HIV and tuberculosis co-infection in England, Wales and Northern Ireland:

Prevalence, risk factors and transmission

Joanne Rebecca Winter, MSc

University College London

Institute for Global Health

PhD Thesis

November 2017

Declaration

I, Joanne Winter, 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

In high-income countries, HIV and tuberculosis are concentrated in hard-to-reach populations. The epidemiology of HIV and tuberculosis co-infection has not been recently described in the UK, and the role of HIV in tuberculosis transmission in low-burden settings is unclear.

A systematic review of risk factors for latent tuberculosis infection and active tuberculosis disease was undertaken. The national surveillance datasets for HIV and tuberculosis were linked, and these datasets were used to investigate risk factors for developing tuberculosis for people living with HIV, and to describe trends in HIV co-infection among tuberculosis patients. Strain typing data on *Mycobacterium tuberculosis* complex isolates from tuberculosis patients were used to examine the role of HIV in tuberculosis transmission.

In England, Wales and Northern Ireland, 6.0% of people diagnosed with HIV between 2000 and 2014 had a tuberculosis diagnosis during this time period, and 5.4% of tuberculosis patients were co-infected with HIV. The number and proportion of tuberculosis patients co-infected with HIV declined from 2005 to 2014. The strongest risk factors for tuberculosis among people with HIV were black African ethnicity, birth in a country with high tuberculosis incidence, and HIV acquisition through injecting drug use. High CD4 count and initiating anti-retroviral therapy were both highly protective against tuberculosis. Among tuberculosis patients, drug misuse was the only social risk factor associated with HIV co-infection. Tuberculosis patients with HIV had fewer subsequent clustered cases than HIV-negative tuberculosis patients, and tuberculosis patients with HIV were more often the result of reactivation of latent tuberculosis than recent infection.

Co-infection with tuberculosis and HIV has declined, but further reductions are necessary. Increasing screening for HIV and latent tuberculosis in high-risk populations such as people of black African ethnicity, people born in high-incidence countries, and people who inject drugs, could reduce tuberculosis in people living with HIV.

Table of Contents

List of Figures

List of Tables

Abbreviations

AIDS	Acquired immunodeficiency syndrome
AKA	Also known as
ART	Anti-retroviral therapy
BCG	Bacillus Calmette-Guérin
BHIVA	The British HIV Association
CDC	Centers for disease control and prevention
CI	Confidence interval
DNA	Deoxyribonucleic acid
DTH	Delayed-type hypersensitivity
ECDC	European Centre for Disease Surveillance and Control
EEA	European Economic Area
EMS	Enhanced matching system
ESMI	Enhanced surveillance of mycobacterial infections
ETS	Enhanced tuberculosis surveillance
EU	European Union
HANDD	HIV and AIDS new diagnoses database
HARS	HIV and AIDS reporting system
HIV	Human immunodeficiency virus
IGRA	Interferon-gamma release assay
IMD	Index of multiple deprivation
IQR	Inter-quartile range
IRIS	Immune reconstitution inflammatory syndrome
IRR	Incidence rate ratio
LTBI	Latent tuberculosis infection
M. tb	*Mycobacterium tuberculosis*
MIRU-VNTR	Mycobacterial interspersed repetitive units – variable number tandem repeats
MSM	Men who have sex with men
MTBC	*Mycobacterium tuberculosis* complex
NHS	National Health Service
NICE	The National Institute for Health and Care Excellence
OR	Odds ratio
PHE	Public Health England
PII	Personally identifying information
PLHIV	People living with HIV
PWID	People who inject drugs

PY	Person-years
QFT-GIT	QuantiFERON-TB Gold-in-Tube
RFLP	Restriction fragment length polymorphism
RNA	Ribonucleic acid
SNP	Single nucleotide polymorphism
SOPHID	Survey of prevalent HIV infections diagnosed
TB	Tuberculosis
TESSy	The European Surveillance System
TST	Tuberculin skin test
UK	United Kingdom
UNAIDS	The Joint United Nations Programme on HIV/AIDS
USA	United States of America
WGS	Whole-genome sequencing
WHO	World Health Organization

