary TB. As the characteristics of 'all TB' and 'unclassified TB' were similar, I analysed in more detail only those cases sub-classified into extra-pulmonary and pulmonary TB.

Figure 5.4 shows how GPs recorded the specific types of TB from 1990 to 2002. Over the 12 years analysed there was no significant improvement in the specificity of diagnostic coding by GPs.

Figure 5.4 Diagnosis by TB type, by year – 1990 to 2002



5.4.6 Consultation Rate by TB Type

Table 5.6 shows the GP consultation rate for different types of TB during the two time periods of interest, 12-18 months and 0-6 months pre-TB. These varied from 3.12 to 3.37 in the 12-18 months pre-TB and from 4.83 to 5.49 0-6 months pre-TB. Rate ratios for each TB-type show that patients consulted between 51% and 63% more in the 6 months immediately pre-TB depending on their TB-type (see Table 5.6). Further analysis of consultation data showed that there was no significant difference between different TB-types and multiple GP consultations pre-TB (see Appendix 5)

Table 5.6 GP consultation rates by TB type comparing 12-18 and 0-6 month's pre-TB

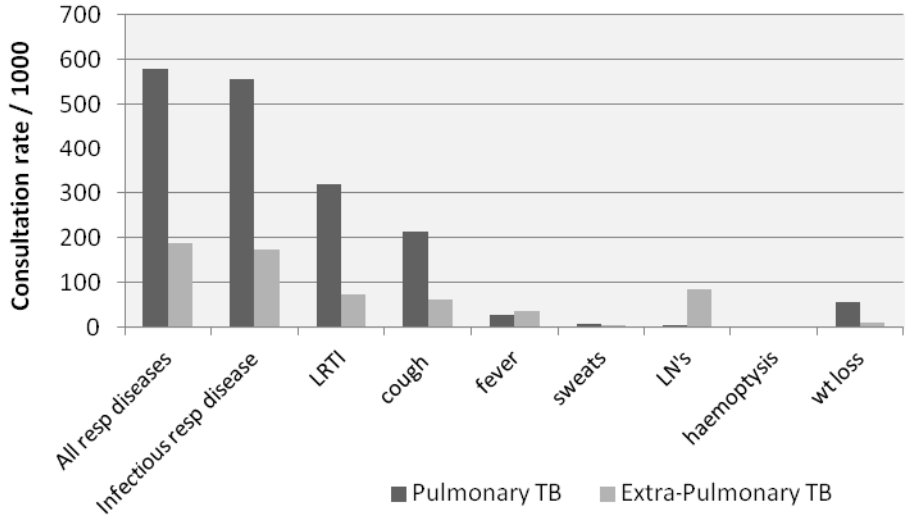
TB type	Number of patients	CASES 12-18 mths pre-TB Consults / case	CASES 0-6 mths pre-TB Consults / case	Rate Ratio
ALL TB	3032	3.12	5.10	**1.63**
Pulmonary TB	759	3.37	5.49	**1.63**
Extra-pulmonary TB	483	3.19	4.83	**1.51**

5.4.6.1 Symptom and Disease Consultations by TB Type

For patients presenting with different types of TB, consultations were analysed by classic TB symptoms and respiratory diagnoses in the 6 months prior to a diagnosis (see Figure 5.5). Consultations for cough were much more frequent for pulmonary TB than for extra-pulmonary TB and consultations for lymphadenopathy were much more common in extra-pulmonary TB (85 per 1000) than pulmonary (5 per 1000). Weight loss in extra-pulmonary TB was seen in 10 per 1000 as compared to 57 per 1000 in pulmonary TB. Consultation rates for fever and sweats were very similar for all types of TB. Consultation rates for 'all respiratory diseases', whether infectious or non-infectious, were more than twice as high for pulmonary TB as they were for extra-pulmonary TB.

So, in summary, patients with pulmonary TB had higher consultation rates for infectious respiratory disease, cough and weight loss. Patients with extra-pulmonary TB had higher consultation rates for lymphadenopathy.

Figure 5.5 Consultation rates for respiratory diseases & TB symptoms 0-6 months pre-TB analysed by TB type

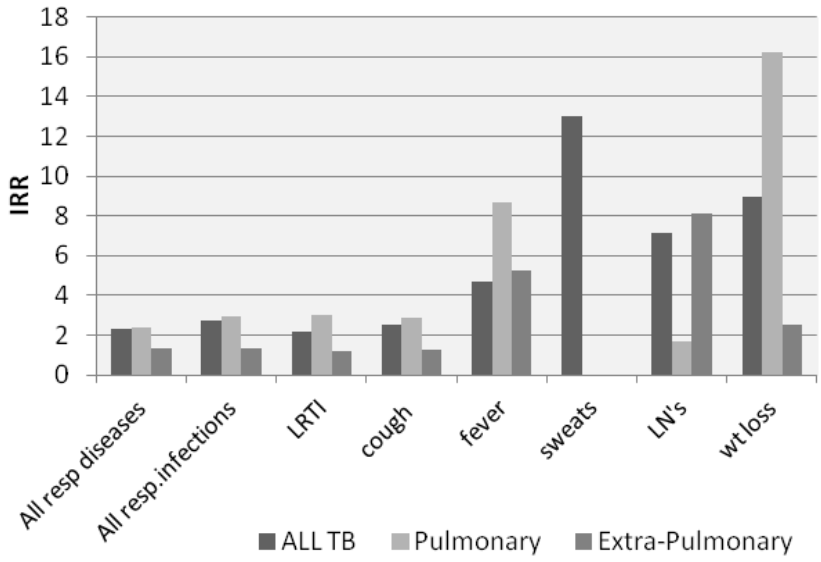


GP consultations for the classic symptoms of TB and respiratory disease were then compared in the two time periods of interest, 0-6 months pre-TB and 12-18 months pre-TB (a TB-free period), using a GEE model. The incidence rate ratio revealed that although sweats, lymph nodes and weight loss were useful indicators of TB in general terms, specific symptoms were more suggestive of particular types of TB.

For extra-pulmonary TB, consultation rates for lymphadenopathy were 8 times higher 0-6 months pre-TB than in a TB-free period (IRR 8.11; 95%CI: 2.38-27.69); and consultation rates for fever were 5 times higher (IRR 5.25). This suggests that lymph nodes are more predictive of extra-pulmonary TB than fever and weight loss (see figure 5.6 & 5.7).

For pulmonary TB, consultation rates for fever were 9 times higher 0-6 months pre-TB than in a TB-free period (IRR 8.67; 95%CI: 2.58-29.07); and consultations for weight loss were 16 times higher (IRR 16.25; 95%CI: 5.80-45.56). This suggests that fever and weight loss were more predictive of pulmonary TB than lymph nodes (see figure 5.6) although wide confidence intervals are noted.

Figure 5.6 Comparison of incidence rate ratios for consultations 0-6 months compared to 12-18 months for patients with different types of TB



** See Appendix 6 for tabulation of IRRs that create this graph*

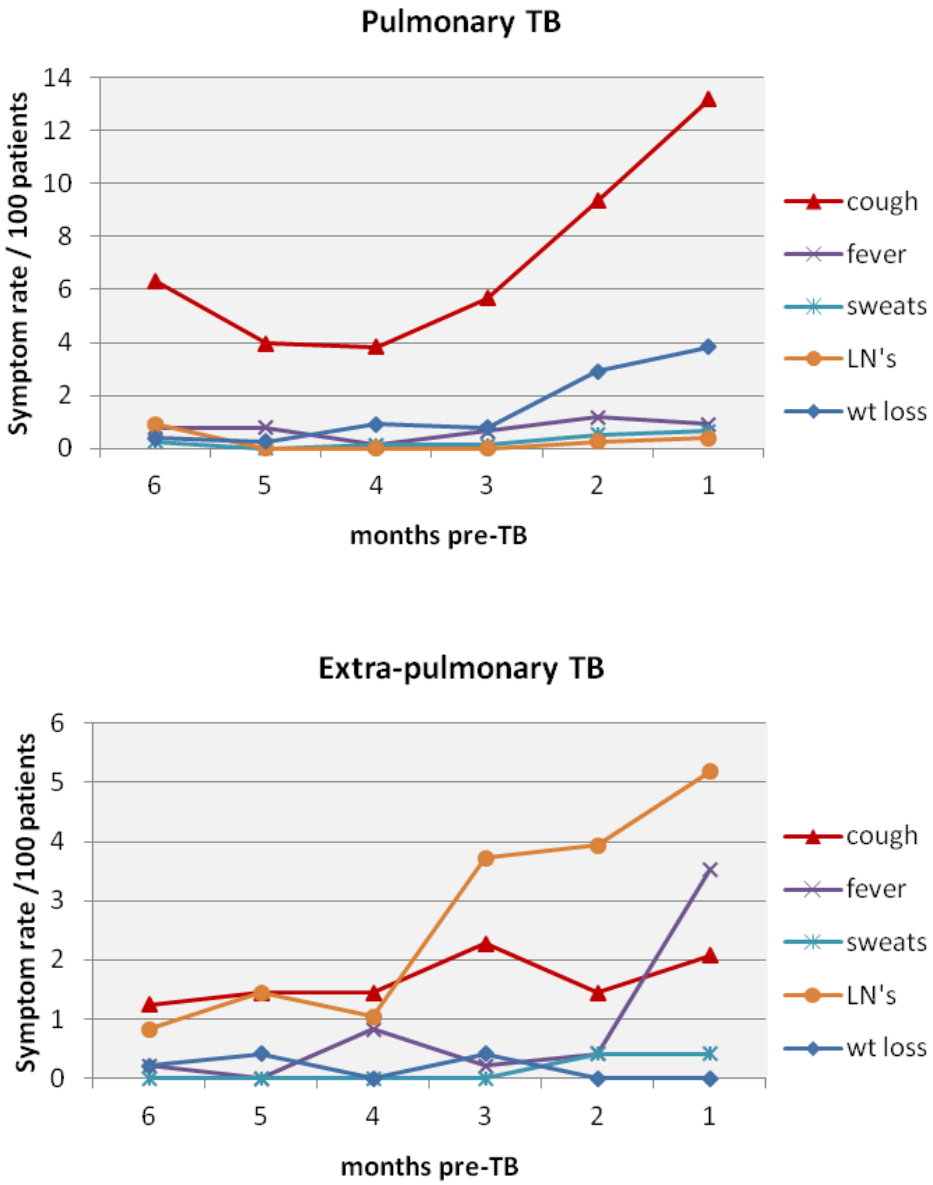
*** Data for consultations on sweating was not subdivided as case numbers were very small*

This study shows that a diagnosis of pulmonary TB was associated with a two fold increase in consultation rate for 'any respiratory disease' in the 6 months pre-TB as compared to a TB-free period (IRR 2.38); and associated with a 3 fold increase in 'respiratory infections and cough' (IRR 2.92) or LRTI consultations (IRR 3.04).

5.4.6.2 Symptom and Disease Consultations by Month and TB Type

Consultations by month in the 6 months pre-TB for respiratory diseases, respiratory infections and symptoms are difficult to interpret due to the small number of patients in each sub-group. For pulmonary TB a rise in consultations for cough and weight loss started about 3 months pre-TB. For extra-pulmonary TB a rise in consultations for lymphadenopathy was seen from 3 months but for fever from only 1 month pre-diagnosis (see Figure 5.7).

Figure 5.7 **Classic TB symptom rates by month pre-TB for 'pulmonary' and 'extra-pulmonary' TB**



5.5 Discussion

5.5.1 Main Findings

TB cases consulted their GP nearly twice as often in the six months before diagnosis as compared to themselves in a TB-free period. Consultation rates for the cardinal symptoms of TB and respiratory infections were also significantly higher; as were repeat consultations. The insidious, non-specific onset of TB is confirmed and a delay in

diagnosis of up to four months. This self-controlled study confirms the findings of the case-control study discussed in Chapter 4, however, the measures of effect are smaller suggesting that control selection bias or confounding by chronic disease, ethnicity or socioeconomic status influenced the results of the case-control study.

Multiple visits were seen equally for different pathological sites of TB but certain symptoms were more common: for pulmonary TB, cough and weight loss were more common and for extra-pulmonary TB, lymphadenopathy was more common.

5.5.2 Self Controlled Methodology

By using a self controlled (case only) study to examine consultation behaviour I overcame unmeasured confounders, such as ethnicity, social class, underlying chronic disease and smoking. Using a self controlled study seemed an appropriate solution when over two thirds of UK TB patients are from an ethnic minority [47] and the GPRD does not contain any ethnicity data. Unmeasured confounders may have influenced the results of my case control study, but by using this methodology I also eliminated control selection bias and allowed a case to be its own control during a disease free period. Self controlled analysis enables an assessment of the additional consulting behaviour associated with TB which cannot be analysed using a case-control study.

The analysis was undertaken using the Generalised Estimating Equation (GEE) model which generates a summary measure, the incidence rate ratio (IRR). The IRR represents statistically the comparison between consultations occurring 12-18 months pre-TB (when the patient is TB-free) and 0-6 months pre-TB (when a patient may be developing TB). The use of 12-18 months as a TB-free time period against which to compare patients becoming symptomatic with TB is validated by Figure 4.6 in Chapter 4 as the increase in case consultation rate is seen from 7 months pre-TB.

This self-controlled study is more of a descriptive study looking at consultation rates, and comparisons of these in different time periods, than a study assessing the risk or incidence of TB. Cases act as their own controls in a TB-free period 12-18 months pre-TB so the study cannot be used to assess the odds of TB as if it were a case-control study.

5.5.3 Consultations

A diagnosis of TB was associated with a 63% increase in consultations, a threefold increase in TB symptoms and a threefold increase in respiratory infections in the six months pre-diagnosis as compared with a TB-free period. Consulting rates for the cardinal symptoms of TB were significantly higher with sweats, lymph nodes and weight loss being particularly elevated. Multiple GP consultations pre-diagnosis could provide an opportunity to diagnose TB earlier but early symptoms are often non-specific. This study highlights 'alert' or 'red-flag' symptoms and diagnoses associated with TB which might help GPs make an earlier diagnosis and supports the need for awareness raising of these among GPs to reduce diagnostic delay. It is important to note that symptoms are only recorded in a minority of cases (27%), and so although useful, it is as essential to improve GPs knowledge of local TB epidemiology, risk factors for TB and the potential significance of repeat respiratory consultations to help diagnose TB earlier. One recommendation that might improve the symptom data would be to undertake an analysis of free-text data to increase our understanding of the significance of their association with TB.

Clearly, an increased consultation rate would be noticed more easily by the same GP but research suggests that continuity of care by the same GP is decreasing.[172] This potentially makes diagnosing a disease such as TB more difficult.

What this self controlled study also shows is that patients are likely to be presenting multiple times to GPs with symptoms potentially compatible with TB but which may be being misdiagnosed as respiratory disease. In the 6 months pre-TB nearly a third of TB patients were diagnosed with a respiratory infection and 11% with three or more lower respiratory infections. Therefore a clearer diagnostic algorithm for GPs and when to 'think to investigate for TB' may be helpful and this work may be useful to inform this.

5.5.4 Stratification of Patients by TB Type

Research evidence suggests that cases of pulmonary and extra-pulmonary TB present differently. I undertook further analysis to understand if different types of TB present to primary care in unique ways that might aid faster diagnosis by GPs.

This research confirms research by others that factors associated with extra-pulmonary TB include female gender[173-175] and younger age.[173-176] I also found that patients with extra-pulmonary TB had less chronic disease than pulmonary cases; this may be explained by their younger age.

For pulmonary TB a 63% increase in consultation rate is seen over the two time periods studied. For extra-pulmonary TB a slightly lower increase in consultation rate at 51% is seen. This may represent the fact that pulmonary cases are more symptomatic and so seek medical care or may represent the fact that pulmonary TB cases have more chronic disease and so see their GP more often for this reason.

Comparison, using the incidence rate ratio, of data from the two time periods 0-6 months and 12-18 months pre-TB revealed that although sweats, lymphadenopathy and weight loss were useful for diagnosing TB in general terms, specific symptoms were more associated with particular types of TB; such as cough and weight loss for pulmonary TB and lymphadenopathy for extra-pulmonary TB (see Appendix 6). However, counts are small once TB-type is sub-divided and confidence limits wide. It is therefore difficult to state how conclusive this evidence is. It is also likely that other symptoms are hidden in the free-text fields that I could not formally examine.

Extra-pulmonary TB is not infectious to others but is, however, more difficult to recognise and diagnose. In the UK the clinical presentation of tuberculosis is changing with an increasing proportion of patients now presenting with extra-pulmonary TB.[173] Awareness of this needs raising in primary care so that clinicians consider extra-pulmonary tuberculosis. This research may help this as it clearly shows that patients with extra-pulmonary TB have more consultations for lymphadenopathy and fever and that there is no significant association with any respiratory disease, infection or cough.

When the GPRD data are stratified by the pathological site of TB (pulmonary or extra-pulmonary) it becomes evident that GPs record diagnoses differently from national notification data (NOIDS returns).[47] This may be because GPs use electronic software that offers drop down diagnostic lists for fast data entry. A large proportion of TB patients (58%) are recorded by their GPs as having 'Tuberculosis;' these patients are not classified further into pulmonary or extra-pulmonary TB. I have termed this 'unclassified' TB' (i.e. site not otherwise specified). The high number of patients with

'unclassified TB' has probably arisen because 'Tuberculosis' appears at the top of the computer generated list of TB codes and pulmonary and extra-pulmonary or site specific TB are lower down. My results show that for TB cases recorded in the GPRD, 25% had 'pulmonary TB' and 16% 'extra-pulmonary TB'. National rates show that in 2007, 56% of TB was pulmonary and 44% extra-pulmonary.[47] This suggests that GPs are probably recording both pulmonary and extra-pulmonary TB within 'unclassified TB' but are managing to record approximately similar proportions of pulmonary and extra-pulmonary TB to national levels if 'unclassified TB' is excluded.

Clinical data from primary care is essential for research of this type. The GPRD provides this in an easily searchable form. But, as data entry is occurring during every-day, busy GP surgeries, primarily as a record of a diagnosis and actions taken, the specificity of recording by GPs is not necessarily a priority. It would be useful to encourage greater diagnostic specificity in practices supplying data to the GPRD to assist research.

The recording by GPs of different types of TB changed little over the time period of this study. In the last few years GP coding has improved markedly,[177] firstly due to practice computerisation and secondly due to financial incentives and the need to supply data to Primary Care Trusts for additional payments and the Quality and Outcomes Framework (QOF - a component of the new General Medical Services contract for general practices, introduced from 1 April 2004). One therefore would presume the recording of TB may have also improved, although the recording of TB is not specifically part of QOF.

Overall, this analysis by TB-type is not as accurate as that available from national enhanced TB surveillance data where TB type is specified for each reported patient.[47] What this research adds is an angle on consultation history and an understanding of consultation rates and type for pulmonary and extra-pulmonary TB plus useful diagnostic indicators.

5.5.5 Limitations of This Study

The limitations of this study are similar to those encountered in the case-control study and are discussed in detail in Chapter 4, Section 4.6.9. These include misclassification by GPs and exclusion of patients who are not registered with a GP. The GPRD lacks

ethnicity data which is likely to confound any study on TB as over two thirds of UK TB patients are from an ethnic minority.[47] In this self-controlled study I have overcome this, however, by using the case as their own control in a time period 18 months prior to diagnosis. Likewise the impacts of socioeconomic status are also overcome by the self controlled methodology.

A further limitation of this study is the low number of symptoms recorded, particularly when analysis of symptoms by TB-type is undertaken. The significance therefore of the symptom results could be questioned. This study, however, uses statistical techniques to compare symptoms in different time periods to try and overcome the difference in absolute numbers. Additionally, when GP recording of 'symptoms' is compared to their recording of 'diagnoses' there is obvious disparity with many more 'specific diagnoses' being recorded than 'symptoms.' This could be explained by the fact that GPs may record a specific diagnosis while asking patients about symptoms that they then record in the free-text field and so analysis of symptoms alone may not be that helpful. To overcome this discrepancy I compared a variable 'respiratory infections including cough' in the two different time periods to bring the symptoms and diagnoses together more appropriately.

5.6 Conclusion

Overall, the results of this self-controlled study confirm those of the case-control study discussed in Chapter 4. They show an increase in consultation rates for symptoms, respiratory infections and in general during the 6-months pre-diagnosis. The measures of effect are smaller in the self-controlled study and this suggests that control selection bias or confounding by chronic disease, ethnicity or socioeconomic status may have had an impact on the results of the case-control study.

Chapter 6

Morbidity Associated with TB: A Self Controlled Study

6.1 Introduction

Long-term morbidity following a diagnosis of TB is little studied; data is lacking because once patients are discharged from hospital outpatients at satisfactory completion of treatment further follow-up is rare. A small number of studies has looked at associations between TB and malignancy[162-164] and TB and chronic respiratory disease.[165,166,178,179] Research, assessing pulmonary function after pulmonary TB, demonstrates variable patterns and severity of pulmonary impairment. Research by Pasipanodya shows that survivors of symptomatic TB have long term lung damage and are 5 times more likely to have abnormal lung function tests than those with latent TB.[166] To examine morbidity following TB I conducted a self controlled study using the GPRD, a longitudinal dataset with up to 20 years of data.