Publications and Presentations

One peer-reviewed journal article has been published from the work done as part of this PhD:

1. Winter JR, Stagg HR, Smith CJ, Brown AE, Lalor MK, Lipman M, Pozniak A, Skingsley A, Kirwan P, Yin Z, Thomas HL, Delpech V, Abubakar I. **Injecting drug use predicts active tuberculosis in a national cohort of people living with HIV**. AIDS. 2017;**31**(17):2403-2413. doi: 10.1097/QAD.0000000000001635. **(Chapter 4)**

Two other journal articles with the results from **Chapter 2** and **Chapter 5** are currently under peer review.

Two abstracts were accepted for oral presentations:

1. Winter JR, Brown J, Stagg HR, Lalor MK, Delpech V, Lipman M, Abubakar I. **Changing diagnostic pattern of HIV and tuberculosis co-infection in England, Wales and Northern Ireland, 2000-2014**. 2017. *British Thoracic Society Winter Meeting, London, UK.* **(Chapter 5)**
2. Winter JR, Stagg HR, Smith CJ, Lalor MK, Davidson A, Delpech V, Abubakar I. **The role of HIV co-infection in tuberculosis transmission in the low-incidence setting of England, Wales and Northern Ireland, 2010-2014**. 2017. *The Union World Conference on Lung Health, Guadalajara, Mexico.* **(Chapter 6)**

Three abstracts were accepted for poster presentations:

1. Winter JR, Stagg HR, Smith CJ, Brown AE, Lalor MK, Thomas HL, Delpech V, Abubakar I. **Social risk factors associated with HIV infection in UK tuberculosis cases, 2010-2014**. 2017. *Conference on Retroviruses and Opportunistic Infections, Seattle, USA.* **(Chapter 5)**
2. Winter JR, Stagg HR, Smith C, Lalor MK, Skingsley A, Thomas HL, Abubakar I, Delpech V. **Risk factors for developing tuberculosis among the national HIV-positive cohort over 15 years in England, Wales and Northern Ireland**. 2016. *The Union World Conference on Lung Health, Liverpool, UK.* **(Chapter 4)**
3. Winter, JR, Delpech V, Kirwan P, Stagg HR, Venugopalan S, Skingsley A, Lalor MK, Thomas, HL, Abubakar, I. **Linkage of UK HIV and tuberculosis data using probabilistic and deterministic methods**. 2016. *Conference on Retroviruses and Opportunistic Infections, Boston, USA.* **(Chapter 3)**

Acknowledgements

Thank you first and foremost to my supervisors, Professor Ibrahim Abubakar, Dr Valerie Delpech, Dr Helen Stagg and Dr Colette Smith, for all their support, encouragement and constructive criticism throughout the course of this PhD. None of this would have been possible without you.

Thank you also to everyone in the tuberculosis and HIV sections at Public Health England for all your work collecting the data I used. Particular thanks to Maeve Lalor, Peter Kirwan, Zheng Yin and Jennifer Davidson for answering my endless questions! I am enormously grateful to everyone who worked with me on the matching algorithm. The deterministic algorithm in Chapter 3 (the hierarchy of matching criteria, additional rules and final hybrid method) was developed by discussion with data managers and epidemiologists at Public Health England (Peter Kirwan, David Quinn, Srividhya Venugopalan, Andrew Skingsley, Alison Brown, Valerie Delpech, Ibrahim Abubakar, Lucy Thomas, Maeve Lalor, Zheng Yin and Jennifer Davidson). The Stata code used in the deterministic and hybrid methods was written by Peter Kirwan and myself. The original matched dataset which was used for comparison in Chapter 3 was compiled by Kunju Shaji and Zheng Yin in 2012. My sincere thanks to Aishatu Adamu for screening 50% of the search results for the systematic review in Chapter 2.

Thank you to the UCL IMPACT scheme and Public Health England for financial support.

I would also like to thank everyone who has shared their knowledge with me throughout this PhD, in particular Marc Lipman, Anton Pozniak and Dominik Zenner. Thank you Hollie, for sharing the journey! Finally, I am grateful to my friends and family for their support – in particular Libby, for making my life easier, and John, for encouraging me to make life difficult.