6.2 Aims & Objectives

- To investigate the long-term outcomes (morbidity) of patients with TB
- To assess /examine if patients are more likely to suffer chronic respiratory problems following an episode of pulmonary tuberculosis

6.3 Methods

Study Design

A self controlled (case-case) study[140,142,149,150,151,152,153] based on the 'case-series' method was conducted to investigate consultation behaviour 12-18 months prior to a diagnosis of TB as compared to 12-18 months post-TB.

Setting

I examined data, as previously described in Chapter 4, Section 4.3, from 743 general practices in the UK who provided data to the General Practice Research Database (GPRD) during the time period April 1990 to April 2002.

Data Source

I used the GPRD (see Chapter 4, Section 4.3).

Participants

Patients were selected from the GPRD between 1990 and 2002;

Cases: were all GPRD patients with a medical code for active TB

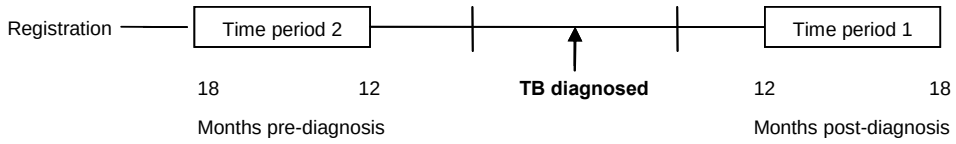
Controls: cases acted as their own control in a different time period i.e. self controls.

'Exclusion' criteria: I excluded patients who had less than 18 months of pre-diagnosis data and patients with less than 18 months of post-diagnosis data.

Measurements

To investigate the long-term outcomes of TB patients I compared consultation behaviour in two time periods: time period 1: 12-18 months post-diagnosis and time period 2: 12-18 months pre-diagnosis. Time period 2 acting as the control period (see diagram below). I looked at total consultation rates, respiratory consultation rates and the prescribing of antibiotics and inhalers during the two time periods of interest. I used the Generalised Estimating Equation (GEE) model[22,24,25,147] (see Chapter 3, Section 3.6.3 for further details) to compare total counts for each individual, and then calculated a summary measure for all individuals, the incidence rate ratio (IRR) with 95% confidence interval.

Figure 6.1 Diagram of comparison time periods for self-controlled morbidity study



Outcomes

The outcome of interest was measures of morbidity due to TB. Morbidity being defined as any departure subjective or objective from a state of physical well-being.[9]

Statistical Analysis

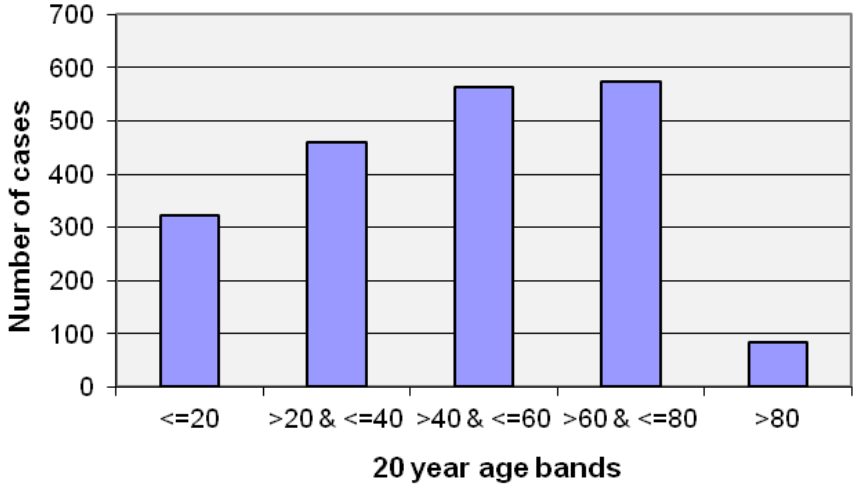
Statistical analysis was performed using STATA version 9.0 (Stata Corp, College Station, TX, USA).[155] I used the Generalised Estimating Equation (GEE) model to generate an incidence rate ratio (IRR) with 95% confidence interval (see Chapter 3 – section 3.6.2 for further details).[22,24,25,147] This method is the same as described for the pre-TB self-controlled study in Chapter 5, Section 5.3.

6.4 Results

6.4.1 Characteristics of Study Participant

2002 TB patients with more than 18 months of follow-up data were identified in the time period 1990-2002. There were similar numbers of males and females (1002 vs. 1000); the mean age at diagnosis was 46 years (SD 22.54; Figure 6.2).

Figure 6.2 Age distribution of TB cases with more than 18months of follow-up



At diagnosis 46% (n=915) of TB patients had one or more chronic diseases (of four analysed); 36% (n=724) had chronic respiratory disease (including asthma, COPD etc),

10% (n=198) chronic heart disease, 7% (n=140) diabetes and 2% (n=32) chronic renal disease. In addition, 17% (n=342) of cases had been, or were current, smokers.

For cases with over 18 months of follow-up data (n=2002): 24% (477) were coded by GPs to have pulmonary TB, 17% (340) had extra-pulmonary TB, 59% (1775) had 'unclassified TB' and 0.5% (10) TB at an unknown site.

6.4.2 Consultation Behaviour

No significant difference was seen in the overall consultation rate 12-18 months pre-TB and 12-18 months post-TB (p=0.328): 6558 consultations occurred 18-12 months pre-diagnosis and 6617 consultations 12-18 months post-diagnosis (IRR 1.03; 95% CI:0.97–1.09).

GP consultations for cough and specific respiratory diseases are shown in Table 6.1. No significant difference was seen in the consultation rates for cough, respiratory disease, respiratory infection and LRTI before and after diagnosis.

Table 6.1 A comparison of GP consultations for patients' 12-18 months pre- and 12-18 months post-TB expressed as an IRR

CONSULTATIONS	**18-12 months pre-TB** No.	**12-18 months post-TB** No.	**Incidence Rate Ratio* IRR (**95% CI)	***p* value**
Cough	218	251	**1.15** (0.93 – 1.43)	0.20
Respiratory disease	882	946	**1.07** (0.93 – 1.23)	0.32
Respiratory infections	594	672	**1.13** (0.98 – 1.31)	0.10
LRTI	295	296	**1.00** (0.82 – 1.23)	0.97

**Statistical analysis using GEE Model*

6.4.3 Analysis of Prescribing Data

Similar numbers of patients were prescribed antibiotics pre- & post-TB: 622 patients pre-TB and 632 patients post-TB. In terms of scripts: 1070 antibiotic prescriptions were issued pre-TB and 1181 post-TB. Antibiotic prescription rates rose from 0.53 pre-TB to 0.59 post-TB but this was not a statistically significant rise (IRR 1.10; 95% CI:0.99-1.22, p=0.06).

Of 2002 patients, 174 patients pre-TB and 242 patients post-TB were prescribed inhalers. A total of 739 inhaler prescriptions were issued pre-TB and 965 post-TB. Inhaler prescription rates rose from 0.37 to 0.48 post-TB, a statistically significant rise (IRR:1.31; 95%CI:1.08-1.58, p=0.006).

As the only significant difference between patients 12-18 months pre-TB and 12-18 months post-TB was the prescription of inhalers, I analysed this group in more detail. 313 people were prescribed inhalers in one or both of the time periods under comparison, see Figure 6.3.

Figure 6.3 Venn Diagram of inhaler prescribing pre and post-TB

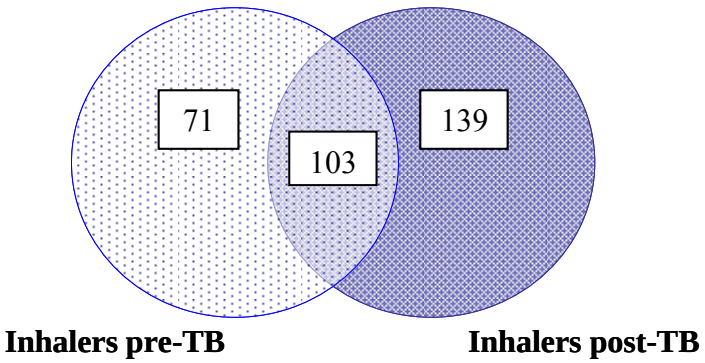


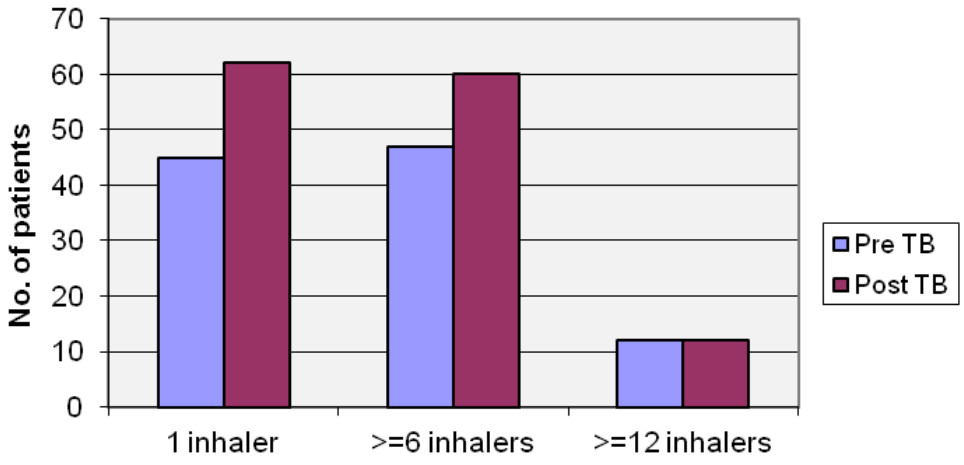
Figure 6.3 shows that twice as many people (n=139) were prescribed inhalers as a new prescription post-TB and that a third (103) remained on inhalers before and after TB. Of the 242 patients prescribed inhalers post-TB 102 (42%) had chronic respiratory disease (asthma or COPD), 63 (26%) had a cough, 82 (34%) had a LRTI and 32 (13%) had shortness of breath.

The mean age of TB cases prescribed an inhaler post-TB was 56 years, i.e. 10 years older than all TB cases.

Figure 6.4 charts the quantity of inhalers prescribed pre and post-TB. It shows higher inhaler prescribing in each of the groups analysed post-TB, apart from in those with 12 or more inhaler prescriptions. In the 6 month time periods of analysis, twelve individuals were prescribed inhalers 12 or more times pre-TB and 12 individuals 12 or more times post-TB. Only one patient was prescribed 12 inhalers pre and post-TB. For patients prescribed 12 or more inhalers post-TB, 7 had no inhaler pre-TB, 3 (a quarter)

had received 6 or more inhalers pre-TB and 2 less than 6 inhalers pre-TB; i.e. different individuals were prescribed multiple inhalers pre and post-TB. I used 12 or more inhalers over a 6 month period as a measure of severe chronic respiratory disease.

Figure 6.4 Inhaler prescriptions for TB patients 12-18 months pre and post-TB



Patients who were prescribed 12 or more inhalers post-TB were on average 10 years older than those prescribed 12 or more pre-TB. Half the patients prescribed 6 or more inhalers pre-TB also received them post-TB whereas only a third of those who received them post-TB had also received them pre-TB.

6.4.4 Analysis of Consultations by TB Type

No significant difference was seen in the overall consultation rate or for specific consultations and diagnoses when the data was re-analysed by TB type. See Appendix 7 – for summary tables of consultation data by TB type.

6.5 Discussion

This study suggests that TB patients who survive at least 18 months after diagnosis have a good long-term prognosis. Morbidity was low and the only significant finding post-TB was an increase in inhaler prescribing by 31% (p=0.006).

Following TB there was no significant increase in morbidity as measured by general consultation rates, consultation rates for LRTI and general respiratory infections, or for cough or antibiotic prescriptions. The hypothesis that, following an episode of pulmonary tuberculosis, patients are more likely to suffer chronic respiratory problems,[165, 166] is not clearly supported by this research. Pasipanodya showed that over 50% of the 107 patients he studied had significantly impaired lung function 5 months after starting TB treatment. However, others have shown less impairment, 18% in a South African study[165] and in a Canadian study "minimal functional impairment" 6 months post diangosis.[178] These studies measured impairment of lung function, if inhaler use is assumed to correlate with worsening lung function then my study gives some support to this research. However, if lung function is severely impaired it would seem likely that consultation rates, cough and other measures of lung function would also increase: here they have not. Of note, is the fact that I assessed patients 12-18 months post TB. Potentially patients with severe lung disease secondary to TB might have died by a year post diagnosis or perhaps the patients themselves showed further improvement in their lung function. In addition, in Pasipanodya's study population, over 50% had been or were current smokers whereas in this study 17% were recorded as having been or were current smokers (although under-ascertainment of smoking status is a problem in the GPRD).

The fact that 36% of TB cases surviving at least 18 months had chronic respiratory disease when diagnosed with TB (cf. 28% of all GPRD TB cases) but no increase in consultation rate post-TB suggests that the chronic respiratory disease was a long standing illness giving an equal number of consultations before and after TB and was not directly affected by TB. An alternative explanation could be that patients have been misclassified pre-diagnosis by GPs as having chronic respiratory disease when in fact they had been exhibiting the early symptoms and signs of TB.

I used antibiotic and inhaler prescriptions as a proxy measure of respiratory disease. Inhalers are generally used for treatment of chronic obstructive airways disease or asthma. To understand the 31% increase in inhaler prescribing post-TB I undertook further analysis. This showed that twice as many patients were prescribed inhalers as a 'new' prescription post-TB and a third remained on inhalers before and after TB. Those that remained on inhalers post-TB may represent patients with a true chronic respiratory disease at diagnosis, and this fits with the 36% of TB patients known to have chronic

respiratory disease at diagnosis. The rise in inhaler prescribing post-TB may therefore represent a true increase in respiratory morbidity post-TB, but this is not confirmed by the consultation data.

There is strong evidence for an association between TB and particular conditions that can lead to increased morbidity, for example: TB and diabetes,[160,161] TB and malignancy,[162-164] and TB and chronic respiratory disease.[165,166] This research shows that 18 months post-TB, however, there is no increase in morbidity due to TB or these conditions as reflected in the stable consultation rates.

This study by using cases as their own controls suggests that patients who survive at least 18 months after a TB diagnosis have morbidity similar to patients without TB. I did consider conducting a cohort study to study this in more detail, but this would have been complicated by unmeasured confounders and the progressive loss of cases either through death or loss to follow-up as patients move GP. In view of the results of this study is seems a cohort study would have added little additional morbidity information.

Strengths and Limitations

To overcome the effects of unmeasured confounders such as ethnicity and socioeconomic status I undertook a self-controlled study. This gives strengthen to the study findings.

By selecting TB patients with at least 18 months of data pre- & post-diagnosis I excluded patients that died in the 18 months immediately after diagnosis, that were lost to follow-up and those that had less than 18 months of GP registration. This latter group would potentially have included newly arrived immigrants. The results of this study are therefore likely to be an under-estimate of long-term morbidity due to these exclusions.

6.6 Conclusion

In TB patients surviving for at least 18 months, consultation patterns returned to levels seen prior to TB disease, suggesting that TB does not result in long-lasting ill health or disability if successfully treated.

Chapter 7

Mortality Associated with TB: A Cohort Study

7.1 Introduction

Mortality following a diagnosis of tuberculosis (TB) is little studied and mortality rates vary considerably.[16,26-33] Analysing TB-patient deaths and risk factors for death, may assist in gaining a better understanding of how to improve the care and the outcomes of patients with TB.

In the UK TB case fatality rates are usually estimated indirectly by comparing the number of TB notifications, or reports to the national surveillance system, with the number of deaths certified as due to TB in any given year. This calculation gives a case fatality rate of around 9%.[180] However, this indirect method of calculating case fatality may give an inaccurate impression and may underestimate mortality. This is because death certificates can be an unreliable source of the actual cause of death and result in misclassification.[181,182] In addition, death from other causes may be increased as a result of having had TB but this is little studied. In the UK a number of small studies[16, 18, 27, 183] have looked at risk factors for death in TB patients but these are still poorly understood.

By studying TB mortality it may be possible to develop more appropriate ways to diagnose and control TB and reduce this potentially preventable cause of death.

7.2 Aims & Objectives

- To describe and quantify mortality from TB
- To calculate the mortality rate in patients with TB
- To examine if TB indirectly increases the risk of death from non-TB causes
- To describe risk factors for death in TB patients

7.3 Methods

Study Design

A cohort study was used to investigate patient mortality following a diagnosis of TB.

Setting and Data Source

I examined data, as previously described in Chapter 4, Section 4.3, from 743 general practices in the UK who provided data to the General Practice Research Database (GPRD) during the time period April 1990 to April 2002.

Participants

Participants in this study were the same as those in the case-control study discussed in Chapter 4; that is all patients diagnosed with active TB from GPRD practices between 1990 and 2002 and a random sample of five patients without-TB for every patient with TB, matched on age band, GP practice and registration at the time of diagnosis of the matched case. Patients without TB were assigned a random index date to equate to their case's date of TB diagnosis. All patients were followed up from diagnosis (or index date) to time of death, transfer out of GP practice or the end of the study period.

Measurements

All patients with a diagnosis of TB were identified using the Read and OXMIS code list (see Appendix 2); as described in Chapter 3, Section 3.4.1.

Mortality was defined as death from any cause during the follow-up period - the period from diagnosis (or predefined index date for participants without TB) to the end of the study period, April 2002. Deaths were defined as TB-related if TB was listed as a specific cause of death (see Chapter 3, Section 3.7.8 for more detail on determining cause of death). Non-TB deaths were defined as those not directly due to TB but due to another cause: I grouped these into deaths due to cardiovascular disease, respiratory disease (excluding TB), cancer, other diseases (including diabetes and renal failure) and unknown. Over 12,000 death codes exist in the GPRD dataset and I re-allocated these to one of these six generic groups to aid analysis of the cause of death.

Patient deaths were identified from the GPRD "patient records" file. All death dates and patient identifiers (ids) were extracted and all consultation records pre- and post- death

then retrieved from the associated "medical records" file. Survival after diagnosis was determined from GP consultations recorded post-diagnosis in the GPRD "medical records" file.

Outcomes

The primary outcomes of interest were death and the time from TB-diagnosis to death. This was calculated by subtracting the date of death from the date on which TB was diagnosed. In the survival analysis, patients were censored at death, or date of last GP visit if they were lost to follow-up, or at the date of transfer to another GP practice. The secondary outcomes of interest were cause of death, possible risk factors for death and frequency of death among patients with TB and those without.

Bias

I tried to reduce ethnicity and social status bias by matching patients on GP practice.

Statistical Analysis

Single variable analysis was undertaken. Categorical variables were compared using the Chi Squared test. A Cox proportional hazards model was used to compare the survival of patients with TB to those without.[154] This model was used to help account for patients who did not provide complete data throughout the follow-up period (censored observations) and was adjusted for sex and chronic disease. Hazard ratios and their 95% confidence intervals were calculated and a Cox proportional hazards survival curve plotted. All p values <0.05 were considered statistically significant and statistical analysis was performed using STATA version 9.0 (Stata Corp, College Station, TX, USA).

I compared demographic characteristics of TB patients and patients without TB that died during the study period, I compared demographic characteristics of TB patients that died and TB patients that survived and analysed GPRD consultation data to assess risk factors for death and cause of death. Further analysis was undertaken of the relationship between TB and cancer and of those patients that died prior to, or very shortly after, a diagnosis of TB. For the latter group, analysis was conducted to determine if these patients were in some way different to the TB cases that survived.

7.4 Results

7.4.1 Characteristics of Study Participants

Over eighteen thousand individuals (3032 with TB and 15,160 without TB) met the inclusion criteria in the time period 1990-2002. Patient characteristics are shown in Table 7.1. Thirty-five percent of TB patients had one or more chronic diseases compared to 18% of individuals without TB and 15% of TB patients had smoked compared to 11% of individuals without TB.

Table 7.1 Characteristics of study participants

Characteristic	Individuals with TB (n=3,032) No. (%)	Individuals without TB (n=15,160) No. (%)
Sex		
Female	1490 (49%)	7833 (52%)
Male	1542 (51%)	7327 (48%)
Age		
<=20	427 (14%)	2135 (14%)
>20 - <=40	651 (22%)	3255 (22%)
>40 - <=60	803 (26%)	4015 (26%)
>60 - <=80	929 (31%)	4645 (31%)
>80	222 (7%)	1110 (7%)
Chronic disease		
Any	1061 (35%)	2729 (18%)
Respiratory	861 (28%)	1837 (12%)
Renal	31 (1%)	42 (0.3%)
Cardiac	213 (7%)	708 (5%)
Diabetes	162 (5%)	475 (3%)

7.4.2 Patient Contributions to the Dataset

The mean follow-up time for patients with TB was 38 months and the mean follow-up time for patients without TB was 35 months. There were 9,591 person-years of follow-up for TB cases and 43,953 person-years of follow-up for patients without TB. Fourteen percent (n=437) of TB patients left their GP during the study period; the outcome of these patients was therefore unknown and they are censored at the last known GP visit for the purposes of the Cox proportional hazard calculation.

7.4.3 Analysis of Overall Mortality Data and Survival

Over the follow-up period 17.3% (n=525) of TB patients died and 8.9% (n=1,352) of individuals without TB died (Table 7.2). The overall mortality rate in TB patients was 54.7/1000 person-years at risk (PYAR), compared to 30.8/1000 PYAR in patients without TB. This gives an incidence rate ratio for death of 1.78 (95% CI 1.61-1.97). The TB-specific death rate was 11.5/1000 PYAR.

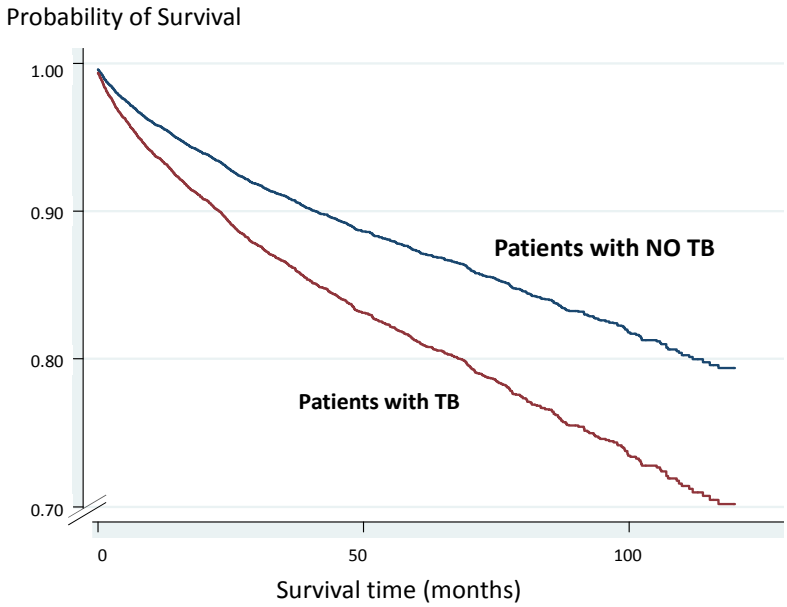
For TB patients, just over half of all deaths, 54% (n=284/525) occurred within a year of diagnosis. Proportionally, three times as many deaths occurred in cases in the first year post-TB as compared to individuals without TB. Further analysis of the data showed that 87 patients (representing 79% of all deaths from TB-disease) died prior to, or very shortly after being diagnosed with TB. This group is further analysed in section 7.4.9.

Table 7.2 Crude comparison of deaths for individuals with TB and those without

	Individuals with TB (n=3,032)	**Individuals without TB** (n=15,160)
Total deaths	17.3% (525)	8.9% (1,352)
Deaths within 1yr of diagnosis	9.4% (284)	3.6% (540)
Deaths >1 yr from diagnosis	7.9% (241)	5.4% (812)

The probability of survival after a diagnosis of TB, estimated using the Cox proportional hazards regression method, is shown in Figure 7.1. There is a marked difference in survival for patients with TB as compared to those without; a sustained and increasing difference is shown across the study period. Survival rates were lower at 1, 5 and 10 years for TB patients compared to patients without TB (see Figure 7.1). The death rate for TB patients was highest in the year following diagnosis. The all-cause case fatality rate of TB patients within a year of diagnosis was 6.5%, mortality at 5 years after diagnosis was nearly 20% and at 10 years 30%.

Figure 7.1 Cox Model Survival Curve for patients with and without TB over a 10 year follow-up period adjusted for sex and chronic disease



	1yr	5yrs	8yrs	10yrs	
% survival	93.5%	81.2%	74.5%	70.1%	**- for TB patients**
No. at risk	2268	726	184	23	
% survival	95.7%	87.2%	82.4%	79.2%	**- for non-TB patients**
No. at risk	10291	3229	1038	259	

7.4.4 Analysis Using a Cox Proportional Hazards Model

The Cox proportional hazards model showed that TB patients had a higher probability of death compared to patients without TB (adjusted Hazard Ratio (HR):1.53; 95%CI:1.39-1.69) even after adjusting for other known risk factors for death; male gender (adjusted HR:1.18) and chronic disease (adjusted HR:2.46) (see Table 7.3). The risk of death was apparently unaffected by smoking (HR:1.00), although from research presented earlier in this thesis I know smoking is under-ascertained in the GPRD.

Table 7.3 Unadjusted and adjusted hazard ratios for death in TB patients versus non-TB patients using a Cox Proportional Hazards Model

Variable	Unadjusted Hazard Ratio (95% CI)	Adjusted Hazard Ratio (95% CI)	*p* value
Patient with TB *(vs not TB)*	1.81 (1.65-1.99)	**1.53*** (1.39-1.69)	< 0.001
Male (*vs* female)	1.20 (1.10-1.31)	1.18[+] (1.07 – 1.30)	0.001
Any chronic disease (*vs* no chronic disease)	2.64 (2.39-2.90)	2.46[+] (2.23 – 2.72)	< 0.001

** HR adjusted for sex and chronic disease*

[+] *HR adjusted for other variables in the model*

In Chapter 4, Section 4.5.8, I showed that chronic respiratory disease was the most significant chronic disease of four common chronic diseases analysed. I therefore repeated the Cox proportional hazards model exchanging the variable 'any chronic disease' for 'chronic respiratory disease' (see Table 7.4). This led to an increase in the adjusted hazard ratio for death for patients with TB as compared to those without to adjusted HR:1.62 (95% CI:1.47–1.78).

Table 7.4 Unadjusted and adjusted hazard ratios for death in TB patients versus non-TB patients using a Cox Proportional Hazards Model, including chronic respiratory disease

Variable	Unadjusted Hazard Ratio (95% CI)	Adjusted Hazard Ratio (95% CI)	*p* value
Patient with TB *(vs not TB)*	1.81 (1.65-1.99)	**1.62*** (1.47–1.78)	< 0.001
Male (*vs* female)	1.20 (1.10-1.31)	1.19[+] (1.09 – 1.31)	< 0.001
Chronic respiratory disease (*vs* no chronic resp. disease)	2.07 (1.87-2.30)	1.89[+] (1.70– 2.10)	< 0.001

** HR adjusted for sex and chronic disease*

[+] *HR adjusted for other variables in the model*

7.4.5 A Comparison of TB Patients and Patients without TB Who Died during the Study Period

Table 7.5 shows a comparison of TB patients who died as compared to patients without TB who died during the study period. A significant difference was seen in gender, age and chronic disease profiles. TB cases that died were more likely to have chronic respiratory disease, and a significantly greater proportion of deaths occurred in TB cases under 60 year olds.

Table 7.5 Single variable analysis of risk factors for death in patients with TB & those without

Characteristic	**Patients with TB who died** (n=525) No. & % of those dying	**Patients without TB who died** (n=1352) No. & % of those dying	***p* value**
Sex			
Male	326 (62.1%)	670 (49.6%)	<0.001
Female	199 (37.9%)	682 (50.4%)	
Age (years)			
<20	2 (0.4%)	3 (0.2%)	<0.001
>20- 40	14 (2.7%)	12 (0.9%)	
>40-60	82 (15.6%)	124 (9.2%)	
>60-80	311 (59.2%)	811 (60.0%)	
>80	116 (22.1%)	402 (29.7%)	
Chronic disease			
Any	289 (55.0%)	457 (33.8%)	<0.001
Respiratory	238 (45.3%)	249 (18.4%)	<0.001
Renal	14 (2.7%)	21 (1.6%)	0.11
Cardiac	73 (13.9%)	176 (13.0%)	0.61
Diabetes	38 (7.2%)	103 (7.6%)	0.78

7.4.6 A Comparison of TB Patients that Died and TB Patients that Survived

Table 7.6 shows that for TB patients that died as compared to those that survived, significantly more were male; they were more likely to have a chronic disease, particularly chronic respiratory disease and have pulmonary TB. Consultation rates in the 6 months pre-TB were also compared and showed that TB cases that died consulted nearly 40% more often than those that survived, a crude rate ratio of 1.37.

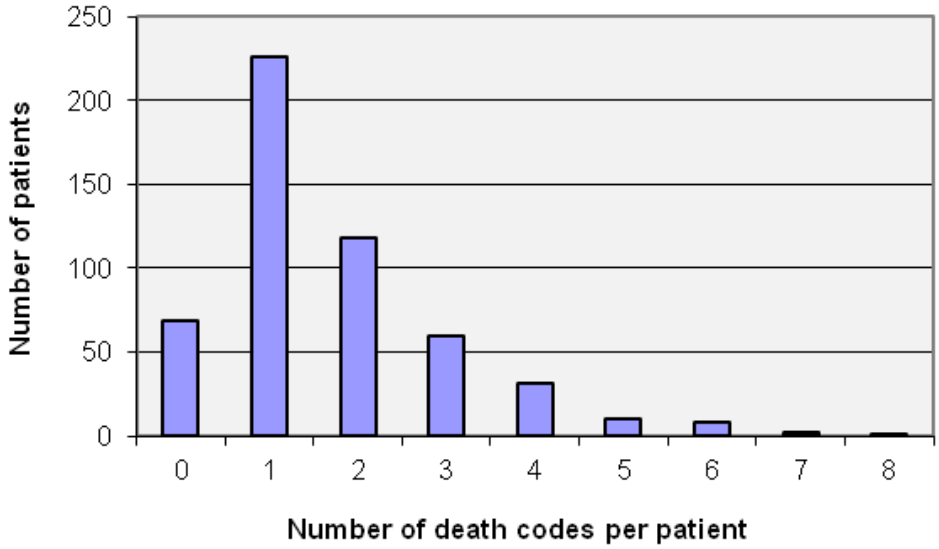
Table 7.6 Single variable analysis of TB patients that died and those that survived

Characteristic	**TB cases that died** (n=525) No. & % of those dying	**TB cases that survived** (n=2070) No. & % of those surviving	***p* value**
Consultation rate in 6mths pre-diagnosis	7.78	5.69	<0.001
Sex			
Male	326 (62.1%)	1012 (48.9%)	<0.001
Female	199 (37.9%)	1058 (51.1%)	
Age (years)			
<20	2 (0.4%)	346 (16.7%)	<0.001
>20- 40	14 (2.7%)	501 (24.2%)	
>40-60	82 (15.6%)	628 (30.3%)	
>60-80	311 (59.2%)	513 (24.8%)	
>80	116 (22.1%)	82 (4.0%)	
Chronic disease			
Any	289 (55.0%)	664 (32.1%)	<0.001
Respiratory	238 (45.3%)	528 (25.5%)	<0.001
Renal	14 (2.7%)	14 (0.7%)	<0.001
Cardiac	73 (13.9%)	122 (5.9%)	<0.001
Diabetes	38 (7.2%)	101 (4.9%)	0.032
TB Type			
Pulmonary	200 (38.1%)	451 (21.8%)	<0.001
Extra-pulmonary	57 (10.9%)	360 (17.4%)	
Unclassified TB	268 (51.0%)	1,259 (60.3%)	

7.4.7 Analysis of Cause of Death in TB Patients

The median number of diagnostic codes recorded in the GPRD on the date of death (death codes) was one (range 0-8) (see Figure 7.2). For the minority of cases with more than one code (n=230) the codes were reviewed to assess the most likely cause of death; see Chapter 3, Section 3.7.8 for methodology.

Figure 7.2 Number of death codes per TB patient



Of the TB patients that died: 21% were recorded as dying of TB, 21% of cancer, 16% of cardiovascular disease, 14% of respiratory (non-TB) disease and 5% of other causes. For 23% of deaths the cause was unknown. Of the 21% (110) deaths specifically attributable to TB the majority, 58 (53%), were in patients with pulmonary TB, 18 (16%) were in patients with extra-pulmonary TB and 34 (31%) were in patients with 'unclassified' TB.

There was a significant difference between the cause of death for TB patients dying within 18 months of diagnosis and those dying after 18 months (see Table 7.7). Nearly all the deaths recorded as being due to TB occurred in the 18 months following diagnosis. A comparison of deaths before and after 18 months was chosen as, by 18 months post-diagnosis, all patients should be well and TB-free.

Table 7.7 Cause of death in TB patients

Cause of death	Total No. (%) of TB patients dying (n=525)	TB patients dying 0-18 months post diagnosis No. (%) (n=318)	TB patients dying >18 months post diagnosis No. (%) (n=207)
TB	110 (21%)	109 (34%)	1 (0.5%)
Unknown	118 (23%)	62 (20%)	56 (27%)
Cancer	110 (21%)	70 (22%)	40 (19%)
Cardiovascular disease	86 (16%)	33 (10%)	53 (26%)
Respiratory disease (not TB)	74 (14%)	33 (10%)	41 (20%)
Other	27 (5%)	11 (4%)	16 (8%)

In the 18 months post diagnosis, one third of deaths (n=109; 34%) were attributed specifically to TB, with most of these occurring prior to or soon after diagnosis (80%; n=87/109). In addition, considerably fewer TB patients died of cardiovascular and respiratory disease during this time period.

The majority of deaths that occurred more than 18 months after a diagnosis of TB were due to cardiovascular disease, respiratory disease and cancer, where the cause of death was known. Only one TB patient died from TB disease during this time period.

A comparison of the mortality rates by cause of death for TB and non-TB patients (Table 7.8) showed that the rate of death due to cancer in patients with underlying TB was nearly double that of patients without TB (IRR:1.89), as was the rate of death from respiratory disease other than TB (IRR:1.87); but the rate of death from cardiovascular disease was unaffected (RR:0.94).

Table 7.8 Mortality rate and incidence rate ratio by cause of death for patients with TB and those without

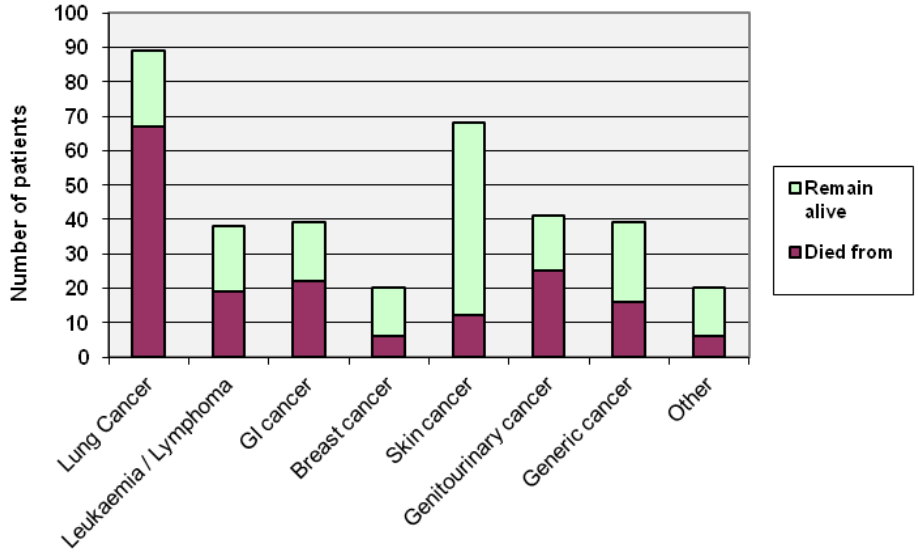
Cause of death	TB patients (n=3,032)		Non-TB patients (n=15,160)			
	No. (%) dying	Death rate /1000 PYAR	No. (%) dying	Death rate /1000 PYAR	Incidence Rate Ratio (95% CI)	*p* value
TB	110 (21%)	11.5	-	-		
Unknown	118 (23%)	12.3	412 (31%)	9.4	**1.31** (1.06-1.61)	<0.001
Cancer	110 (21%)	11.5	267 (20%)	6.1	**1.89** (1.50-2.37)	<0.001
Cardiovascular disease	86 (16%)	9.0	420 (31%)	9.6	**0.94** (0.74-1.19)	0.3
Resp. disease (not TB)	74 (14%)	7.7	181 (13%)	4.1	**1.87** (1.41-2.47)	<0.001
Other	27 (5%)	2.8	72 (5%)	1.6	**1.72** (1.06-2.71)	<0.001
Total deaths	**525**	54.7	**1352**	30.8	**1.78** (1.61-1.97)	<0.001

7.4.8 Cancer as a Cause of Death

Cancer was the second commonest cause of death in TB patients where cause of death was known. In patients with underlying TB, the mortality rate from cancer was 11.5/1000 PYAR. Of 3032 TB cases, 347 (11.4%) had a diagnosis of cancer. 173 patients (49.9%) were diagnosed with cancer pre-TB and of these 16% died; of the 174 patients diagnosed with cancer after their TB, 46% died. This suggests that TB patients who were diagnosed with cancer post-TB had a much worse prognosis than those diagnosed pre-TB. Nearly a third of all TB cases with cancer died from their cancer (n=110) during the follow-up period; two thirds within 18 months of their TB diagnosis and a third after 18 months.

The commonest form of cancer seen in TB patients (Figure 7.3) was lung cancer (25%) followed by skin cancer (19%). Of the TB patients with lung cancer, (a cancer with known poor prognosis) 76% died during the period of analysis as compared to 40% of those with other malignancies.

Figure 7.3 Types of cancer seen in GPRD TB patients



7.4.9 Death Prior to, or Very Shortly after, a Diagnosis of TB

Eighty seven patients were recorded in the GPRD as "dying on the day their TB was diagnosed" (i.e. the date of recorded patient death was the same as the date of TB diagnosis). These 87 patients represent 79% of all deaths from TB-disease and a third (31%) of all TB patient deaths occurring within a year of diagnosis.

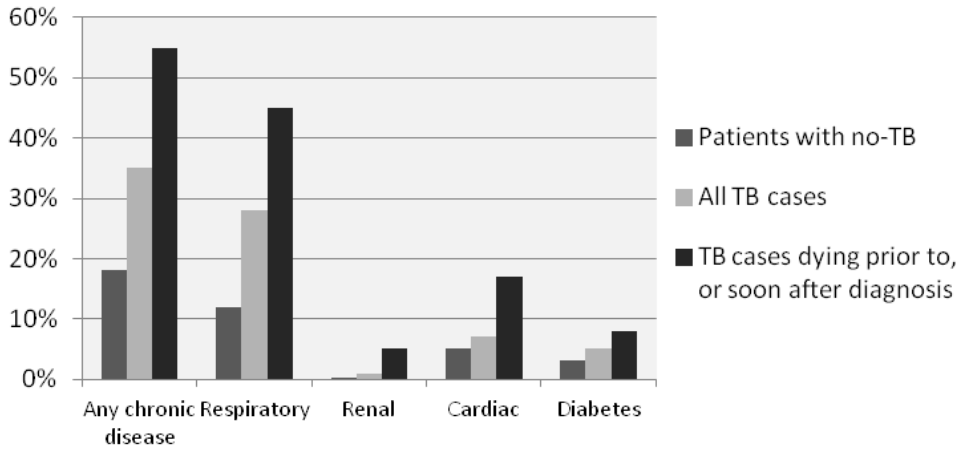
Demographic and other characteristics of the 87 TB cases that died prior to, or very shortly after, a diagnosis of TB are shown in Table 7.9. Those who died were more likely to have pulmonary TB, be male, be aged over 60 and have one or more chronic diseases with chronic respiratory disease being the most common.

Table 7.9 Characteristics of TB patients who died prior to, or shortly after, a diagnosis of TB

Characteristics	TB Cases No. & (%) (n=87)	All other TB Cases No. & (%) (n=2,945)
Sex		
Female	39 (45%)	1451 (49%)
Male	48 (55%)	1494 (51%)
Age		
<=20	0 (0%)	427 (14%)
>20 - 40	1 (1%)	650 (22%)
>40 - 60	7 (8%)	796 (27%)
>60 - 80	51 (59%)	878 (30%)
>80	28 (32%)	194 (7%)
Chronic Disease		
Any	48 (55%)	1013 (34%)
Respiratory	39 (45%)	822 (28%)
Renal	4 (5%)	27 (1%)
Cardiac	15 (17%)	198 (7%)
Diabetes	7 (8%)	155 (5%)
Ever smoked	13 (15%)	426 (14%)
TB Type		
Pulmonary	44 (51%)	715 (24%)
Extra-pulmonary	16 (18%)	467 (16%)
Unclassified TB	27 (31%)	1747 (59%)

Figure 7.4 shows that the proportion of cases with any chronic disease is highest in patients that died prior to, or shortly after, a diagnosis of TB and lowest in those without TB.

Figure 7.4 Chronic diseases in different TB patient populations



In the 6 months pre-diagnosis, the median number of consultations for cases diagnosed with TB at or near death was six (IQR=4 to 11; range 0-33); this was very similar to that for all-TB cases (n=3032) at five. Only 8% of these 87 cases were not seen pre-diagnosis. Symptoms compatible with TB were reported by 18 of 87 patients, these patients were evenly distributed between pulmonary, extra-pulmonary and 'unclassified' TB; the majority had cough, followed by fever and weight loss. Of the 87 TB patients that died at, or near diagnosis, 35% had a chest infection in the 6 months prior to death and 14% had consultations for chronic respiratory disease.