Ethical Approval and Patient Consent

The work in Chapter 2 was a review of previously published data, and therefore ethical approval was not required. The work presented in Chapters 3-6 was approved by the UCL student Research Ethics Committee (5683/001). Patient consent was not required because Public Health England has authority under the Health and Social Care Act 2012 to hold and analyse national surveillance data for public health and research purposes.

1 General Introduction

Tuberculosis (TB) and disease caused by the human immunodeficiency virus (HIV) are two of the leading infectious causes of death worldwide.[1-3] There is considerable geographic and sociodemographic overlap between the two epidemics; 11% of all new TB cases in 2015 were in people living with HIV (PLHIV). The two diseases have a synergistic relationship, resulting in greater morbidity and mortality among co-infected patients. This thesis aims to improve our understanding of the epidemiology of TB-HIV co-infection in a low-incidence setting, and the role of HIV in TB transmission, in England, Wales and Northern Ireland. The introduction outlines the burden and epidemiology of HIV, TB and TB-HIV, globally and specifically in low-incidence settings. I describe the transmission, pathogenesis and treatment of HIV and TB, and discuss the tools that have been developed to study and better understand TB transmission. The introduction concludes with the objectives for my thesis.

1.1 HIV infection

HIV is a retrovirus which targets the human immune system, impairing immune function and causing the condition acquired immunodeficiency syndrome (AIDS). Infection with HIV is lifelong.

1.1.1 HIV pathogenesis and disease

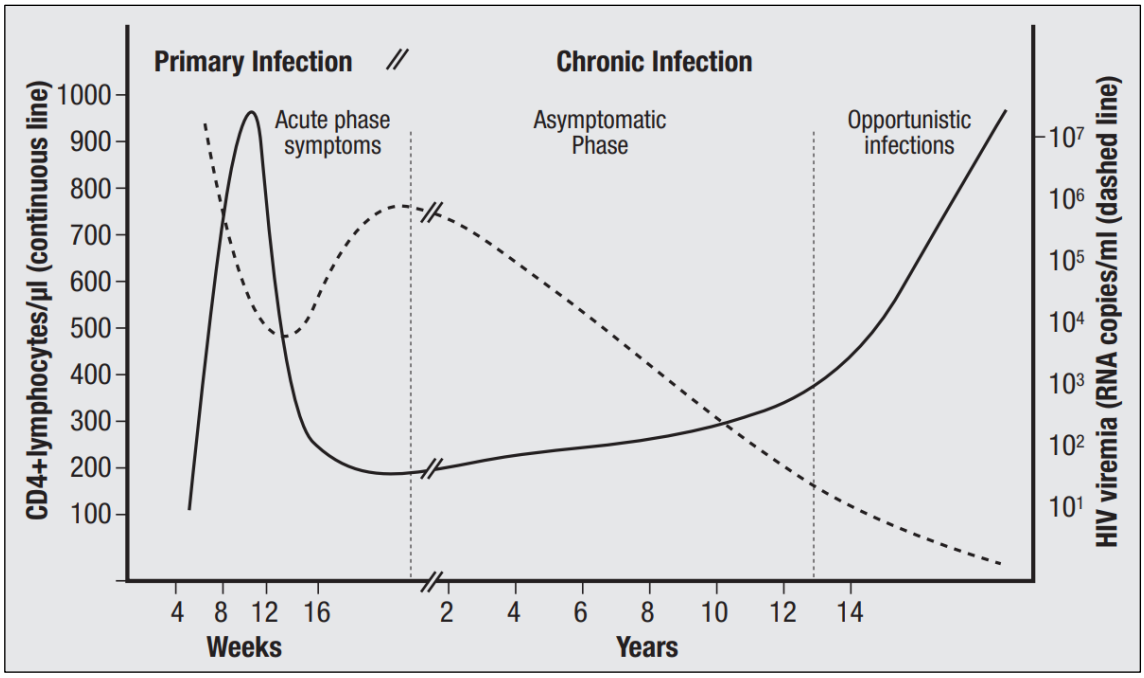
HIV infects cells of the immune system, particularly $CD4^+$ T helper cells.[4, 5] This ultimately leads to the depletion of $CD4^+$ cells, predominantly through pyroptosis (an inflammatory form of programmed cell death),[6] but also through apoptosis, viral killing of infected cells, or the killing of infected cells by cytotoxic $CD8^+$ lymphocytes. $CD4^+$ depletion results in the loss of cell-mediated immunity, leading to greater susceptibility to opportunistic infections and cancers.

HIV is a retrovirus, transmitted as two copies of a positive-sense single strand of ribonucleic acid (RNA). The virus particle enters the host cell by binding to the CD4

receptor on the host membrane and fusing with the cell membrane. After uncoating, the viral genome is reverse transcribed into double-stranded deoxyribonucleic acid (DNA) and then integrated into the host cell's genome. Viral replication occurs within the nucleus of the host cell; viral RNA is transcribed and exported from the nucleus before being translated into new viral proteins. These are then assembled into viral particles which bud off from the cell membrane and then mature.[7]

There are a number of stages to HIV infection. Acute HIV infection occurs within 2-4 weeks of infection and is frequently asymptomatic, or accompanied by influenza-like illness such as fever, headache, rash or sore throat.[5] This phase is characterized immunologically by a sharp drop in CD4 count and a high viral load (Figure 1.1), resulting in widespread dissemination of the virus. Following infection, rapid viral replication leads to high viral loads and substantial CD4 depletion. This activates $CD8^+$ T cells, which kill HIV-infected cells. The production of antibodies, known as seroconversion, occurs 3-12 weeks after infection.[8]

Figure 1.1: The relationship between CD4 count and viral load during the clinical course of HIV infection.



Source: Fanales-Belasio et al. [9]

Seroconversion is followed by a clinically latent phase of infection, where HIV replicates at very low levels.[5] This phase can last for months to years, and can be

prolonged by the use of anti-retroviral drugs which limit viral replication and maintain CD4 count. Other symptoms follow as the immune system weakens, viral load increases and CD4 count falls. These may include swollen lymph nodes, weight loss, diarrhea, cough and fever.[3]

AIDS is the most severe clinical stage of HIV infection; defined by severely impaired immunity (CD4 count <200 cells/µl) or AIDS indicator diseases, which are usually opportunistic infections, certain cancers, or other severe clinical manifestations. These typically include severe infections such as pneumonia, TB, oesophageal candidiasis, cryptococcal meningitis, or cancers such as Kaposi's sarcoma and lymphomas.[3, 10] The term AIDS has, however, fallen out of fashion as we have come to understand immune impairment and the associated morbidities as a sliding scale rather than an all or nothing event.

1.1.2 HIV diagnosis and treatment

HIV infection can be diagnosed by detecting HIV-specific antibodies in serum or plasma, or by detecting the presence of the HIV p24 antigen or viral nucleic acids. Antibody tests can only detect HIV after seroconversion, once an immune response to HIV has occurred; usually within 3-12 weeks of infection (this has become quicker over time). Antibodies therefore may not be detectable if testing is done before the patient has seroconverted.[3] In contrast, antigens are present in high quantities in the early weeks following infection, and then stop being detectable.[11] Nucleic acid tests can detect virus as soon as 7 days after infection, for as long as the virus is present in the blood, however as they are expensive they are generally only used for patients with a recent high-risk event.[12]

Disease stage is also classified based on combinations of laboratory tests and clinical signs and symptoms. The USA's Centers for Disease Control and Prevention (CDC) use a three-stage system for HIV-infected persons, based on a combination of age, CD4 count, and whether the patient has presented with an AIDS indicator disease.[13] The World Health Organization (WHO) classifies HIV/AIDS as primary HIV

infection (asymptomatic, or acute retroviral syndrome), or one of four clinical stages defined by AIDS indicator diseases of varying severities, ranging from stage 1 (asymptomatic or persistent generalized lymphadenopathy) to stage 4 (most severe, including extra-pulmonary and disseminated TB [discussed in more detail in section 1.2.1]).[14]