Over half of the patients diagnosed with TB at or near death, had pulmonary TB, as compared to only 25% of all TB cases. Analysis of the number of consultations per case in the 6 months prior to a diagnosis of TB showed that patients with pulmonary TB had a median of 7 consultations and those with extra-pulmonary had a median of 9 consultations.

7.5 Discussion

7.5.1 Main Findings

This is the most detailed review of TB patient deaths in the UK to date. The all-cause case fatality rate of TB patients within a year of diagnosis was 6.5%, mortality 5 years after diagnosis was nearly 20% and at 10 years 30%. The majority of TB patient deaths occurred in the first year after diagnosis, nearly a third of these occurred prior to or

shortly after diagnosis. Death from TB-disease specifically, occurred almost exclusively within the first 18 months, with 80% occurring prior to, or very shortly after, a diagnosis of TB suggesting a missed diagnosis, serious delay in diagnosis and/or advanced disease. Excess mortality in TB patients was greatest in the first year after diagnosis, thereafter there was an excess but it was less marked. Patients with TB had a greatly increased risk of death if they had pulmonary TB, were male, aged over 60 and had a chronic disease or cancer. TB was associated with a higher probability of death (adjusted HR:1.53) and the TB-specific death rate in this study was 11.5/1000 PYAR.

7.5.2 Case Fatality Rates and Survival Rates

There was a marked difference in survival of patients with TB as compared to patients without TB. A TB patients chance of survival was 81% at 5 years and 70% at 10 years, survival for those without TB was 87% and 79% respectively; TB patients had an elevated mortality following diagnosis and this risk remained even after treatment completion. This finding is supported by that of Tocque et al.[27]

In this study the 12-month all-cause case fatality rate was 6.5%. Strikingly different TB case fatality rates have been found in various studies in different countries. Many of these studies assess mortality by looking at deaths within a year of diagnosis or during treatment.[16,26, 28, 30, 32, 184] These studies show case fatality rates that vary between 3.6%[16] and 24%.[32] In 2006 the global TB case fatality rate for smear-positive pulmonary TB patients was 4.6%.[185] One of the strengths of my study is that in addition to calculating the case fatality rate at one year I have been able to assess the long-term outcomes of TB patients by using up to 12 years of longitudinal data contained within the GPRD.

This study adds to the existing UK literature and particularly the studies of Crofts et al.[180] Ditah[186] and Martineau[187] who each reviewed national UK TB surveillance and mortality data. It adds both in terms of its size, its length of longitudinal follow-up data and the fact that it uses quality controlled primary care data from the GPRD, rather than national surveillance data, and links this to national mortality data. This study differs from previous studies in that it is the first to use a proper primary method. Previous estimates of case-fatality have been based on triangulation of data; here I have used a cohort study with robust methods. The 12 month case fatality rate in this study (6.5%) is between that found by Crofts[180] and Martineau;[187] Martineau describes a decrease in

case fatality rate from 9.26% in 1998, to 5.59% in 2001; Crofts describes a case fatality rate of 8.9% in 2001/2. A lower case fatality rate may represent better TB diagnosis and control, may represent younger patients with less co-morbidity or may represent incomplete ascertainment. In Crofts conclusion he comments that mortality among TB cases is underestimated in national TB surveillance data (in fact by 25%) but by triangulating three different data sources a higher estimate of 8.9% was found. Overall therefore, it seems that the case fatality rate of 6.5% found in my study is similar to the case fatality rate calculated from other single national UK data sources.

This study supports the findings of others[16, 26, 27] and shows that TB increases the risk of all-cause mortality compared to the general population. Tocque et al. found 24% all-cause mortality in 8 years, with 31% of cases dying from TB;[27] Bakhshi found 3.6% all-cause mortality in 6 years.[16] My study, which is considerably larger than these, showed a 20% all-cause mortality at 5 years and 30% at 10 years with 1 in 5 deaths attributable to TB disease. Most TB patients die of something other than TB and have a higher death rate even after successful treatment.

In this study half of all TB patient deaths occurred in the first year after diagnosis. This research shows that TB patients who died, consulted more often in the 6 months pre-TB than those that survived (RR:1.37). If excess consultations are taken to be a proxy measure of delay in diagnosis, then diagnostic delay appears to have contributed to death.

It is important for national TB programmes to understand TB patient deaths so that they can assess how effective their TB management and control has been. This is particularly true in high income countries where death is often the most commonly reported reason for TB treatment failure. In the UK, death whilst on treatment is the main reason that we fail to meet the WHO's global target of 85% treatment success.[1, 186, 188]

7.5.3 Risk Factors for Death

This study confirms, and adds to what other smaller studies have found, that specific risk factors are associated with increased mortality in TB patients. These risk factors include: male gender,[16,27,30,31,33,94, 184] older age,[16, 26, 30-32, 183, 184, 189] pulmonary disease,[30, 184,190] presence of malignancy or HIV[27,30,31,94] and having an underlying medical

condition[32] (particularly chronic respiratory disease.[16]) Sterling also showed that the more independent risk factors a patient has, the greater their risk of death.[31]

A few studies have looked at social determinants of ill health and found associations between alcohol or drugs and TB.[26,30,31] This was not possible in my study, nor was analysis of the impact of ethnicity on death. Results differ in studies that assess the impact of ethnicity on death; Pablos-Mendez found an association between TB death and ethnic minority group,[94] where as Bakhshi and Sterling did not.[16,31]

Another risk factor for death is treatment delay. Very few studies have looked at treatment delay and its impact on death. Pablos-Mendez in New York, however, found higher death rates in those who were delayed in starting treatment.[94] Although I cannot directly show that repeat consultations prior to diagnosis constitute treatment delay, it is likely that this is the case, as TB patients consulted their GP nearly three times more often (OR:2.68), and with symptoms and respiratory conditions compatible with TB in the 6 months pre-diagnosis (see Chapter 4, Section 4.5.3). Moreover, a comparison of TB cases that died, with TB cases that survived, revealed that those that died consulted nearly 40% more often in the 6 months pre-TB than those that survived. This suggests a delay in diagnosis that impacted on survival. Repeat consultation behaviour pre-diagnosis has been little studied and this information is one of the strengths of using the GPRD dataset which contains such data. Diagnostic delay is likely to be one of the main contributors leading to death from TB, particularly for those who die prior to, or soon after, diagnosis. Diagnostic delay, whether patient or health care worker delay, is a factor that can be changed and reduced by the provision of TB information and education to patients and health care workers.

This study gives a unique insight into the risk factors for death from TB in a large sample population. Single variable and multivariable analysis of TB patients that died and those who survived helps us to understand the possible risk factors that are associated with death in TB patients. Knowledge of these factors in conjunction with an increase in consultation rates and repeat chest infections could help health care professionals diagnose TB earlier, and lead to quicker treatment and decreased mortality. Specifically in the UK, consideration of a diagnosis of TB in older, male patients with a known chronic disease (particularly chronic respiratory disease) and cancer may help to reduce TB related deaths.

7.5.4 Co-Morbidity and TB Deaths

The analyses presented in this thesis show that chronic disease is associated with an increased risk of developing TB, diagnostic delay and an increased likelihood of death. This supports the work of others.[32,123] The importance of co-morbidity is shown here by a 2.5 times greater risk of death in TB patients with a chronic disease (adjusted HR:2.46). One third of TB cases analysed had a chronic disease at diagnosis compared to 18% of patients without TB and 55% of those dying prior to, or very shortly after diagnosis. This suggests that TB cases are generally sicker (as measured by rates of chronic disease) than cases without TB. This may be one explanation as to why patients are more susceptible to contracting TB and why TB cases have a higher death rate than cases without TB. Another explanation might be that TB symptoms have been wrongly assigned to a chronic respiratory disease and so delayed the diagnosis of TB which can present in a similar way.

Of all the chronic diseases analysed, chronic respiratory disease had the greatest impact on the risk of death for TB patients'. Survival was 62% worse for these patients; a fact that needs to be considered and acted upon by GPs when treating patients with TB.

7.5.5 Cause of Death

A detailed analysis of cause of death was not possible, as the GPRD provides secondary data and there is no direct access to the actual records or patient letters. Over a fifth of deaths were directly attributable to TB, the majority occurred early and were due to pulmonary TB. The latter finding is interesting as theoretically these patients would have been more likely to have symptoms compatible with an early diagnosis. Cause of death differed significantly ($p<0.001$) depending on how soon it occurred after diagnosis; for those dying early (within 18 months of diagnosis), most died from TB disease or cancer, for those dying later most died of cardiovascular disease, respiratory disease or cancer. In addition, this study showed that the increased risk of death for TB patients continued long-term.

This study suggests that the single most important factor causing death from 'TB disease' was the failure to diagnose (and therefore treat) patients with TB earlier.

7.5.6 Cancer and TB

Death from cancer is associated with TB. This is confirmed by this and other studies.[27,30,31] Here, cancer doubled the rate of death in TB patients as compared to patients without TB. In this study 10% of TB cases had a diagnosis of cancer; half were diagnosed with cancer prior to their TB and half after. This research suggests that TB patients diagnosed with cancer post-TB have a much worse prognosis; nearly three times as many TB patients died if their cancer was diagnosed after their TB than before. Tocque et al. found that malignancy was the cause of death in 28% of TB patients that died, and that lung cancer was particularly common.[27] Similarly my study shows the importance of malignancy as a cause of death (seen in 21%) and that lung cancer was the most common form of cancer (seen in 25%). Research suggests that cancer and its treatment lead to immune-suppression and this in turn can lead to reactivation of TB in those already infected or increase susceptibility to new infection. For those cases diagnosed with TB prior to their cancer, TB might be acting as a marker of underlying immune-suppression associated with cancer.

7.5.7 Death Prior to, or Shortly after Diagnosis

Seventeen percent of TB cases that died are recorded to have died prior to, or shortly after, diagnosis. From GPRD data it is not possible to distinguish whether the diagnosis of TB was made clinically or at post-mortem. The description "death prior to, or very shortly after a diagnosis of TB" is probably an artefact of how GPs record patient death and diagnosis in their records. It could represent those patients who died prior to a formal diagnosis of TB and then were subsequently diagnosed at post-mortem or following a positive microbiological culture; or it could represent patients who were diagnosed late and died while in hospital shortly after diagnosis. In the latter group GPs are likely to be informed of the patient's diagnosis and death at the same time and so record these in the patient's notes as having occurred on the same date.

Croft et al.[180] made a similar finding that 25% of TB patients were diagnosed post-death or on the day they died. Any factors that differentiate this group from TB cases that survive could be used to help healthcare workers investigate and diagnose TB earlier and so reduce this mortality. Those who died at diagnosis were more likely to have pulmonary TB, be male, aged over 60 and have one or more chronic diseases; chronic respiratory disease being the most common. Risk factors to help diagnose TB prior to

death appear to be no different to those associated with increased mortality in all TB patients, whether diagnosed pre or post death.

80% of all deaths directly attributable to TB-disease occur prior to, or shortly after, diagnosis and the majority of these patients are seen multiple times with a median of 6 consults in the 6 months pre-diagnosis. These patients are therefore likely to have had advanced undiagnosed TB suggesting a serious delay in diagnosis.

Diagnosis at death suggests an unacceptable and potentially avoidable delay in diagnosis and treatment to such a degree that it has resulted in death. Ways to reduce this delay need to be further investigated but would probably need to take a two-pronged approach: firstly to educate patients of the signs and symptoms of TB so that they present earlier to a health care professional, and secondly to educate health care professionals in primary care and accident and emergency departments to "THINK TB".

7.5.8 Study Limitations

This study has several limitations:

1. Determining the cause of death by reviewing GP data records is subject to possible misclassification. Some patients had multiple entries on the day of death. I estimated a single cause of death by reviewing all multiple entries on the day of death and allocating one using a pre-determined method (see Chapter 3, Section 3.7.8). By following a strict protocol I tried to reduce misclassification.
2. The mean patient follow-up period was 38 months, although up to 10 years of follow-up time was available for some. I analysed deaths within the time period available for each patient; survival data is therefore not based on a life-time but rather the time available in the dataset. The Cox proportional hazards model allows for this.

3. The data available in the GPRD did not allow me to assess some important known risk factors for TB such as ethnicity and social class. To lessen the effect of these possible confounders I randomly selected patients without TB from the same GP practice as the case, making the assumption that patients registered with the same GP are likely to be from similar ethnic and social groups.
4. Fourteen percent (n=437) of TB patients were 'transferred out,' i.e. left the GP practice in which they were first diagnosed during the study period; the outcome of these patients is therefore unknown. Their exclusion from the final analysis may add a degree of bias to the results.
5. Smoking was poorly recorded in the GPRD dataset used (only 15% of patients were recorded to have ever-smoked, where as the current percentage of active smokers in UK society is more like 25%). This under-ascertainment has been noted in other studies and is equally poor in both the exposed and the unexposed groups. Smoking may therefore be inadequately controlled for in the multivariable analysis. The risk of death in the Cox proportional hazards model was unaffected by smoking (hazard ratio of 1.00); it is important however for smoking to remain in the model as it is a risk factor for chronic disease and TB[169, 170] and therefore a potential confounder.

7.6 Conclusion

TB leads to increased mortality and a case fatality rate at one year of 6.5%. Mortality is particularly high in the first year following diagnosis. Undiagnosed TB, pulmonary TB, older age, male gender, malignancy and chronic disease are all associated with a worse prognosis. A TB patients chance of survival was 81% at 5 years and 70% at 10 years.

Death from TB-disease occurred almost exclusively within the first 18 months following diagnosis, with 80% occurring prior to, or very shortly after a diagnosis of TB. This suggests a missed diagnosis or serious delay in diagnosis and advanced disease. Strikingly, the excess mortality of TB patients persists well beyond 5 years

even after successful treatment. In patients with TB the rate of death due to cancer was nearly double as was the rate of death from respiratory disease other than TB.

Interventions to improve early diagnosis and treatment of TB may help reduce mortality as well as interventions to increase patient awareness of symptoms and when to seek health care advice. Awareness needs to be raised in health care professionals, not only of the risk factors for TB disease and death, but also that TB patients with a diagnosis of chronic respiratory disease and cancer have a higher risk of death.

By having a greater understanding of TB deaths it should be possible to develop more appropriate ways to diagnose and control TB and reduce this potentially preventable cause of death.

Chapter 8

Discussion

8.1 Introduction

This thesis describes the epidemiology of TB from a primary care perspective using data from the General Practice Research Database. It is the largest study of TB in primary care to date and uses classical epidemiological methods in a case-control and a cohort study, and a more recently developed but well established method in two self-controlled studies. It investigates the epidemiology of TB by assessing primary care consultations prior to a diagnosis of TB and morbidity and mortality following TB. In so doing it examines the role of primary care in detection, management and patient outcomes.

This research is important because TB is a growing public health problem;[2,68,70] and most TB research focuses on secondary care. This leads to the role of primary care being poorly understood.[8,10-13] Primary care has a vital role in identifying TB early to reduce patient morbidity, mortality and ongoing disease transmission. To diagnose TB early it is essential to understand its epidemiology from a primary care perspective, the diagnostic process in primary care, the relationship between the onset of disease and the start of treatment and the outcomes for TB patients. The results of this thesis increase our understanding of these aspects of TB and could be used to inform ways to diagnose TB earlier and improve our prevention and control of TB in the UK.

8.2 The Epidemiology of TB from a Primary Care Perspective

Epidemiology as defined in chapter 1 is "the study of the distribution and determinants of health-related states or events in specified populations, and the application of this study to the control of health problems."[9] This thesis analyses the epidemiology of TB from a primary care perspective: pre-TB, via a case-control and a self-controlled study, and morbidity and mortality post-TB, via a self-controlled and a cohort study. By using

the GPRD as the data source, a longitudinal picture of the course of TB disease and its impact on patients can be obtained: this is unique.

8.2.1 The Epidemiology of TB Pre-Diagnosis

In the 6 months prior to a diagnosis of TB the research presented in this thesis shows that there are clear differences between cases and controls. The case-control study reveals that TB cases consult three times more often than controls, with this increasingly evident from 4 months pre-TB; cases have higher rates of consultation for TB symptoms and respiratory infections, and twice the amount of chronic disease with a particularly strong association between TB and chronic respiratory disease.

The self-controlled study reveals similar findings but of a smaller magnitude: cases consult twice as often as controls and show a 3.5 fold increase in consultations for the classic symptoms of TB and a nearly 3 fold increase in consultations for 'respiratory infections' (see Table 8.1).

The case-control and the self-controlled studies use different methodologies but both assess consultation behaviour in the 6 months pre-TB. Consultation behaviour was assessed in cases and controls (whether randomly selected controls for the case-control study or self controls for the self-controlled study) by comparing consultation rates, the frequency of consultations for TB symptoms and respiratory diagnoses and the relative importance of these symptoms and respiratory diagnoses and a later diagnosis of TB. The relative importance of symptoms and respiratory diagnoses are documented in the case-control study by the 'odds of TB' and in the self-controlled study by the 'incidence rate ratio' (IRR).

The odds ratios for TB in the case-control study and the IRRs for TB in the self-controlled study were different (see Table 8.1). For example, for cases with 'any' of the symptoms of TB the odds of TB was 8.37 times higher, whereas the IRR was 3.48 times higher. Although not directly comparable these differences may have occurred by chance, may represent a true difference between the samples compared (such as more chronic disease in cases), may represent an error in the results due to bias (for example in the different selection of cases and controls or the impact of the lack of ethnicity / life style information) or could be due to confounding (see Section 8.4 for a critique of the methods used).

Table 8.1 Comparison of case-control and self-controlled study results

Symptoms / Respiratory diseases	**Case-control study** **OR** (95% CI)	**Self-controlled study** **IRR** (95% CI)
Cough	**5.20** (4.56 – 5.93)*	**2.51** (2.15 - 2.93)*
Fever	**9.66** (6.92 – 13.50)*	**4.71** (2.99 - 7.42)*
Sweating	**70.00** (16.68 – 293.82)*	**13.00** (3.86 - 43.80)*
Lymphadenopathy	**32.90** (21.61 – 50.11)*	**7.15** (4.16 - 12.29)*
Haemoptysis	**32.50** (7.33 – 144.02)*	**-**
Weight loss	**14.05** (9.80 – 20.15)*	**8.93** (5.19 - 15.38)*
Any TB symptom	**8.37** (7.43 – 9.41)*	**3.48** (3.03 – 4.00)*
All respiratory diseases (incl. resp. infection)	**6.81** (6.19-7.51)*	**2.28** (2.07 – 2.50)*
All Respiratory infections	**7.31** (6.61 – 8.09)*	**2.73** (2.47 – 3.01)*
LRTI	**7.81** (6.92 – 8.82)*	**2.20** (1.95 – 2.49)*

**p<0.001*

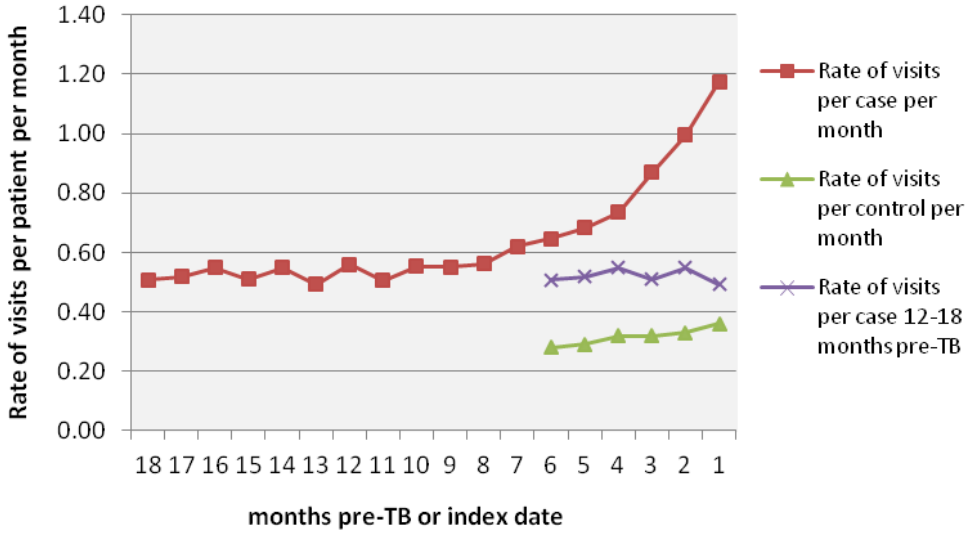
The important result of these studies pre-TB, is the fact that both studies show that in the 6 months before diagnosis, cases consult differently and more often than controls. This suggests that there is a delay in the diagnosis of TB in primary care but also that opportunities exist to diagnose TB earlier. Chapters 4 and 5 show that from 4 months pre-diagnosis there is a clear window of opportunity in which GPs could intervene and diagnose TB earlier. Dissemination of the findings of these studies to GPs may raise awareness of the different presentations of TB and the strength of their association which in turn may assist GPs to make an earlier diagnosis.