There is currently no cure or effective vaccine for HIV infection. HIV infection can be treated with anti-retroviral drugs, which target various stages of the viral replication cycle. The primary function of anti-retroviral therapy (ART) is to reduce the viral load, preventing CD4 decline or allowing CD4 recovery to restore immune function. ART is highly effective at slowing disease progression and reducing morbidity and mortality, particularly when started as soon as possible following HIV diagnosis and maintained.[15, 16]

An additional benefit of ART for PLHIV is that it can reduce the risk of onwards transmission of HIV; this is referred to as treatment as prevention. The efficacy of treatment as prevention is dependent on treatment adherence and maintaining an undetectable viral load.[17, 18] However, there have been no documented cases of HIV transmission from people with an undetectable viral load and the preventive action of ART against HIV transmission is an important contributor to ending the HIV epidemic.

ART can also be used prophylactically by HIV-negative persons to reduce the risk of being infected with HIV. Post-exposure prophylaxis involves administering ART to protect against HIV transmission after an event with high risk of transmission (described in more detail in section 1.1.3), such as a needle-stick injury with contaminated blood, or unprotected sex with someone known to be HIV-positive.[19, 20] Similarly people who are HIV-negative, but are at very high risk of acquiring HIV infection, can take ART regularly to lower their chances of becoming infected. This is referred to as pre-exposure prophylaxis. Pre-exposure prophylaxis is between 44-86%

effective when compared to placebo or no prophylaxis, and efficacy is strongly correlated with adherence.[21]

1.1.3 Transmission of HIV

Infection with HIV occurs through the transfer of bodily fluids, including blood, semen, vaginal fluids and breastmilk. This means that HIV can be transmitted through sexual contact, blood-borne transmission, or vertically from mother to child during childbirth.[3]

The risk of sexual transmission of HIV varies; the risk of infection (per sexual act) is higher for sex between men than for heterosexual contact, and transmission from men to women is more likely than from women to men. The risk is low for sex between women, or for oral sex acts, but anal sex carries a higher risk of HIV transmission.[22, 23] Viral load is a measure of the number of copies of the HIV virus in the blood. Higher viral loads are associated with increased risk of transmission, particularly during the acute phase of infection (see section 1.1.1), due to the combination of peaking viral load and the fact that most people are unaware of their HIV infection at this stage. The presence of other sexually transmitted infections are associated with higher likelihood of HIV transmission due to damaged mucosal surfaces, particularly genital ulcers,[23] although STIs such as gonorrhea and chlamydia also increase the risk.[24]

Blood-borne transmission can occur through a number of routes; most commonly through needle sharing during injecting drug use, needle stick injury or unclean medical injections or transfusion of contaminated blood or blood products. Less common routes include through organ or tissue transplantation, or through mucus membrane exposure to contaminated blood. Vertical transmission of HIV from mother to child may occur during pregnancy, but most commonly occurs during birth, or through breastfeeding.[3, 25]

It has been demonstrated that HIV-diagnosed persons who are on ART and have an undetectable viral load do not transmit HIV.[18]

1.1.4 HIV epidemiology

1.1.4.1 The global burden of HIV

Between 2010 and 2016, the estimated number of PLHIV globally increased by around a third, from an estimated 27.7 million in 2000 to 36.7 million in 2016, of which 34.5 million were adults (17.8 million [52%] women) and 2.1 million children.[26] There was an 11% decrease in the estimated number of new HIV infections in adults from 2010 to 2016 (1.9 million to 1.7 million) and a 47% decrease in new infections among children (from 300,000 in 2010 to 160,000 in 2016).