The case-control and self-controlled studies suggest that patients who consult on multiple occasions over a 6 month period, particularly with cough or recurrent respiratory infections, and without explanatory long-term chronic disease have a greater likelihood of TB and should be carefully assessed by GPs for TB.

Of note, there was a clear difference in case and control consultations even at 6 months pre-TB (a time when cases are generally considered to be TB-free). There was a two-fold difference in case and control consultation rates and a 27% difference between case and self-control consultation rates (see Figure 8.1). This difference in consultation rate between cases, controls and self-controls may reflect the higher rates of chronic disease

in cases (adjusted OR:2.69) and the therefore higher rates of consultation through routine chronic disease follow-up. Further research is needed to understand the association between chronic respiratory disease and TB.

Figure 8.1 Consultation rates in the months pre-TB or index-date for cases, controls and self-controls superimposed for comparison



This thesis presents analysis of pulmonary and extra-pulmonary TB and shows that these subtypes of TB present differently. The research literature suggests that patients with extra-pulmonary TB experience longer delays prior to diagnosis.[118] Understanding this and how cases present is important to assist faster diagnosis by GPs. My research showed that patients with extra-pulmonary TB were younger, had less chronic disease and were more likely to be female than those with pulmonary TB. This confirms work by others.[173-176] As extra-pulmonary TB is on the rise in the UK[47] a higher index of suspicion by health care workers is needed. My research clearly showed that patients with pulmonary and extra-pulmonary TB consulted differently with consultations for lymphadenopathy and fever being more common in extra-pulmonary TB, and cough and weight loss in pulmonary TB. Interestingly multiple visits are seen equally in pulmonary and extra-pulmonary TB.

Male gender, smoking and individual chronic diseases (particularly chronic respiratory disease) were shown by this research to be independent risk factors for TB; this confirms and extends the research of others.[47,158-160,165-167,169] Raising awareness of these risk factors along with those found by Story et al.,[133] which included born abroad,

homelessness and problem drug use, could help primary care practitioners diagnose TB earlier and reduce poor outcomes.

8.2.2 The Epidemiology of TB Post-Diagnosis

Morbidity following TB was examined in a self-controlled study. This showed that patients who survive at least 18 months post-TB have a good long-term prognosis; morbidity was low and the only significant finding was an increase in inhaler prescribing (p=0.006). Overall, it appears that TB patients once successfully treated return to pre-TB consultation rates and morbidity levels.

Mortality was examined in a cohort study which showed that TB leads to increased mortality (adjusted HR:1.53) and a case fatality rate at one year of 6.5%. Mortality was particularly high in the first year following diagnosis and in large part this was due to death from undiagnosed and untreated TB. Undiagnosed TB, pulmonary TB, age over 60, male gender, malignancy and chronic disease were all found to be associated with a worse prognosis. This research adds substantially and uniquely to the existing UK literature on risk factors for death in TB patients.[16,18,27,183] A TB patients chance of survival was 81% at 5 years and 70% at 10 years, survival for those without TB was 87% and 79% respectively. Death from TB disease occurred almost exclusively within the first 18 months following diagnosis, with 80% occurring prior to, or very shortly after, a diagnosis of TB. This suggests a missed diagnosis or serious delay in diagnosis and advanced disease. Strikingly, the excess mortality of TB patients persists well beyond 5 years, even after successful treatment. Few other studies have looked at this length of follow-up.[16,27] Excess mortality beyond 18 months is mostly due to underlying chronic disease or cancer. Further research is needed to understand these associations in greater depth.

This thesis shows that chronic disease appears to increase the risk of developing TB, delay diagnosis and increase the likelihood of death. The latter association confirms work by Fielder that underlying medical conditions are independent risk factors for death.[32] It is important that GPs are made aware of the association between chronic disease and TB mortality as some of the symptoms of TB and chronic respiratory disease are obviously similar.

For those patients that die at or near diagnosis it was not possible from GPRD data to distinguish whether the TB diagnosis was made clinically or at post-mortem. It is a sad reflection of UK health care at the start of the 21st century, when TB diagnostics have been available for over 100 years and good treatment for at least 50,[191] that 17% of TB cases that died were diagnosed with TB only at or near death this represents 80% of those dying of TB-disease. Diagnosis at death suggests an unacceptable and potentially avoidable delay in diagnosis and treatment to such a degree that it has resulted in death. Ways to reduce this delay need to be further investigated but would probably need to take a two-pronged approach: firstly to educate patients of the signs and symptoms of TB and where to seek help so that they present earlier to health care professionals, and secondly to educate health care professionals in primary care and accident and emergency departments to "THINK TB". Clinicians should be encouraged to have a high index of suspicion for TB and consider TB in any differential diagnosis for a variety of conditions especially in older patients with recent chest infections, recurrent non-specific consultations or chronic respiratory disease that is not responding to the usual forms of treatment. Increasing awareness among health care workers of 'red flag' symptoms as discussed in Chapters 4 and 5 plus linking these to risk factors for death identified by this research – male gender, older age, chronic disease, multiple consultations and chest infections could help increase investigation, diagnosis and treatment of TB and in turn lead to reduced mortality. Khan et al.[189] showed that physician experience and knowledge of TB led to a decrease in deaths. This is evidence to support the idea that TB education for health care workers improves patient survival.

8.3 Diagnostic Delay

Diagnosing TB in low prevalence settings, and in primary care, is challenging as the research presented in this thesis shows and is consistent with research conducted by others.[8, 19, 20, 24, 109, 112] By extrapolation, a diagnostic delay of up to 4 months is seen here. There are several possible reasons why TB is diagnosed late in the UK: firstly, patients may be unaware of the significance of their symptoms; secondly, doctors are less experienced at recognising and managing TB;[74] and thirdly, the geographic distribution varies with high and increasing rates seen in London and other large cities but much lower rates elsewhere.[73] The diagnostic challenge that TB poses in primary

care is in part due to its variable and non-specific presentation. The difficulty in recognising that early symptoms are due to TB is likely to account for much of the diagnostic delay seen in this research. Cough is a common problem in primary care and for the majority it will not represent TB. This argument is well discussed by Nardell in his editorial: "Needles in haystacks: diagnosing tuberculosis under low prevalence conditions."[192] Ultimately, the non-specific nature of the symptoms and signs of TB makes a degree of diagnostic delay inevitable no matter what interventions are used.

Understanding the risk factors that lead to a delay in TB diagnosis are crucial to reducing delay. I discussed these in depth in Chapter 2. Known risk factors for diagnostic delay include: sputum smear negative TB,[111,121, 122] extra-pulmonary TB,[20,121] no respiratory symptoms,[111,114] a history of asthma,[114] HIV infection,[122] no chest x-ray at the first visit,[23,114] ethnic group (white>ethnic),[19] older age,[25,111] female gender[19,25, 111] and an alternative diagnosis for which an antibiotic is prescribed.[114,129]

Diagnostic delay is preventable or at least reducible and ways in which it might be reduced in a primary care setting include:

- Educate the public on TB disease and transmission
- Initiate targeted population interventions to improve TB knowledge
- Re-educate healthcare workers (particularly those in primary care) on the signs and symptoms of TB and risk factors associated with TB
- Raise the level of suspicion for TB in health care professionals.
- Search out active cases via contact tracing investigations
- Provide UK primary care physicians with clear diagnostic tips / an algorithm for diagnosing TB (see Section 8.6 for more details)

This research shows that TB remains a diagnostic challenge and that diagnostic delay exists in the UK. A delay in diagnosis is an issue for public health and the prevention and control of TB. This research suggests that more needs to be done to raise awareness among GPs of TB, its risk factors and other key diagnostic features in the hope that this will lead to earlier diagnosis and treatment; and in turn reduce the likelihood of TB transmission (see Section 8.6).

8.4 A Critique of the Methods Used

The General Practice Research Database is a powerful tool for epidemiological and primary care research. The strengths of the GPRD for this thesis include its size, its comprehensive coverage of the UK population (so findings are generalisable within the UK setting),[76] the fact that it contains continuous, longitudinal primary care data which is high-quality and validated[80] and that it has been previously used in epidemiological studies, particularly respiratory studies.[78,80,83,90-92] Its major limitations for a study looking at TB, however, are its lack of ethnicity data, HIV data and variations in recording. To overcome the lack of ethnicity and socioeconomic data, I used a self-controlled methodology in addition to a case-control methodology. Additional limitations include the fact that the GPRD consists of observational data, there is no requirement to enter diagnoses for minor consultations, data recording and analysis is complex, it has poor smoking data (a potential confounder) and the details of consultations may be hidden in the unsearchable free-text field. Finally, the geographic distribution of GP practices submitting data to the GPRD may have meant that the patients analysed in this research were a slightly different group to 'all' UK TB patients. This is because, the GPRD only contains 12% of London GP practices and yet 40% of the UKs TB cases are resident in London. However, as the GPRD dataset does not provide the actual geographic locality of each patient I may well have analysed more patients resident in London and urban settings due to the fact that TB is more prevalent in these areas. On balance, the GPRD provided a unique resource for assessing TB epidemiology and TB patients pre-diagnosis and long-term.

An initial weakness of the overall study design was the fact that it excluded patients who were not registered with a GP, such as new-entrants and homeless patients who are at higher risk of TB. However a recent study in London has identified that the vast majority of TB patients are registered with a GP.[105]

To understand the diagnosis and management of TB in primary care and to identify risk factors for TB I conducted a case-control study as it offered an efficient study design and allowed me to take multiple controls. Due to confounding, however, a case-control study may not show the true association between an exposure and a disease or in this study between consultations pre-diagnosis and disease. At the time of analysis the GPRD dataset did not contain any ethnicity data (a major confounder for TB when over

two thirds of UK TB occurs in people from ethnic minorities) so although the dataset was large, confounding may have affected the results. To try to minimise confounding I selected controls from the same GP practice and in a similar age band to TB cases. I chose controls from the same GP practice, with the assumption that patients registered with the same GP were more likely to be from a similar social background and ethnic group; although realising that considerable confounding was likely to still exist.

To confirm the significance of the results of my case-control study I conducted a self-controlled (case-only) study as a second method to examine consultation behaviour pre-TB. Self matching of cases eliminated control selection bias and removed confounding by subject characteristics such as gender, ethnicity, chronic illness, social class and life-style.[83,140-142,143] Since the self-controlled analysis included only cases with at least 18 months of data prior to a TB diagnosis and only one control per case (that is itself) it was less powered than the case-control study. I used the self-controlled methodology mostly in a descriptive way, rather than analytically, to further understand the presentation of TB cases in primary care.

By using two different methods to essentially measure the same thing, the relationship between consultation behaviour prior to diagnosis and disease, I attempted to understand and minimise the impact of control selection bias and possible confounders. In general, the results of the self-controlled study did not differ greatly from those of the case-control study. The self-controlled study confirmed the significance of increased consultation rates, a diagnostic delay of up to 4 months and other findings of the case-control study and showed confounding and control selection bias could be usefully minimised by this method. The most striking difference between the two studies was the magnitude of the findings in the case-control study.

Published research clearly shows that TB is more common in ethnic minorities and lower socioeconomic groups. I favour the results of the self-controlled study as the truer representation of the impact of TB disease on patients. This is because the self-controlled study, removes the effect of ethnicity and socioeconomic status, in addition, to removing the impact of chronic disease and smoking.

I have discussed most of the strengths and limitations of each study within the discussion section of the relevant chapters and so will not do so again here.

8.5 Implications for Patients

This research shows a delay in diagnosis of up to four months was not unusual but that patients in the main presented to their GP multiple times before diagnosis. There is therefore a window of opportunity for diagnosis. Increasing patient's knowledge through education campaigns of TB, its symptoms, signs and risk factors might help patients raise the possibility of TB as a diagnosis with their GP.

This thesis provides new information on the impact of TB on patient's lives. It shows that TB does not influence long-term morbidity and that mortality, if it occurs from TB-disease is most likely early, where as long-term mortality although slightly elevated is similar in cause to the general population. This information if disseminated could provide patients with a more realistic picture of the implications of their disease.

This research suggests that TB screening should be actively encouraged in patients with unexplained, recurrent, nonspecific symptoms or chest infections and any patient with symptoms suggestive of TB.

8.6 Implications for Primary Care

There is little research on TB in primary care. A common finding it that GPs have a major role to play in the early diagnosis and treatment of TB.[8,10,101,103,108] This thesis strongly supports the need for earlier diagnosis in primary care. To help GPs make an earlier diagnosis, awareness of TB needs to be raised. This research suggests that awareness of the following could be used as 'red flags' to raise a GP's level of suspicion for TB:

- Recurrent chest infections over a few months at a level higher than expected for this patient, or more than three chest infections in a 6 month period
- Respiratory symptoms not responding to antibiotics
- Recurrent cough (to trigger assessment for the other classic symptoms of TB, a chest x-ray and sputum specimen)
- Any of the rarer TB symptoms i.e. weight loss, haemoptysis or lymphadenopathy

- Repeat, unexplained visits to the GP for non-specific symptoms
- Repeat, unexplained visits to the GP in which no diagnosis is made
- Multiple visits over a few months at a level higher than expected for this patient
- In GP practices, or geographic areas, with a higher proportion of ethnic minorities, TB should be high on any list of differentials for patients with an unexplained or unresolved cough or any of the classic symptoms of TB.

Combining the knowledge of these 'red-flags' with the other features highlighted in this thesis such as the risk factors for TB (male gender, older age, suffering a chronic disease), the strength of association between clinical presentations and TB and raising awareness of newer investigations and the local referral process may help GPs make an earlier diagnosis. Practically the recurrent nature of non-specific, non-resolving symptoms or recurrent respiratory infections, particularly in patients with a known chronic disease, would be something that GP computer systems could flag and in so doing raise a GPs level of suspicion for TB. Earlier investigation could be encouraged by the use of an electronic prompt based on an algorithm that draws together the electronic back-data held on a patient such as consultation rates, symptoms, signs and risk factors for TB.

Informing primary care professionals of the above, of the current epidemiology, risk factors and clinical presentations of TB and the process of local referral may lead to improved knowledge of the disease as well as earlier diagnosis. This should be a high priority in the UK, or at least in areas of higher incidence, and is supported by the work contained within this thesis and the limited existing literature on TB in primary care.[11-13,74] There are numerous ways of raising awareness in primary care professionals: online educational packages, articles in peer reviewed journals, courses, and locally run TB awareness training.

This thesis shows that patients consult their GP multiple times pre-TB. These visits therefore provide an opportunity to potentially diagnose TB earlier. The non-specific nature of the early symptoms of TB can, however, make diagnosis harder for GPs. This research highlights the strength of association between TB and its symptoms and TB and respiratory infections which might help GPs make an earlier diagnosis. In terms of

strength of association, symptoms other than cough are more suggestive of TB, and 'alert' symptoms include fever, weight loss and lymphadenopathy plus recurrent (greater than three), lower respiratory infections over a 6 month period. Consultations for fever and weight loss rise only in the last 1 month pre-diagnosis; it therefore seems unlikely that presentation with these will reduce existing diagnostic delay.

Raising awareness of TB symptoms alone may only help GPs diagnose a minority of cases as only 27% of cases in this study had symptoms recorded in a searchable form. It is likely however, that the coded symptoms are not a true reflection of the symptoms complained of, due to the fact, that GPs tend to record 'diagnoses' rather than 'symptoms' (as discussed in Chapter 4). Further symptom information is likely to be recorded in the free-text field of the GPRD, but this is not easily searchable. This work suggests that raising awareness of TB symptoms and their strength of association is important but needs to be part of a package of measures to raise awareness among GPs of the local epidemiology of TB, risk factors for TB, other disease associations and the process of local referral for diagnostic tests. Further research is needed to assess the free-text fields of the GPRD for further information on symptoms and the strength of their association to TB.

This study confirms smoking as a risk factor for TB. The risk of smoking in TB cases is one of the most modifiable of all the risk factors. Smoking not only increases the risk of developing TB[167,168] but also of suffering long-term lung damage.[166] Primary care practitioners by encouraging and supporting TB patients who smoke to stop could therefore have a key role in improving the outcomes of these patients.

Raising awareness through improved education among health professionals is essential,[11-13] and particularly for general practitioners, to help them diagnose TB earlier. Continuing professional development online modules may be useful for this. Davies in a personal paper "The challenge of tuberculosis"[74] stated that "the most important challenge for TB care in the UK today is the re-education of the medical and allied professions in diagnosis and management of the disease." This statement is very much supported by my research. TB awareness-raising among health care professionals was highlighted as an area of need in the Chief Medical Officer's TB Action Plan. As primary care is often the first place a symptomatic patient is seen it is important to raise awareness of TB in primary care. Using the findings of my research I have written an

online educational package about TB for GPs available on doctors.net and for nurses available on nursing.net. These online educational packages aim to improve knowledge, diagnosis and treatment of TB cases. In addition, in my position as a Consultant in Communicable Disease Control I lecture to NW London GPs, practice nurses and health visitors on TB, its symptoms, signs and control to try and raise awareness locally. Education of health care workers should not only be to improve their knowledge of TB but also to raise awareness of the risk factors for TB disease and death and those groups most affected.

8.7 Implications for TB Prevention and Control

This thesis examines the epidemiology of TB and provides a better understanding of the diagnosis, management and outcomes of TB. By understanding the role of primary care in TB diagnosis, reviewing the problem of late diagnosis and assessing the long-term outcomes of patients in terms of morbidity and mortality, recommendations can be made to improve TB control and prevention in the UK.

It is important for national TB programmes to understand TB patient deaths so that they can assess how effective their TB management and control programmes have been. This is particularly true in high income countries where death is often the most commonly reported reason for TB treatment failure.[47]

My cohort study looking at mortality gives a unique insight into the risk factors for death from TB in a large sample population. Knowledge of these risk factors in conjunction with an increase in consultation rates and repeat chest infections could help health care professionals diagnose TB earlier, and lead to quicker treatment and decreased mortality. Specifically in the UK, consideration of a diagnosis of TB in older, male patients with a known chronic disease (particularly chronic respiratory disease) may help to reduce TB mortality.

This research shows that there is a high burden of potentially preventable mortality as shown by the 80% of deaths from TB-disease that occur shortly before, or after diagnosis. The information on mortality and its risk factors discussed in this thesis could be used to inform the development of strategies to minimise mortality.

For clinicians based in both primary and secondary care, this research highlights the relative importance of specific symptoms and signs as diagnostic markers for TB. For patients, this work could mean faster diagnosis, quicker treatment and reduced transmission.

Serious consideration needs to be given to how to prevent the high early mortality in TB patients. The research presented here suggests that diagnosing TB earlier (to reduce morbidity and mortality) is not just about educating GPs and patients about the symptoms and signs of disease but also what appears to be needed are other ways to improve case detection. Examples of this could include increased screening of patient's for TB on registration with a new GP, particularly if patients are from an ethnic minority, newly arrived immigrants, have existing chronic disease or are smokers; or encouraging earlier investigation of patient's who present with symptoms and signs compatible with TB, or unexplained respiratory disease; possibly via an electronic prompt within a GPs computer system that encourages consideration of TB.

With the increasing use and availability in hospitals of new blood tests for TB (the interferon gamma release assays - IGRAs)[193] it may be appropriate to review their use and availability in primary care. IGRA tests indicate mycobacterial infection but cannot distinguish between latent from active TB. This is an area where further research is needed by national bodies to determine the best use of IGRAs and their cost effectiveness.

8.8 Unanswered Questions and Future Research

As with any research, constraints of time place limits on what can be achieved. Further research that could usefully follow from this thesis includes:

1. An analysis of the GPRD free-text fields to assess what specific TB symptoms GPs recorded during the six months pre-TB; or to assess if TB was mentioned in the differential diagnosis prior to it appearing as the ‘actual’ diagnosis. This would help determine if diagnostic delay is in fact shorter than this research suggests.

2. A cohort study using GP practices to measure the positive predictive value of different symptoms and the incidence of TB; two things unobtainable from the studies discussed in this thesis.
3. A critical case review of patients dying from TB disease, looking at patient pathways and avoidable causes.
4. The development and trialling of an electronic decision tool for TB, for use in primary care, and based on these research findings.

8.9 Conclusion

This thesis presents the largest study of TB in primary care to date. By using the General Practice Research Database and up to 12 years of longitudinal data I have gained new insight into TB in the UK. As a consequence this research provides a wealth of information on TB from a primary care perspective. It increases our understanding of the early progression of TB, the risk factors for TB, the strength of association between TB, its symptoms and respiratory diagnoses plus quantification of the long-term outcomes of TB both in terms of morbidity and mortality. This research offers a unique view of the epidemiology of TB and the role of primary care in its diagnosis and control. It fills a gap in the research literature on the role of primary care in TB and the long-term outcomes of TB patients. Overall, this research indicates a need for improved patient and professional awareness of TB and emphasises the need for earlier diagnosis. The dissemination of the findings of this work could shorten diagnostic delay and improve patient outcomes.

Appendix 1

Detailed Search Strategies for Literature Review

TB and Primary Care

No.	Database	Search term
1	EMBASE, MEDLINE, CINAHL	exp TUBERCULOSIS/
2	EMBASE, MEDLINE, CINAHL	"primary care".ti,ab
3	EMBASE, MEDLINE, CINAHL	"general practice".ti,ab
4	EMBASE, MEDLINE, CINAHL	"family practice".ti,ab
5	EMBASE, MEDLINE, CINAHL	"family practitioner".ti,ab
6	EMBASE, MEDLINE, CINAHL	"general practitioner".ti,ab
7	EMBASE, MEDLINE, CINAHL	"family physician".ti,ab
8	EMBASE, MEDLINE, CINAHL	exp FAMILY PRACTICE/
9	EMBASE, MEDLINE, CINAHL	exp PHYSICIANS, FAMILY/
10	EMBASE, MEDLINE, CINAHL	2 OR 3 OR 4 OR 5 OR 6 OR 7 OR 8 OR 9
11	EMBASE, MEDLINE, CINAHL	1 AND 10
12	EMBASE, MEDLINE, CINAHL	11 [Limit to: Publication Year 1980-2009]
13	EMBASE, MEDLINE, CINAHL	12 [Limit to: Humans and English Language] for EMBASE, MEDLINE

TB and Diagnostic Delay

No.	Database	Search term
1	EMBASE, MEDLINE	exp TUBERCULOSIS/
2	EMBASE, MEDLINE	diagnos*.ti,ab
3	EMBASE, MEDLINE	delay*.ti,ab
4	EMBASE, MEDLINE	late*.ti,ab
5	EMBASE, MEDLINE	3 or 4
6	EMBASE, MEDLINE	2 and 5
7	EMBASE, MEDLINE	1 and 6
8	EMBASE, MEDLINE	7 [Limit to: Publication Year 1980-2009 and Humans and English Language]

Appendix 2

TB Medical Codes Used in GPRD Search

medcode	pcpsmedcode	flag	description
010 A	100123	O	TUBERCULOSIS WITH ASBESTOSIS
010 C	100124	O	SILICOTUBERCULOSIS
010 M	100125	O	TUBERCULOSIS WITH MINERS' LUNG
011	100126	O	TUBERCULOSIS PULMONARY
011 A	100127	O	TUBERCULOSIS
011 B	100128	O	TUBERCULOUS ABSCESS LUNG
011 C	100129	O	TUBERCULOSIS MILIARY LUNG
011 D	100130	O	TUBERCULOUS BRONCHIECTASIS
011 P	100132	O	TUBERCULOSIS PRIMARY PULMONARY
0120P	100133	O	T B PRIMARY RESPIRATORY (TUBERCULOSIS)
0121	100134	O	TUBERCULOSIS WITH PLEURISY
0121A	100135	O	TUBERCULOUS ADHESIONS PLEURA
0121B	100136	O	TUBERCULOUS EMPYEMA
0121C	100137	O	TUBERCULOUS PLEURAL EFFUSION
0123	100140	O	TUBERCULOUS LARYNGITIS
0124A	100141	O	TUBERCULOMA
0129	100142	O	TUBERCULOUS LYMPHADENOPATHY
0129A	100143	O	TUBERCULOUS NASAL SINUS
0129B	100144	O	TUBERCULOUS LYMPH NODES HILAR
0129BM	100145	O	TUBERCULOUS LYMPH NODES MEDIASTINAL
0130	100146	O	TUBERCULOUS MENINGITIS
0130M	100147	O	TUBERCULOUS MENINGOMYELITIS
014	100148	O	TUBERCULOSIS ABDOMINAL
014 A	100149	O	TUBERCULOUS ABSCESS ISCHIORECTAL
014 B	100150	O	TUBERCULOUS LYMPH NODES MESENTERIC
014 C	100151	O	TUBERCULOUS PERITONITIS
014 D	100152	O	TUBERCULOSIS INTESTINAL
014 DC	100153	O	TUBERCULOSIS ILEOCAECAL
014 H	100154	O	TUBERCULOUS HEPATITIS
014 RP	100155	O	TUBERCULOUS ABSCESS RETROPERITONEAL
0150A	100156	O	TUBERCULOSIS SPINE
0150AA	100157	O	TUBERCULOUS ABSCESS JOINT VERTEBRAL
0150AB	100158	O	TUBERCULOUS ABSCESS VERTEBRA
0150B	100159	O	TUBERCULOUS ABSCESS LUMBAR
0150C	100160	O	TUBERCULOUS ABSCESS ILIOPSOAS
0150D	100161	O	TUBERCULOUS ABSCESS SACRUM
0150G	100164	O	TUBERCULOUS SPONDYLITIS
0151	100165	O	TUBERCULOUS JOINT HIP
0152	100166	O	TUBERCULOUS JOINT KNEE
0158A	100167	O	TUBERCULOUS JOINT
0158B	100168	O	TUBERCULOUS BONE
0158C	100169	O	TUBERCULOUS MASTOIDITIS
0158D	100170	O	TUBERCULOUS DACTYLITIS
0159	100171	O	TUBERCULOSIS SKELETAL

medcode	pcpsmedcode	flag	description
0159A	100172	O	TUBERCULOUS OSTEOMYELITIS
0159B	100173	O	TUBERCULOUS ARTHRITIS
0159C	100174	O	TUBERCULOUS TENOSYNOVITIS
0159CW	100175	O	T B ABSCESS CHEST WALL (TUBERCULOSIS)
016	100176	O	TUBERCULOSIS GENITO-URINARY SYSTEM
016 A	100177	O	TUBERCULOUS ORCHITIS
016 BL	100178	O	TUBERCULOSIS URINARY BLADDER
016 EP	100179	O	TUBERCULOSIS EPIDIDYMIS
016 F	100180	O	TUBERCULOSIS FALLOPIAN TUBE
016 K	100181	O	TUBERCULOSIS KIDNEY
016 P	100182	O	TUBERCULOSIS PROSTATE
016 PA	100183	O	TUBERCULOSIS SEMINAL VESICLES
016 PB	100184	O	TUBERCULOSIS PROSTATE & SEMINAL VESICLES
016 PN	100185	O	TUBERCULOSIS PENIS
016 TE	100186	O	TUBERCULOSIS URETER
016 TH	100187	O	TUBERCULOSIS URETHRAL
0170A	100188	O	TUBERCULOUS ERYTHEMA NODOSUM
0170B	100189	O	TUBERCULOSIS LUPUS VULGARIS
0171	100196	O	TUBERCULOUS ADENITIS
0171A	100197	O	TUBERCULOSIS AXILLA GLAN
0171NG	100198	O	TUBERCULOUS GLANDS NECK
0172A	100199	O	TUBERCULOUS CHOROIDITIS
0172L	100200	O	TUBERCULOSIS ACHROACYTOSIS LACHRYMAL GLA
0179	100201	O	TUBERCULOUS PERICARDITIS
0179AR	100202	O	T B RELAPSE (TUBERCULOSIS)
0179B	100203	O	TUBERCULOSIS BREAST
0179C	100204	O	TUBERCULOSIS STOMACH
0179D	100205	O	TUBERCULOSIS ADRENAL GLANDS
0179DA	100206	O	TUBERCULOUS ADDISON'S DISEASE
0181	100207	O	TUBERCULOSIS MILIARY
0189C	100208	O	CHRONIC MILIARY TUBERCULOSIS
5202CT	105094	O	OCCLUSAL TUBERCULUM
4E38.00	124561	R	Sputum: tubercle on Z-N stain
65V9.00	127765	R	Notification of tuberculosis
65V9.11	127766	R	TB - tuberculosis notification
A1...00	147742	R	Tuberculosis
A10..00	147743	R	Primary tuberculous infection
A100.00	147744	R	Primary tuberculous complex
A101.00	147745	R	Tuberculous pleurisy in primary progressive tuberculosis
A10y.00	147746	R	Other primary progressive tuberculosis
A10z.00	147747	R	Primary tuberculous infection NOS
A11..00	147748	R	Pulmonary tuberculosis
A11..11	147749	R	Lung tuberculosis
A110.00	147750	R	Infiltrative lung tuberculosis
A111.00	147751	R	Nodular lung tuberculosis
A112.00	147752	R	Tuberculosis of lung with cavitation
A113.00	147753	R	Tuberculosis of bronchus
A114.00	147754	R	Tuberculous fibrosis of lung
A115.00	147755	R	Tuberculous bronchiectasis
A116.00	147756	R	Tuberculous pneumonia

medcode	pcpsmedcode	flag	description
A117.00	147757	R	Tuberculous pneumothorax
A11y.00	147758	R	Other specified pulmonary tuberculosis
A11z.00	147759	R	Pulmonary tuberculosis NOS
A12..00	147760	R	Other respiratory tuberculosis
A120.00	147761	R	Tuberculous pleurisy
A120000	147762	R	Tuberculosis of pleura
A120100	147763	R	Tuberculous empyema
A120200	147764	R	Tuberculous hydrothorax
A120z00	147765	R	Tuberculous pleurisy NOS
A121.00	147766	R	Tuberculosis of intrathoracic lymph nodes
A121000	147767	R	Tuberculosis of hilar lymph nodes
A121100	147768	R	Tuberculosis of mediastinal lymph nodes
A121200	147769	R	Tuberculosis of tracheobronchial lymph nodes
A121z00	147770	R	Tuberculosis of intrathoracic lymph nodes NOS
A122.00	147771	R	Isolated tracheal or bronchial tuberculosis
A122000	147772	R	Isolated tracheal tuberculosis
A122100	147773	R	Isolated bronchial tuberculosis
A122z00	147774	R	Isolated tracheal or bronchial tuberculosis NOS
A123.00	147775	R	Tuberculous laryngitis
A124.00	147776	R	Resp TB bacteriologically and histologically confirmed
A124000	147777	R	TB lung confirm sputum microscopy with or without culture
A124100	147778	R	Tuberculosis of lung, confirmed by culture only
A124200	147779	R	Tuberculosis of lung, confirmed histologically
A124300	147780	R	Tuberculosis of lung, confirmed by unspecified means
A124400	147781	R	TB intrathoracic lymph nodes confirm bact histologically
A124500	147782	R	Tuberculosis of larynx, trachea & bronchus conf bact/hist'y
A124600	147783	R	Tuberculous pleurisy, conf bacteriologically/histologically
A124700	147784	R	Primary respiratory TB confirm bact and histologically
A125.00	147785	R	Respiratory TB not confirmed bact or histologically
A125000	147786	R	Tuberculosis of lung, bacteriologically & histolog'y neg
A125100	147787	R	Tuberculosis lung bact and histological examin not done
A125200	147788	R	Prim respiratory TB without mention of bact or hist confirm
A125X00	147789	R	Resp TB unspcf,w'out mention/bacterial or histol confrmtn
A12y.00	147790	R	Other specified respiratory tuberculosis
A12y000	147791	R	Tuberculosis of mediastinum
A12y100	147792	R	Tuberculosis of nasopharynx
A12y200	147793	R	Tuberculosis of nasal septum
A12y300	147794	R	Tuberculosis of nasal sinus
A12yz00	147795	R	Other specified respiratory tuberculosis NOS
A13..00	147796	R	Tuberculosis of meninges and central nervous system
A130.00	147797	R	Tuberculous meningitis
A130000	147798	R	Tuberculosis of cerebral meninges
A130100	147799	R	Tuberculosis of spinal meninges
A130200	147800	R	Tuberculous leptomeningitis
A130300	147801	R	Tuberculous meningoencephalitis
A130z00	147802	R	Tuberculous meningitis NOS
A131.00	147803	R	Tuberculoma of meninges
A132.00	147804	R	Tuberculoma of brain
A133.00	147805	R	Tuberculous abscess of brain
A134.00	147806	R	Tuberculoma of spinal cord

medcode	pcpsmedcode	flag	description
A135.00	147807	R	Tuberculous abscess of spinal cord
A136.00	147808	R	Tuberculous encephalitis or myelitis
A136000	147809	R	Tuberculous encephalitis
A136100	147810	R	Tuberculous myelitis
A136z00	147811	R	Tuberculous encephalitis or myelitis NOS
A13y.00	147812	R	Other specified tuberculosis of central nervous system
A13z.00	147813	R	Tuberculosis of central nervous system NOS
A14..00	147814	R	Tuberculosis of intestines, peritoneum and mesenteric glands
A140.00	147815	R	Tuberculous peritonitis
A14y.00	147820	R	Other gastrointestinal tract tuberculosis
A14y000	147821	R	Tuberculosis of anus
A14y100	147822	R	Tuberculosis of large intestine
A14y200	147823	R	Tuberculosis of small intestine
A14y300	147824	R	Tuberculosis of mesenteric lymph glands
A14y400	147825	R	Tuberculosis of rectum
A14y500	147826	R	Tuberculosis of retroperitoneal lymph nodes
A14yz00	147827	R	Other gastrointestinal tract tuberculosis NOS
A14z.00	147828	R	Tuberculosis of gastrointestinal tract NOS
A15..00	147829	R	Tuberculosis of bones and joints
A15..11	147830	R	Tuberculous osteomylelytis
A15..12	147831	R	Tuberculous arthritis
A15..13	147832	R	Tuberculous synovitis
A150.00	147833	R	Tuberculosis of vertebral column - Pott's
A151.00	147834	R	Tuberculosis of hip
A152.00	147835	R	Tuberculosis of knee
A153.00	147836	R	Tuberculosis limb bones - Tuberculous dactylitis
A154.00	147837	R	Tuberculous mastoiditis
A15x.00	147838	R	Tuberculosis of other specified bones
A15y.00	147839	R	Tuberculosis of other specified joint
A15z.00	147840	R	Tuberculosis of bones or joints NOS
A16..00	147841	R	Tuberculosis of genitourinary system
A160.00	147842	R	Tuberculosis of kidney
A160.11	147843	R	Renal tuberculosis
A160000	147844	R	Tuberculous nephropathy
A160100	147845	R	Tuberculous pyelitis
A160200	147846	R	Tuberculous pyelonephritis
A160z00	147847	R	Tuberculosis of kidney NOS
A161.00	147848	R	Tuberculosis of bladder
A162.00	147849	R	Tuberculosis of ureter
A163.00	147850	R	Tuberculosis of other urinary organs
A164.00	147851	R	Tuberculosis of epididymis
A165.00	147852	R	Tuberculosis of other male genital organs
A165000	147853	R	Tuberculosis of prostate
A165100	147854	R	Tuberculosis seminal vesicle
A165200	147855	R	Tuberculosis of testis
A165z00	147856	R	Tuberculosis of other male genital organs NOS
A166.00	147857	R	Tuberculous oophoritis or salpingitis
A166000	147858	R	Tuberculous oophoritis
A166100	147859	R	Tuberculous salpingitis
A166111	147860	R	Fallopian tube tuberculosis

medcode	pcpsmedcode	flag	description
A166z00	147861	R	Tuberculous oophoritis or salpingitis NOS
A167.00	147862	R	Tuberculosis of other female genital organs
A167000	147863	R	Tuberculous cervicitis
A167100	147864	R	Tuberculous endometritis
A167z00	147865	R	Tuberculosis of other female genital organs NOS
A16z.00	147866	R	Genitourinary tuberculosis NOS
A17..00	147867	R	Tuberculosis of other organs
A170.00	147868	R	Tuberculosis of skin and subcutaneous tissue
A170100	147871	R	Tuberculosis - lupus vulgaris
A170z00	147879	R	Tuberculosis of skin and subcutaneous tissue NOS
A171.00	147880	R	Tuberculosis with erythema nodosum hypersensitivity reaction
A171100	147882	R	Tuberculous erythema nodosum
A171z00	147883	R	Erythema nodosum with tuberculosis NOS
A172.00	147884	R	Tuberculosis of peripheral lymph nodes
A172000	147885	R	Tuberculous - cervical lymphadenitis
A172011	147886	R	Scrofula - tuberculous cervical lymph nodes
A172100	147887	R	Scrofulous tuberculous abscess
A172200	147888	R	Tuberculous adenitis
A172z00	147889	R	Tuberculosis of peripheral lymph nodes NOS
A173.00	147890	R	Tuberculosis of eye
A173000	147891	R	Tuberculous chorioretinitis
A173100	147892	R	Tuberculous episcleritis
A173200	147893	R	Tuberculous interstitial keratitis
A173300	147894	R	Tuberculous chronic iridocyclitis
A173400	147895	R	Tuberculous keratoconjunctivitis
A173z00	147896	R	Tuberculosis of eye NOS
A174.00	147897	R	Tuberculosis of ear
A175.00	147898	R	Tuberculosis of thyroid gland
A176.00	147899	R	Tuberculosis of adrenal glands - Addison's disease
A177.00	147900	R	Tuberculosis spleen
A178.00	147901	R	Tuberculosis oesophagus
A17y.00	147902	R	Tuberculosis of other specified organs
A17y000	147903	R	Tuberculosis endocardium
A17y100	147904	R	Tuberculosis myocardium
A17y200	147905	R	Tuberculosis pericardium
A17y300	147906	R	Tuberculosis of stomach
A17y400	147907	R	Tuberculosis of liver
A17yz00	147908	R	Tuberculosis of other specified organs NOS
A17z.00	147909	R	Tuberculosis of other organs NOS
A18..00	147910	R	Miliary tuberculosis
A180.00	147911	R	Acute miliary tuberculosis
A180000	147912	R	Acute miliary tuberculosis of a single specified site
A180100	147913	R	Acute miliary tuberculosis of multiple sites
A18y.00	147914	R	Other specified miliary tuberculosis
A18z.00	147915	R	Miliary tuberculosis NOS
A1y..00	147916	R	Other specified tuberculosis
A1z..00	147917	R	Tuberculosis NOS
Ayu1.00	149556	R	[X]Tuberculosis
Ayu1000	149557	R	[X]Other resp tubercul,confirmd bacteriologicly+histologicly
Ayu1100	149558	R	[X]Resp tuberculos unspcfd,confirmd bacteriolog+histologicly

medcode	pcpsmedcode	flag	description
Ayu1200	149559	R	[X]Oth resp tubercul, w'out m/bacteriol or histol confirmatn
Ayu1400	149561	R	[X]Other tuberculosis of nervous system
Ayu1500	149562	R	[X]Tuberculosis of nervous system, unspecified
Ayu1600	149563	R	[X]Tuberculosis of other specified organs
Ayu1700	149564	R	[X]Acute miliary tuberculosis, unspecified
Ayu1800	149565	R	[X]Other miliary tuberculosis
Ayu1900	149566	R	[X]Miliary tuberculosis, unspecified
AyuJ000	149819	R	[X]Sequelae of central nervous system tuberculosis
AyuJ100	149820	R	[X]Sequelae of genitourinary tuberculosis
AyuJ200	149821	R	[X]Sequelae of tuberculosis of bones and joints
AyuJ300	149822	R	[X]Sequelae of tuberculosis of other organs
AyuJ400	149823	R	[X]Sequelae of respiratory and unspecified tuberculosis
F004.00	157993	R	Meningitis - tuberculous
F033300	158093	R	Encephalitis due to tuberculosis
F033311	158094	R	Tuberculous encephalitis
F040600	158155	R	Tuberculous intracranial abscess
F041300	158161	R	Tuberculous intraspinal abscess
F4A5500	159718	R	Keratitis due to tuberculosis
G500300	161326	R	Acute pericarditis - tuberculous
G520600	161374	R	Acute myocarditis - tuberculous
H450.00	162924	R	Pneumoconiosis associated with tuberculosis
J550200	164628	R	Peritonitis - tuberculous
Jyu9300	165197	R	[X]Tuberculous disorders of intestine and mesentery
K154800	165894	R	Cystitis in tuberculosis
K214300	166196	R	Prostatitis in tuberculosis
K43..00	167129	R	Female tuberculous pelvic inflammatory disease
K803	168481	O	EXCISION TUBERCULOUS LESION JOINT
L173.00	170891	R	Maternal tuberculosis in pregnancy/childbirth/puerperium
L173000	170892	R	Maternal tuberculosis,unspec whether in pregnancy/puerperium
L173100	170893	R	Maternal tuberculosis during pregnancy - baby delivered
L173200	170894	R	Maternal tuberculosis in puerperium - baby delivered
L173300	170895	R	Maternal tuberculosis in pregnancy - baby not yet delivered
L173400	170896	R	Maternal tuberculosis in puerperium - baby previously deliv.
L173z00	170897	R	Maternal tuberculosis in pregnancy/childbirth/puerperium NOS
N018.00	174422	R	Tuberculous arthritis
N22yD00	176138	R	Tuberculous infection of tendon sheath
N304.00	176470	R	Tuberculosis of spine (Pott's)
N304.11	176471	R	Tuberculosis of spine
N304000	176472	R	Tuberculosis of cervical spine
N304100	176473	R	Tuberculosis of thoracic spine
N304200	176474	R	Tuberculosis of lumbar spine
N304300	176475	R	Tuberculosis of sacrum/coccyx
N305.00	176476	R	Tuberculosis of limb bones
N305000	176477	R	Tuberculosis of unspecified limb bone
N305100	176478	R	Tuberculosis of the upper arm bone
N305200	176479	R	Tuberculosis of the forearm bone
N305300	176480	R	Tuberculosis of the pelvic and thigh bones
N305400	176481	R	Tuberculosis of the lower leg bone
N305500	176482	R	Tuberculosis of other limb bones
N305600	176483	R	Tuberculosis of multiple limb bones

medcode	pcpsmedcode	flag	description
N305z00	176484	R	Tuberculosis of limb bones NOS
N306.00	176485	R	Tuberculosis of other bones
N306000	176486	R	Tuberculosis of bone, site unspecified
N306100	176487	R	Tuberculosis of the bones of the shoulder region
N306200	176488	R	Tuberculosis of the bones of the hand
N306300	176489	R	Tuberculosis of the bones of the ankle and foot
N306400	176490	R	Tuberculosis of the bones of other sites
N306500	176491	R	Tuberculosis of the bones of multiple sites
N306z00	176492	R	Tuberculosis of bone NOS
Q402400	180831	R	Congenital tuberculosis