This growing population of PLHIV and the decrease in the number of new infections is partly explained by increased survival and decreases in HIV transmission as a result of widespread access to anti-retroviral therapy (ART). ART restores immune function,[27] decreases morbidity[16] and mortality,[28] and prevents transmission of HIV.[18] The estimated number of people accessing ART globally increased from 685,000 (3% of all PLHIV) in 2000 to 19.5 million (53%) in 2016.[26] Widespread educational campaigns to raise awareness and understanding of HIV infection and treatment, and promotion of preventive measures such as condom use and access to needle exchange programs, have also contributed to the decrease in new infections.[29]

People living in Africa are disproportionately affected by HIV infection, accounting for 70% of all PLHIV and two thirds of all new infections.[26] This is primarily due to poverty and a lack of health services supporting HIV prevention and treatment, compounded by socioeconomic and sociocultural factors which widen health inequalities.[30] However, HIV is also an important public health concern in high-income countries. In 2016 there were an estimated 2.1 million PLHIV in western and central Europe and North America, with 73,000 new infections and 18,000 AIDS-

related deaths (a decrease of 32% since 2010). An estimated 1.7 million PLHIV (81%) were on ART.[26] HIV also causes substantial financial costs; healthcare costs are high [31] and decreasing mortality rates mean that PLHIV are living longer and non-communicable disease are becoming an increasing problem in PLHIV.[32]

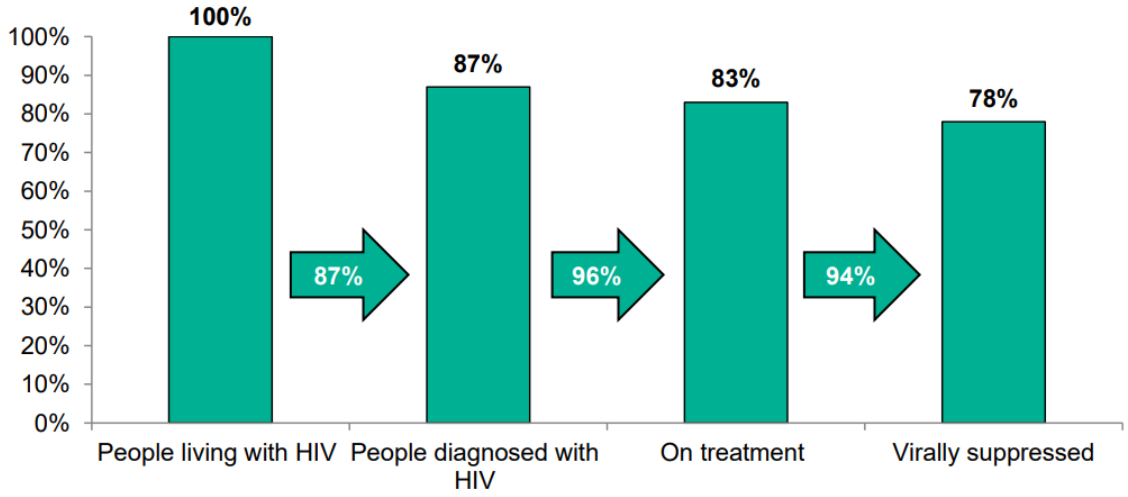
Important risk factors for HIV globally include men who have sex with men (MSM), people who inject drugs (PWID), people in prison or other congregate settings, transgender people, and sex workers and their clients.[3] Unprotected sex, other STIs, sharing needles or other injecting equipment or drug solutions, unsafe blood transfusions or organ transplantations, and accidental needle stick injuries, are also risk factors for transmission of HIV.[3]

1.1.4.2 HIV in the UK

Reporting of data on HIV infections in the United Kingdom (UK) is described in detail in section 3.1.2.2. As of 2015, there were an estimated 101,200 people living with HIV in the UK, although an estimated 13,500 were not aware of their HIV infection. This gives an overall prevalence of 160 PLHIV per 100,000 population; 210 per 100,000 in adults aged 15-74. There were 6,095 new diagnoses of HIV in 2015, a diagnosis rate of 11.4/100,000 population.[10]

The proportions of PLHIV who are diagnosed, linked to care, have initiated ART, and are virally suppressed are high in the UK (Figure 1.2). Late diagnosis is common; 39% of diagnoses in 2015 were at a late stage of infection (CD4 count less than 350 cells/μl), although this had declined from 56% in 2006.[10] Late diagnosis is associated with worse outcomes, including a ten-fold increase in the risk of death in the first year following HIV diagnosis.

Figure 1.2: Continuum of HIV care in the UK, 2015.