Appendix 3

Drugs Used to Treat TB – BNF Code Group 05.01.09.00

Drug Code	Drug name
96773997	ETHAMBUTOL
96773998	ETHAMBUTOL
93406992	PYRAZINAMIDE
95366998	PYRAZINAMIDE
95314997	RIFAMPICIN+ ISONIAZID
95314998	RIFAMPICIN+ ISONIAZID
98684997	RIFAMPICIN+ ISONIAZID
98694998	RIFAMPICIN+ ISONIAZID
99178998	RIFAMPICIN+ ISONIAZID

Appendix 4

Top 40 GPRD Non-GP Consultations during 6 Months Pre-TB

GPRD consultation diagnosis	OXMIS or Read Code	Frequency
Did not attend - no reason	R	2050
Blood sample -> Lab NOS	R	1916
Reviewed at hospital	R	1727
Blood test	O	1587
Injection given	R	1363
Reviewed at hospital	O	1235
Influenza vaccination	R	1190
Incoming mail NOS	R	1183
DNA (did not arrive)	O	1041
X-ray chest	O	886
Full blood count	O	844
Discharged from hospital	R	835
Certificates - administration	R	713
Influenza vaccine given	O	709
Consultant letter from:	R	599
Seen in general medical clinic	R	581
Letter received	O	567
Letter sent to consultant	R	516
Urine test	O	502
Vaccination prophylactic	O	496
Full blood count - FBC	R	481
Failed encounter	R	458
Immunisations	R	453
Seen in chest clinic	R	451
MSU	O	438
U & E'S (urea & electrolytes)	O	432

GPRD consultation diagnosis	OXMIS or Read Code	Frequency
Injection given	O	413
Test liver function	O	407
Prescription given no examination of patient	R	372
Failure to attend appointment	O	355
Standard chest X-ray	R	340
Seen in hospital out-pat.	R	313
Seen in medical clinic	O	312
Letter sent to outside agency	R	311
Patient reviewed at hospital	R	295
Seen in hospital casualty	R	292
Blood glucose result	R	282
URINE CULTURE	O	275
Seen in ENT clinic	R	259

Appendix 5

Visits per Patient by TB Type over the 6 Month Period Pre-Diagnosis

Visits	ALL TB		Extra-pulmonary TB		Pulmonary TB	
	No.	%	No.	%	No.	%
0 visits pre-TB	291	9.6	44	9.1	62	8.2
1 visit pre-TB	356	11.7	60	12.4	77	10.1
2 visits pre-TB	383	12.6	62	12.8	84	11.1
3 visits pre-TB	354	11.7	58	12.0	103	13.6
>3 visits pre-TB	1648	54.4	259	53.7	433	57.0
>=6 visits pre-TB	**1121**	**37.0**	**136**	**28.2**	**248**	**32.7**

Appendix 6

IRRs and Confidence Intervals for Different TB-Types by Symptoms and Respiratory Diseases

Consultation types	ALL TB		Extra-Pulmonary TB		Pulmonary TB	
	IRR (95% CI)	***p* value**	**IRR** (95% CI)	***p* value**	**IRR** (95% CI)	***p* value**
Cough	2.51 (2.15-2.93)	<0.001	1.26 (0.79-2.01)	0.327	2.87 (2.21-3.72)	<0.001
Fever	4.71 (2.99-7.42)	<0.001	5.25 (1.44-19.18)	0.012	8.67 (2.58-29.07)	<0.001
Sweating	13 (3.86-43.80)	<0.001	-	-	-	-
Lymphadenopathy	7.15 (4.16-12.29)	<0.001	8.11 (2.38-27.69)	0.001	1.67 (0.40-6.98)	0.485
Haemoptysis	-	-	-	-	-	-
Weight loss	8.93 (5.19-15.38)	<0.001	2.5 (0.48-12.90)	0.274	16.25 (5.80-45.56)	<0.001
All respiratory diseases	2.28 (2.07-2.50)	<0.001	1.33 (0.99-1.79)	0.06	2.38 (2.03-2.78)	<0.001
Respiratory inf. (incl. cough)	2.73 (2.47-3.01)	<0.001	1.35 (0.98-1.84)	0.062	2.92 (2.46-3.47)	<0.001
LRTI only	2.20 (1.95-2.49)	<0.001	1.2 (0.82-1.76)	0.353	3.04 (2.42-3.81)	<0.001

Appendix 7

GP Consultations 12-18 Months Pre and 12-18 Months Post-TB

Pulmonary (n=477)

GP consultations per individual	**18-12 months pre-diag**	**12-18 months post-diag**	**IRR**	**Stats using GEE Model** 95% CI	**p value**
Consultations for coughs	64	69	**1.08**	(0.73 – 1.60)	0.71
Consults for resp disease	257	278	**1.08**	(0.85 – 1.38)	0.52
Consults for resp infections	190	201	**1.06**	(0.82 – 1.37)	0.67
Consultations for LRTI	102	104	**1.02**	(0.71 – 1.47)	0.92
Consultations for influenza	7	5	**0.71**	(0.25 – 2.04)	0.53

Comment *– no statistically significant results*

Extra-Pulmonary TB (n=340)

GP consultations per individual	**18-12 months pre-diag**	**12-18 months post-diag**	**IRR**	**Stats using GEE Model** 95% CI	**p value**
Consultations for coughs	27	33	1.22	(0.72 – 2.06)	0.45
Consults for resp disease	102	113	1.11	(0.75 – 1.63)	0.60
Consults for resp infections	81	85	1.05	(0.72 – 1.53)	0.80
Consultations for LRTI	46	35	0.76	(0.45 – 1.28)	0.30
Consultations for influenza	2	7	3.5	(0.60 – 20.39)	0.16

Comment *– no statistically significant results*

APPENDIX 8

Power Calculation for Case Control Study

Assuming that 50% of TB cases consult at least 3 times in the 6 months prior to a diagnosis of TB and that only 12.5% of controls consult at least 3 times in the same time period; in order to have 90% power to observe a significant difference ($p<0.05$) between cases and controls and if 5 controls per case are selected I would need at least 2,749 cases and 13,745 controls to detect an Odds Ratio of 1.15.

This size of study should also have ample power to explore the relatively subtle differences in consultation behaviour and cope with sub-group analysis.

References

1. World Health Organization. 44th World Health Assembly: WHO Report, 1991. WHA44/1991/REC/1. 1991. Geneva, Switzerland: WHO, World Health Organization.

2. Rose AM, Watson JM, Graham C et al. Tuberculosis at the end of the 20th century in England and Wales: results of a national survey in 1998. Thorax 2001; 56(3):173-179.

3. Ormerod LP, Shaw RJ, Mitchell DM. Tuberculosis in the UK, 1994: current issues and future trends. Thorax 1994; 49(11):1085-1089.

4. Department of Health. *Stopping Tuberculosis in England; An action plan from the Chief Medical Officer*. 2004. London, DH.

5. Chemotherapy and management of tuberculosis in the United Kingdom: recommendations 1998. Joint Tuberculosis Committee of the British Thoracic Society. Thorax 1998; 53(7):536-548.

6. National Collaborating Centre for Chronic Conditions. Tuberculosis: clinical diagnosis and management of tuberculosis, and measures for its prevention and control. 2006. London, Royal College of Physicians.

7. Anderson SR, Maguire H, Carless J. Tuberculosis in London: a decade and a half of no decline. Thorax 2007; 62(2):162-167.

8. Metcalf EP, Davies JC, Wood F, Butler CC. Unwrapping the diagnosis of tuberculosis in primary care: a qualitative study. British Journal of General Practice 2007; 57(535):116-122.

9. Last JM. A Dictionary of Epidemiology. 3rd ed. Oxford: Oxford University Press; 1995.

10. Macrorie R. Tuberculosis in primary care. British Journal of General Practice 2007; 57(537):324-325.

11. Griffiths C, Martineau A. The new tuberculosis: raised awareness of tuberculosis is vital in general practice. British Journal of General Practice 2007; 57(535):94-95.

12. Singh S, Madge S, Lipman M. Tuberculosis in primary care. British Journal of General Practice 2002; 52(478):357-358.

13. Davies P. Tuberculosis: Current problems for primary care, a danger of complacency? Primary Care Respiratory Journal 2003; 12(4):105-106.

14. Brandli O. The Clinical Presentation of Tuberculosis. Respiration 1998; 65(2):97-105.

15. Zahar JR, Azoulay E, Klement E et al. Delayed treatment contributes to mortality in ICU patients with severe active pulmonary tuberculosis and acute respiratory failure. Intensive Care Medicine 2001; 27(3):513-520.

16. Bakhshi SS, Hawker J, Ali S. Tuberculosis mortality in notified cases from 1989-1995 in Birmingham. Public Health 1998; 112(3):165-168.

17. Chin DP, Crane CM, Diul MY et al. Spread of Mycobacterium tuberculosis in a Community Implementing Recommended Elements of Tuberculosis Control. JAMA 2000; 283(22):2968-2974.

18. Doherty MJ, Spence DP, Davies PD. Trends in mortality from tuberculosis in England and Wales: effect of age on deaths from non-respiratory disease. Thorax 1995; 50(9):976-979.

19. Rodger A, Jaffar S, Paynter S, Hayward A, Carless J, Maguire H. Delay in the diagnosis of pulmonary tuberculosis, London, 1998-2000: analysis of surveillance data. BMJ 2003; 326(7395):909-910.

20. Lewis KE, Stephens C, Shahidi MM, Packe G. Delay in starting treatment for tuberculosis in east London. Communicable Disease & Public Health 2003; 6(2):133-138.

21. Pirkis JE, Speed BR, Yung AP, Dunt DR, MacIntyre CR, Plant AJ. Time to initiation of anti-tuberculosis treatment. Tuber Lung Dis 1996; 77(5):401-406.

22. Rodrigo T. Proposing indicators for evaluation of tuberculosis control programmes in large cities based on the experience of Barcelona. International Journal of Tuberculosis and Lung Disease 2001; 5(5):432-440.

23. Sherman LF, Fujiwara PI, Cook SV, Bazerman LB, Frieden TR. Patient and health care system delays in the diagnosis and treatment of tuberculosis. Int J Tuberc Lung Dis 1999; 3(12):1088-1095.

24. Wares DF, Ormerod LP, Morton S. Delay in diagnosis of tuberculosis: of remaining concern in England and Wales. J Public Health Med 1999; 21(3):355-356.

25. Ward J, Siskind V, Konstantinos A. Patient and health care system delays in Queensland tuberculosis patients, 1985-1998. International Journal of Tuberculosis and Lung Disease 2001; 5(11):1021-1027.

26. Mathew TA, Ovsyanikova TN, Shin SS et al. Causes of death during tuberculosis treatment in Tomsk Oblast, Russia. International Journal of Tuberculosis and Lung Disease 2006; 10(8):857-863.

27. Tocque K, Convrey RP, Bellis MA, Beeching NJ, Davies PD. Elevated mortality following diagnosis with a treatable disease: tuberculosis. International Journal of Tuberculosis and Lung Disease 2005; 9(7):797-802.

28. Sacks LV, Pendle S. Factors related to in-hospital deaths in patients with tuberculosis. Arch Intern Med 1998; 158(17):1916-1922.

29. Franco J, Blanquer R. Mortality from tuberculosis in Spain from 1970 to 1993: changes in epidemiological trends during the acquired immune-deficiency syndrome epidemic. International Journal of Tuberculosis and Lung Disease 1998; 2(8):663-669.

30. Borgdorff MW, Veen J, Kalisvaart NA, Nagelkerke N. Mortality among tuberculosis patients in The Netherlands in the period 1993-1995. Eur Respir J 1998; 11(4):816-820.

31. Sterling TR, Zhao Z, Khan A et al. Mortality in a large tuberculosis treatment trial: modifiable and non-modifiable risk factors. International Journal of Tuberculosis and Lung Disease 2006; 10(5):542-549.

32. Fielder JF, Chaulk CP, Dalvi M, Gachuhi R, Comstock GW, Sterling TR. A high tuberculosis case-fatality rate in a setting of effective tuberculosis control: implications for acceptable treatment success rates. International Journal of Tuberculosis and Lung Disease 2002; 6(12):1114-1117.

33. Kourbatova EV, Borodulin BE, Borodulina EA, del RC, Blumberg HM, Leonard MK, Jr. Risk factors for mortality among adult patients with newly diagnosed tuberculosis in Samara, Russia. International Journal of Tuberculosis and Lung Disease 2006; 10(11):1224-1230.

34. Musher DM. How Contagious Are Common Respiratory Tract Infections? N Engl J Med 2003; 348(13):1256-1266.

35. Vynnycky E, Fine PEM. Lifetime Risks, Incubation Period, and Serial Interval of Tuberculosis. Am J Epidemiol 2000; 152(3):247-263.

36. Bloom BR, Murray CJ. Tuberculosis: commentary on a reemergent killer. Science 1992; 257(5073):1055-1064.

37. McKinney J.D., Jacobs W.R., Bloom B.R. Persisting problems in tuberculosis. In: Krause RM, Gallin J.I.Fauci A.S., editors. Emerging infections. New York: Academic Press; 1998 p. 51-146.

38. Jacob JT, Mehta AK, Leonard MK. Acute Forms of Tuberculosis in Adults. The American Journal of Medicine 2009; 122(1):12-17.

39. Parrish NM, Dick JD, Bishai WR. Mechanisms of latency in Mycobacterium tuberculosis. Trends in Microbiology 1998; 6(3):107-112.

40. Targeted Tuberculin Testing and Treatment of Latent Tuberculosis Infection. Joint Statement of the American Thoracic Society (ATS) and the Centers for Disease Control and Prevention (CDC). Am J Respir Crit Care Med 2000; 161(4):S221-247.

41. Targeted tuberculin testing and treatment of latent tuberculosis infection. American Thoracic Society. MMWR Recomm Rep 2000; 49(RR-6):1-51.

42. Lalvani A, Pathan AA, McShane H et al. Rapid detection of Mycobacterium tuberculosis infection by enumeration of antigen-specific T cells. Am J Respir Crit Care Med 2001; 163(4):824-828.

43. Frothingham R, Stout JE, Hamilton CD. Current issues in global tuberculosis control. International Journal of Infectious Diseases 2005; 9(6):297-311.

44. Grzybowski S, Barnett GD, Styblo K. Contacts of cases of active pulmonary tuberculosis. Bull Int Union Tuberc 1975; 50(1):90-106.

45. Shaw JB, Wynn-Williams N. Infectivity of pulmonary tuberculosis in relation to sputum status. Am Rev Tuberc 1954; 69(5):724-732.

46. Styblo K. Epidemiology of Tuberculosis. The Hague: Royal Netherlands Tuberculosis Association, 1991.

47. Health Protection Agency. Tuberculosis in the UK: Annual report on tuberculosis surveillance in the UK 2008. 2008. London, Health Protection Agency.

48. Davies PD. Clinical Tuberculosis. London: Arnold; 2003.

49. El-Sony A, Enarson D, Khamis A, Baraka O, Bjune G. Relation of grading of sputum smears with clinical features of tuberculosis patients in routine practice in Sudan. Int J Tuberc Lung Dis 2002; 6(2):91-97.

50. Cutler RR, Baithun SI, Doran HM, Wilson P. Association between the histological diagnosis of tuberculosis and microbiological findings. Tuber Lung Dis 1994; 75(1):75-79.

51. Wilcke JT, Kok-Jensen A. Diagnostic strategy for pulmonary tuberculosis in a low-incidence country: results of chest X-ray and sputum cultured for Mycobacterium tuberculosis. Respir Med 1997; 91(5):281-285.

52. Scott B, Schmid M, Nettleman MD. Early identification and isolation of inpatients at high risk for tuberculosis. Arch Intern Med 1994; 154(3):326-330.

53. Beck JS. Skin changes in the tuberculin test. Tubercle 1991; 72(2):81-87.

54. Ewer K, Deeks J, Alvarez L et al. Comparison of T-cell-based assay with tuberculin skin test for diagnosis of Mycobacterium tuberculosis infection in a school tuberculosis outbreak. The Lancet 2003; 361(9364):1168-1173.

55. Mitchison DA. The Diagnosis and Therapy of Tuberculosis During the Past 100 Years. Am J Respir Crit Care Med 2005; 171(7):699-706.

56. Ormerod LP, McCarthy OR, Rudd RM, Horsfield N. Short course chemotherapy for pulmonary tuberculosis. Respir Med 1991; 85(4):291-294.

57. A controlled trial of 6 months' chemotherapy in pulmonary tuberculosis. Final report: results during the 36 months after the end of chemotherapy and beyond. British Thoracic Society. Br J Dis Chest 1984; 78(4):330-336.

58. Joint Tuberculosis Committee of the British Thoracic Society. Control and prevention of tuberculosis in the United Kingdom: code of practice 2000. Thorax 2000; 55(11):887-901.

59. Story A, van Hest R, Hayward A. Tuberculosis and social exclusion. BMJ 2006; 333(7558):57-58.

60. Watson JM. Tuberculosis in Britain today. BMJ 1993; 306(6872):221-222.

61. Kumar D, Citron KM, Leese J, Watson JM. Tuberculosis among the homeless at a temporary shelter in London: report of a chest x ray screening programme. J Epidemiol Community Health 1995; 49(6):629-633.

62. Melzer M, Storring RA, Bagg LR. Tuberculosis in an area bordering east London: significant local variations when compared to national data. Infection 2000; 28(2):103-105.

63. Bothamley GH, Griffiths C, Beeks M, MacDonald M, Beasley E. Detecting tuberculosis in new arrivals to UK. Failure to register with a general practice compounds the problem. BMJ 2000; 321(7260):570.

64. Health Protection Agency. Annual report on tuberculosis cases reported in England, Wales and Northern Ireland in 2003. 2005. London, Health Protection Agency.

65. Kruijshaar ME, Watson JM, Drobniewski F et al. Increasing antituberculosis drug resistance in the United Kingdom: analysis of national surveillance data. BMJ 2008; 336(7655):1231-1234.

66. Rose AM, Sinka K, Watson JM, Mortimer JY, Charlett A. An estimate of the contribution of HIV infection to the recent rise in tuberculosis in England and Wales. Thorax 2002; 57(5):442-445.

67. Ahmed AB, Abubakar I, Delpech V et al. The growing impact of HIV infection on the epidemiology of tuberculosis in England and Wales: 1999 2003. Thorax 2007; 62(8):672-676.

68. Dye C. Global epidemiology of tuberculosis. The Lancet 2006; 367(9514):938-940.

69. Corbett EL, Watt CJ, Walker N et al. The Growing Burden of Tuberculosis: Global Trends and Interactions With the HIV Epidemic. Arch Intern Med 2003; 163(9):1009-1021.

70. World Health Organization. Global tuberculosis control: Epidemiology, strategy, financing: WHO report 2009. 2009. Geneva, Switzerland, WHO.
Ref Type: Report

71. Hayward AC, Watson JM. Tuberculosis in England and Wales 1982-93: notifications exceeded predictions. Commun Dis Rep 1995; 5:R29-R33.

72. Watson JM, Maguire HC. PHLS work on the surveillance and epidemiology of tuberculosis. Commun Dis Rep CDR Rev 1997; 7(8):R110-R112.

73. Ormerod LP, Charlett A, Gilham C, Darbyshire JH, Watson JM. Geographical distribution of tuberculosis notifications in national surveys of England and Wales in 1988 and 1993: report of the Public Health Laboratory Service/British Thoracic Society/Department of Health Collaborative Group. Thorax 1998; 53(3):176-181.

74. Davies PD. The challenge of tuberculosis. J R Soc Med 2003; 96(6):262-265.

75. Lawson DH, Sherman V, Hollowell J. The General Practice Research Database. QJM 1998; 91(6):445-452.

76. Walley T, Mantgani A. The UK General Practice Research Database. The Lancet 1997; 350(9084):1097-1099.

77. General Practice Research Database: much more than the database. http://www.gprd.com/home/. 27-8-2009. Ref Type: Internet Communication

78. Petersen I, Johnson AM, Islam A, Duckworth G, Livermore DM, Hayward AC. Protective effect of antibiotics against serious complications of common respiratory tract infections: retrospective cohort study with the UK General Practice Research Database. BMJ 2007; 335(7627):982-987.