Source: Public Health England [10]

PLHIV in the UK largely acquired their HIV infection through sexual contact; in 2015 an estimated 47,000 PLHIV were MSM (population HIV prevalence 5,870 per 100,000), and there were 19,600 heterosexual men and 29,900 heterosexual women living with HIV (prevalence 100 per 100,000).[10] 58% of heterosexual PLHIV were black African; the HIV prevalence among black African adults was 2%. There is also a much smaller population of PLHIV who acquired their HIV infection through injecting drug use, approximately 2,500 as of 2015. The estimated HIV prevalence among PWID was 380 per 100,000.[10]

There were 305 diagnoses of AIDS-defining illnesses within 3 months of HIV diagnosis in 2015, a decline of over 50% since 2006.[10] The most common AIDS-defining illnesses were Pneumocystis pneumonia (43%), oesophageal candidiasis (11%) and TB (9%). 594 PLHIV died in 2015. Mortality rates were consistent between MSM and heterosexuals with diagnosed HIV (480/100,000 and 430/100,000 respectively), but higher among PWID (2,500/100,000).[10]

1.2 Tuberculosis

In humans, TB is a disease predominantly caused by the bacteria *Mycobacterium tuberculosis (M. tb)*, but also by other mycobacteria in the *Mycobacterium tuberculosis* complex (*MTBC*, a genetically related group of

Mycobacterium species capable of causing TB) such as *M. bovis*, *M. africanum*, and *M. microti*.

1.2.1 Tuberculosis infection and pathogenesis

TB is an airborne disease, and infection occurs via the lungs. Following infection, a number of scenarios can occur. Infection can be cleared by the immune system; this immune response is not well characterized. If infection is established, primary disease can occur within the first 1-3 years after infection. Alternatively, latent TB infection (LTBI) may be established, where the bacteria lie dormant without replicating, and the patient is asymptomatic and non-contagious.[33] LTBI can reactivate and cause active TB disease.

The lifetime risk of progressing from LTBI to active disease is 5-10%; however this risk is much higher for certain groups of people such as PLHIV, individuals with diabetes, and those who are malnourished or smoke tobacco. The most common presentation is pulmonary disease, affecting the lungs, but TB disease can also manifest in extra-pulmonary forms such as lymphatic TB, pleural TB, skeletal TB, central nervous system TB, meningeal TB, abdominal TB, genitourinary TB, or miliary (disseminated) TB.[34] Extra-pulmonary disease is often harder to identify and diagnose as clinical signs are less specific. Severe forms of extra-pulmonary TB (miliary and meningeal) are associated with worse outcomes.[35, 36]

1.2.2 Tuberculosis diagnosis

1.2.2.1 Diagnosis of latent tuberculosis

As LTBI is asymptomatic and there is no bacterial replication, there are no direct diagnostic tests for LTBI. Instead, diagnostics rely on detecting an immune response to infection and then ruling out active disease. Widely available diagnostic tests for LTBI include the tuberculin skin test (TST) and interferon-gamma release assays (IGRAs); whilst a new skin test (C-Tb) has been evaluated but is not yet widely used.

The TST involves an intradermal injection of tuberculin purified protein derivative, which causes a delayed-type hypersensitivity (DTH) reaction if the patient is infected with TB.[37] The resulting induration is then measured. Guidelines for interpreting the results of the TST vary; the CDC diagnose LTBI based on an induration of >10mm (>5mm for immunosuppressed patients),[38] whilst in the UK there is a dual threshold of 5mm for people who have not received the Bacillus Calmette-Guérin (BCG) vaccine and 15mm for those who have been vaccinated.[39]

The TST is generally widely available (although there have been shortages in recent years), but has some disadvantages. Administering and interpreting the TST requires two visits by the patient 48-72 hours apart,[37] and patients often do not return for their second visit.[40] Additionally, TST sensitivity can be reduced by malnutrition, severe TB disease or immunosuppression, whilst decreased specificity can occur because of infection with non-tuberculous mycobacteria or prior vaccination with the BCG vaccination (currently the only licensed vaccine against TB).[40]