79. Walters K, Rait G, Petersen I, Williams R, Nazareth I. Panic disorder and risk of new onset coronary heart disease, acute myocardial infarction, and cardiac mortality: cohort study using the general practice research database. Eur Heart J 2008; 29(24):2981-2988.

80. Hansell A, Hollowell J, Nichols T, McNiece R, Strachan D. Use of the General Practice Research Database (GPRD) for respiratory epidemiology: a comparison with the 4th Morbidity Survey in General Practice (MSGP4). Thorax 1999; 54(5):413-419.

81. Jones R, Latinovic R, Charlton J, Gulliford MC. Alarm symptoms in early diagnosis of cancer in primary care: cohort study using General Practice Research Database. BMJ 2007;334(7602):1040-1047.

82. Osborn DPJ, Levy G, Nazareth I, Petersen I, Islam A, King MB. Relative Risk of Cardiovascular and Cancer Mortality in People With Severe Mental Illness From the United Kingdom's General Practice Research Database. Arch Gen Psychiatry 2007; 64(2):242-249.

83. Meier CR, Jick SS, Derby LE et al. Acute respiratory-tract infections and risk of first-time acute myocardial infarction. The Lancet 1998; 351(9114):1467-1471.

84. Jick H, Jick SS, Derby LE. Validation of information recorded on general practitioner based computerised data resource in the United Kingdom. BMJ: British Medical Journal 1991; 302(6779):766-768.

85. Hollowell J. The General Practice Research Database: quality of morbidity data. Pop Trend 1997; 87:36-40.

86. Jick H, Terris BZ, Derby LE, Jick SS. Further validation of information recorded on a general practitioner based computerized data resource in the united kingdom. Pharmacoepidemiology & Drug Safety 1992; 1:347-349.

87. Nazareth I, King M, Haines A, Rangel L, Myers S. Accuracy of diagnosis of psychosis on general practice computer system. BMJ: British Medical Journal 1993; 307(6895):32-34.

88. Nazareth I, King M, Haines A, Tai SS, Hall G. Care of schizophrenia in general practice. BMJ: British Medical Journal 1993; 307(6909):910.

89. McCormick A, Fleming D, Charlton J. Morbidity Statistics from General Practice. Fourth National Study 1991-1992. 1995. London, HMSO.

90. Soriano JB, Kiri VA, Maier WC, Strachan D. Increasing prevalence of asthma in UK primary care during the 1990s. The International Journal of Tuberculosis and Lung Disease 2003; 7:415-421.

91. Winchester CC, Macfarlane TV, Thomas M, Price D. Antibiotic Prescribing and Outcomes of Lower Respiratory Tract Infection in UK Primary Care. Chest 2009; 135(5):1163-1172.

92. Pitman RJ, Melegaro A, Gelb D, Siddiqui MR, Gay NJ, Edmunds WJ. Assessing the burden of influenza and other respiratory infections in England and Wales. Journal of Infection 2007; 54(6):530-538.

93. Jick H. A database worth saving. The Lancet 1997; 350(9084):1045-1046.

94. Pablos-Mendez A, Sterling TR, Frieden TR. The relationship between delayed or incomplete treatment and all-cause mortality in patient with tuberculosis. JAMA 1996; 276:1223-1228.

95. Golub JE, Bur S, Cronin WA et al. Delayed tuberculosis diagnosis and tuberculosis transmission. Int J Tuberc Lung Dis 2006; 10(1):24-30.

96. Richardson WS, Wilson MC, Nishikawa J, Hayward RS. The well-built clinical question: a key to evidence-based decisions. ACP J Club 1995; 123(3):A12-A13.

97. Oxman AD, Sackett DL, Guyatt GH. Users' guides to the medical literature. I. How to get started. The Evidence-Based Medicine Working Group. JAMA 1993; 270(17):2093-2095.

98. Scottish Intercollegiate Guidelines Network (SIGN). Methodology Review Group. Report on the review of the method of grading guideline recommendations. 1999. Edinburgh, SIGN.

99. Harbour R, Miller J. A new system for grading recommendations in evidence based guidelines. BMJ 2001; 323(7308):334-336.

100. Griffiths C, Sturdy P, Brewin P et al. Educational outreach to promote screening for tuberculosis in primary care: a cluster randomised controlled trial. Lancet 2007; 369(9572):1528-1534.

101. Singh S. What about GPs in the fight against tuberculosis? Journal of the Royal Society of Medicine 2003; 96(8).

102. White VLC, Paliwalla M, Steves CJ, Jadhav D, Moore-Gillon J. Management of tuberculosis in a British inner-city population. J Public Health 2002; 24(1):49-52.

103. Macrorie R, Cordell A, Hamlet N. Tuberculosis in primary care. British Journal of General Practice 2002; 52(481):674-675.

104. Margolis PA, Lannon CM, Stuart JM, Fried BJ, Keyes-Elstein L, Moore J. Practice based education to improve delivery systems for prevention in primary care: Randomised trial. British Medical Journal 2004; 328(7436):388-392.

105. Bothamley GH, Rowan JP, Griffiths CJ et al. Screening for tuberculosis: The port of arrival scheme compared with screening in general practice and the homeless. Thorax 2002; 57(1):45-49.

106. Van den Bosch CA, Roberts JA. Tuberculosis screening of new entrants; How can it be made more effective? Journal of Public Health Medicine 2000; 22(2):220-223.

107. Hargreaves S, Holmes A, Friedland JS. Refugees, asylum seekers, and general practice: room for improvement?[see comment]. British Journal of General Practice 2000; 50(456):531-532.

108. Levin AC, Gums JG, Grauer K. Tuberculosis: The primary care physician's role in eradication. Postgraduate Medicine 1993; 93(3):46-58.

109. Paynter S, Hayward A, Wilkinson P, Lozewicz S, Coker R. Patient and health service delays in initiating treatment for patients with pulmonary tuberculosis: retrospective cohort study. Int J Tuberc Lung Dis 2004; 8(2):180-185.

110. Asch S, Leake B, Anderson R, Gelberg L. Why Do Symptomatic Patients Delay Obtaining Care for Tuberculosis? Am J Respir Crit Care Med 1998; 157(4):1244-1248.

111. Diez M, Bleda MJ, Alcaide J et al. Determinants of health system delay among confirmed tuberculosis cases in Spain. European Journal of Public Health 2005; 15(4):343-349.

112. Storla DG, Yimer S, Bjune GA. A systematic review of delay in the diagnosis and treatment of tuberculosis. BMC Public Health 2008; 8:15.

113. Golub JE. Impact of empiric antibiotics and chest radiograph on delays in the diagnosis of tuberculosis. International Journal of Tuberculosis and Lung Disease 2005; 9(4):392-397.

114. Golub JE. Patient and health care system delays in pulmonary tuberculosis diagnosis in a low-incidence state. International Journal of Tuberculosis and Lung Disease 2005; 9(9):992-998.

115. Altet-Gomez MN, Alcaide MJ, Canela SJ et al. Pulmonary symptomatic tuberculosis' diagnostic delay study. Arch Bronconeumol 2003; 39(4):146-152.

116. Ohmori M, Ozasa K, Mori T et al. Trends of delays in tuberculosis case finding in Japan and associated factors. International Journal of Tuberculosis and Lung Disease 2005; 9(9):999-1005.

117. Franco J, Blanquer R, Flores J, Fernandez E, Plaza P, Nogueira JM. [Analysis of the diagnostic delay in tuberculosis]. Med Clin (Barc) 1996; 107(12):453-457.

118. Farah MG, Rygh JH, Steen TW, Selmer R, Heldal E, Bjune G. Patient and health care system delays in the start of tuberculosis treatment in Norway. BMC Infect Dis 2006; 6:33.

119. Gagliotti C, Resi D, Moro ML. Delay in the treatment of pulmonary TB in a changing demographic scenario. International Journal of Tuberculosis and Lung Disease 2006; 10(3):305-309.

120. Gulbaran Z, Pretet S, Dusser D. [From first symptom to diagnosis and from diagnosis to treatment of tuberculosis: still a long delay]. Rev Pneumol Clin 1996; 52(1):20-25.

121. Sanz B, Blasco T. Variables associated with diagnostic delay in immigrant groups with tuberculosis in Madrid. International Journal of Tuberculosis and Lung Disease 2007; 11(6):639-646.

122. Greenaway C, Menzies D, Fanning A et al. Delay in Diagnosis among Hospitalized Patients with Active Tuberculosis-Predictors and Outcomes. Am J Respir Crit Care Med 2002; 165(7):927-933.

123. Rao VK, Iademarco EP, Fraser VJ, Kollef MH. The impact of comorbidity on mortality following in-hospital diagnosis of tuberculosis. Chest 1998; 114(5):1244-1252.

124. Kam A, Ford-Jones L, Malloy P, Khan K, Kitai I. Active tuberculosis among adolescents in Toronto, Canada: Clinical features and delays in diagnosis. The Pediatric Infectious Disease Journal 2007; 26(4):355-356.

125. Ho M.-J. Health-seeking patterns among Chinese immigrant patients enrolled in the directly observed therapy program in New York City. The International Journal of Tuberculosis and Lung Disease 2004; 8(11):1355-1359.

126. Moudgil H, Leitch AG. Extra-pulmonary tuberculosis in Lothian 1980-1989: ethnic status and delay from onset of symptoms to diagnosis. Respiratory Medicine 1994; 88(7):507-510.

127. Diez M, Bleda MJ, Alcaide J et al. Determinants of patient delay among tuberculosis cases in Spain. Eur J Public Health 2004; 14(2):151-155.

128. Szczepura A. Access to health care for ethnic minority populations. Postgraduate Medical Journal 2005; 81(953):141-147.

129. Dooley K, Golub J, Goes F, Merz W, Sterling T. Empiric treatment of community–acquired pneumonia with fluoroquinolones, and delays in the treatment of Tuberculosis. Clinical Infectious Diseases 2002; 34(12):1607-1612.

130. Yoon YS, Lee HJ, Yoon HI et al. Impact of fluoroquinolones on the diagnosis of pulmonary tuberculosis initially treated as bacterial pneumonia. The International Journal of Tuberculosis and Lung Disease 2005; 9:1215-1219.

131. Ang D, Hsu AA, Tan BH. Fluoroquinolones may delay the diagnosis of tuberculosis. Singapore Med J 2006; 47(9):747-751.

132. Wang JY, Hsueh PR, Jan IS et al. Empirical treatment with a fluoroquinolone delays the treatment for tuberculosis and is associated with a poor prognosis in endemic areas. Thorax 2006; 61(10):903-908.

133. Story A, Murad S, Roberts W, Verheyen M, Hayward AC, for the London Tuberculosis Nurses Network. Tuberculosis in London: the importance of homelessness, problem drug use and prison. Thorax 2007; 62(8):667-671.

134. García Rodríguez L.A. PGS. Use of the UK General Practice Research Database for pharmacoepidemiology. British Journal of Clinical Pharmacology 1998; 45(5):419-425.

135. Perry J. OXMIS problem codes for primary medical care. Oxford: OXMIS Publications; 1978.

136. Chisholm J. The Read clinical classification. BMJ: British Medical Journal 1990; 300(6732):1092.

137. Gail M, Williams R, Byar DP, Brown C. How many controls? Journal of Chronic Diseases 1976; 29(11):723-731.

138. Rothman KJ, Greenland S. Modern Epidemiology. 2nd edition ed. Philadelphia: Lipincott-Raven; 1998.

139. Woodward M. Epidemiology: study design and data analysis. 2nd ed. London: Chapman and Hall; 2005.

140. Smeeth L, Donnan PT, Cook DG. The use of primary care databases: case-control and case-only designs. Fam Pract 2006; 23(5):597-604.

141. Maclure M. The Case-Crossover Design: A Method for Studying Transient Effects on the Risk of Acute Events. Am J Epidemiol 1991; 133(2):144-153.

142. Farrington CP. Control without separate controls: evaluation of vaccine safety using case-only methods. Vaccine 2004; 22(15-16):2064-2070.

143. Maclure M, Mittleman aMA. Should We Use a Case-Crossover Design? Annual Review of Public Health 2000; 21(1):193-221.

144. Douglas AS, Ali S, Bakhshi SS. Does vitamin D deficiency account for ethnic differences in tuberculosis seasonality in the UK? Ethn Health 1998; 3(4):247-253.

145. Douglas AS, Strachan DP, Maxwell JD. Seasonality of tuberculosis: the reverse of other respiratory diseases in the UK.[see comment]. Thorax 1996; 51(9):944-946.

146. Schaaf HS, Nel ED, Beyers N, Gie RP, Scott F, Donald PR. A decade of experience with Mycobacterium tuberculosis culture from children: a seasonal influence on incidence of childhood tuberculosis. Tuber Lung Dis 1996; 77(1):43-46.

147. Zeger SL, Liang K-Y, Albert PS. Models for Longitudinal Data: A Generalized Estimating Equation Approach. Biometrics 1988; 44(4):1049-1060.

148. Omar RZ, Wright EM, Turner RM, Thompson SG. Analysing repeated measurements data: a practical comparison of methods. Statistics in Medicine 1999; 18:1587-1603.

149. Farrington CP. Relative Incidence Estimation from Case Series for Vaccine Safety Evaluation. Biometrics 1995; 51(1):228-235.

150. Farrington CP, Nash J, Miller E. Case Series Analysis of Adverse Reactions to Vaccines: A Comparative Evaluation. Am J Epidemiol 1996; 143(11):1165-1173.

151. Smeeth L, Thomas SL, Hall AJ, Hubbard R, Farrington P, Vallance P. Risk of Myocardial Infarction and Stroke after Acute Infection or Vaccination. N Engl J Med 2004; 351(25):2611-2618.

152. Whitaker HJ, Farrington CP, Spiessens B, Musonda P. Tutorial in biostatistics: the self-controlled case series method. Statistics in Medicine 2006; 25(10):1768-1797.

153. Douglas IJ, Smeeth L. Exposure to antipsychotics and risk of stroke: self controlled case series study. BMJ 2008; 337(aug28_2):a1227.

154. Kirkwood BR. Essentials of Medical Statistics. Oxford: Blackwell; 1998.

155. StataCorp. Stata Statistical Software: Release 9 (2005) College Station. Texas: StataCorp LP: 2005.

156. Craig SE, Bettinson H, Sabin CA, Gillespie SH, Lipman MC. Think TB! Is the diagnosis of pulmonary tuberculosis delayed by the use of antibiotics? International Journal of Tuberculosis and Lung Disease 2009; 13(2):208-213.

157. Rodger AJ, Story A, Fox Z, Hayward A. HIV prevalence and testing practices among tuberculosis cases in London: a missed opportunity for HIV diagnosis? Thorax 2010; 65(1):63-69.

158. Diwan VK, Thorson A. Sex, gender, and tuberculosis. The Lancet 1999; 353(9157):1000-1001.

159. Holmes CB, Hausler H, Nunn P. A review of sex differences in the epidemiology of tuberculosis. The international journal of tuberculosis and lung disease 1998; 2:96-104.

160. Stevenson CR, Critchley JA, Forouhi NG et al. Diabetes and the risk of tuberculosis: a neglected threat to public health? Chronic Illn 2007; 3(3):228-245.

161. Walker C, Unwin N. Estimates of the impact of diabetes on the incidence of pulmonary tuberculosis in different ethnic groups in England. Thorax 2010; 65:578-581.

162. Kim HR, Hwang SS, Ro Y.K., et al. Solid-organ malignancy as a risk factor for tuberculosis. Respirology 2008; 13(3):413-419.

163. Engels EA, Shen M, Chapman RS et al. Tuberculosis and subsequent risk of lung cancer in Xuanwei, China. International Journal of Cancer 2009; 124(5):1183-1187.

164. Kamboj M, Sepkowitz K. The Risk of Tuberculosis in Patients with Cancer. Clinical Infectious Diseases 2006; 42(11):1592-1595.

165. Hnizdo E, Singh T, Churchyard G. Chronic pulmonary function impairment caused by initial and recurrent pulmonary tuberculosis following treatment. Thorax 2000; 55(1):32-38.

166. Pasipanodya JG, Miller TL, Vecino M et al. Pulmonary Impairment After Tuberculosis. Chest 2007; 131(6):1817-1824.

167. Davies PDO, Yew WW, Ganguly D et al. Smoking and tuberculosis: the epidemiological association and immunopathogenesis. Transactions of the Royal Society of Tropical Medicine and Hygiene 2006; 100(4):291-298.

168. Aubry MC, Wright JL, Myers JL. The pathology of smoking-related lung diseases. Clin Chest Med 2000; 21(1):11-35.

169. Bates MN, Khalakdina A, Pai M, Chang L, Lessa F, Smith KR. Risk of tuberculosis from exposure to tobacco smoke: a systematic review and meta-analysis. Arch Intern Med 2007; 167(4):335-342.

170. Lin HH, Ezzati M, Murray M. Tobacco smoke, indoor air pollution and tuberculosis: a systematic review and meta-analysis. PLoS Med 2007; 4(1):e20.

171. Anderson C, Abubakar I, Maguire H, Sonnenberg P. Survey of tuberculosis incidents in hospital healthcare workers, England and Wales, 2005. J Public Health 2007; 29(3):292-297.

172. R Baker. Will the future GP remain a personal doctor? Br J Gen Pract 1997; 47(425):831-833.

173. Kruijshaar ME, Abubakar I. Increase in extrapulmonary tuberculosis in England and Wales 1999 - 2006. Thorax 2009; 64(12):1090-1095.

174. Forssbohm M, Zwahlen M, Loddenkemper R, Rieder HL. Demographic characteristics of patients with extrapulmonary tuberculosis in Germany. Eur Respir J 2008; 31(1):99-105.

175. Ong A, Creasman J, Hopewell P et al. A Molecular Epidemiological Assessment of Extrapulmonary Tuberculosis in San Francisco. Clinical Infectious Diseases 2004; 38(1):25-31.

176. Gonzalez OY, Adams G, Teeter LD, Bui TT, Musser JM, Graviss EA. Extra-pulmonary manifestations in a large metropolitan area with a low incidence of tuberculosis. Int J Tuberc Lung Dis 2003; 7(12):1178-1185.

177. McGovern MP, Boroujerdi MA, Taylor MW et al. The effect of the UK incentive-based contract on the management of patients with coronary heart disease in primary care. Fam Pract 2008; 25(1):33-39.

178. Long R, Maycher B, Dhar A, Manfreda J, Hershfield E, Anthonisen N. Pulmonary Tuberculosis Treated With Directly Observed Therapy. Chest 1998; 113(4):933-943.

179. Willcox PA, Ferguson AD. Chronic obstructive airways disease following treated pulmonary tuberculosis. Respiratory Medicine 1989; 83(3):195-198.

180. Crofts JP, Pebody R, Grant A, Watson JM, Abubakar I. Estimating tuberculosis case mortality in England and Wales, 2001-2002. International Journal of Tuberculosis and Lung Disease 2008; 12(3):308-313.

181. Heldal E, Naalsund A, Kongerud J, Tverdal A, Boe J. Deaths from active tuberculosis: can we rely on notification and mortality figures? Tuber Lung Dis 1996; 77(3):215-219.

182. Kircher T, Nelson J, Burdo H. The autopsy as a measure of accuracy of the death certificate. N Engl J Med 1985; 313(20):1263-1269.

183. Teale C, Goldman JM, Pearson SB. The association of age with the presentation and outcome of tuberculosis: a five-year survey. Age Ageing 1993; 22(4):289-293.

184. Tan KL, Sin Fai Lam KN, Chew LS. Mortality of patients while on treatment for active tuberculosis. Singapore Med J 1996; 37(3):258-260.

185. World Health Organization. Global tuberculosis control: surveillance, planning, financing. WHO Report, 2006. 2006. Geneva, Switzerland, WHO.

186. Ditah IC, Reacher M, Palmer C et al. Monitoring tuberculosis treatment outcome: analysis of national surveillance data from a clinical perspective. Thorax 2008; 63(5):440-446.

187. Martineau AR, Lowey H, Tocque K, Davies PD. Decreasing tuberculosis case fatality in England and Wales, 1988-2001. International Journal of Tuberculosis and Lung Disease 2004; 8(6):737-742.

188. World Health Organization. Global Tuberculosis Control: WHO Report, 2001. WHO/TB/2001.287. 2001. Geneva, Switzerland: WHO.

189. Khan K, Campbell A, Wallington T, Gardam M. The impact of physician training and experience on the survival of patients with active tuberculosis. CMAJ 2006; 175(7):749-753.

190. Rabaud C, Engozogho M, Dailloux M, Hoen B, May T, Canton P. Tuberculosis in Lorraine, France: study of prognostic factors. International Journal of Tuberculosis and Lung Disease 1997; 1(3):246-249.

191. Reichman LB. Tuberculosis elimination--what's to stop us? International Journal of Tuberculosis and Lung Disease 1997; 1(1):3-11.

192. Nardell EA. Needles in haystacks: Diagnosing tuberculosis under low prevalence conditions. Tubercle and Lung Disease 1996; 77(5):389-390.

193. CDC. Updated guidelines for using Interferon Gamma Release Assays to detect Mycobacterium tuberculosis infection. MMWR 2010; 59(No. RR-5):1-25.