IGRA blood tests are more specific due to a lack of cross-reactivity, detecting the interferon-gamma immune response to specific *Mycobacterium tuberculosis (M. tb)* antigens which are not present in BCG or most non-tuberculous mycobacteria,[40] and consequently their results are not confounded by prior BCG vaccination. Commercially available IGRAs include the QuantiFERON-TB Gold In-Tube (QFT-GIT), the newer QuantiFERON-TB Gold-Plus, and the T-SPOT.TB. As well as greater specificity, other advantages of IGRA tests include greater sensitivity in PLHIV with low CD4 counts, only requiring a single patient visit, the fact that results can be available within 24 hours, and that results cannot be affect by the perception or bias of healthcare workers (although they can return an indeterminate or borderline result).[37, 41] However, IGRAs are more expensive, more technically complex to conduct, involve venepuncture, and may be less sensitive in high-burden settings.[40, 42]

The C-Tb skin test is specific to *M. tb* and is applied in the same way as the TST, but which is more specific than the TST and has comparable sensitivity to the

QFT-GIT in both HIV-positive and HIV-negative patients.[43, 44] However, like IGRAs it lacks sensitivity to *M. tb* at very low CD4 counts,[44] and it is not yet widely used.

1.2.2.2 Active tuberculosis diagnosis

Active TB is usually suspected from clinical signs or symptoms, and confirmed by laboratory testing. Symptoms of active TB vary depending on the site of disease, but commonly include cough (with sputum and, at times, blood), fatigue, fever, night sweats, loss of appetite and weight loss. In cases of pulmonary disease, a chest x-ray can show nodules, lesions or cavitation that are symptomatic of TB; although not all patients with TB show abnormalities in the lungs, and some lesions are not specific to TB.

The gold-standard for TB diagnosis is culture confirmation of *MTBC*; a positive culture confirms the presence of TB bacteria. Culturing *MTBC* can take 2-4 weeks, therefore if TB is suspected it is usually recommended to begin treatment before the results of the culture are available.

Other methods of diagnosing active TB include sputum smear microscopy, histology, or PCR testing. Sputum smear microscopy involves checking for the presence of acid-fast bacilli on a sputum smear. Whilst this method lacks sensitivity (particularly in immunocompromised individuals) and specificity [45] (as not all acid-fast bacilli are *MTBC*, and it cannot differentiate between live and dead bacteria) it is nonetheless a useful diagnostic tool and is often used as a proxy for infectiousness of TB.

The GeneXpert MTB/RIF® test is a nucleic acid amplification test which has been recommended by the WHO since 2010.[45] It can detect the presence of live or dead *M. tb* bacteria, and simultaneously assesses whether the bacteria are resistant to rifampicin, an important first line anti-TB drug.[46] It is reliable and quick, with results available in under two hours, however it is also expensive and reliant on good healthcare infrastructure.[47]

1.2.3 Treatment of tuberculosis

Treatment of LTBI can prevent progression to active TB, reducing the risk by at least 60%.[48] Testing and treatment for LTBI is recommended by the WHO in specific high-risk populations in high- and middle-income countries with TB incidence below 100/100,000 population.[49] Systematic screening (with IGRA or TST) and treatment is recommended for PLHIV, contacts of pulmonary TB cases, and other immunocompromised individuals, and should also be considered for prisoners, homeless persons, people who use illicit drugs, healthcare workers, immigrants from countries with high TB incidence.[49] Standard treatment regimens for LTBI consist of 6 or 9 months of isoniazid, or other combinations of isoniazid and rifapentine or rifampicin, or rifampicin alone. Regimens including rifampicin or rifapentine should be prescribed with caution for PLHIV on ART, as there is potential for drug-drug interactions.[49]

Active TB is treatable and can be cured with antibiotics. Standard treatment guidelines for drug-sensitive active TB are 2 months of isoniazid, rifampicin, pyrazinamide and ethambutol, followed by 4 months of isoniazid and rifampicin.[50, 51] Drug resistance is an increasing problem, and this regimen can be adapted and/or extended if necessary, guided by the results of drug sensitivity testing and the site of disease. TB can be mono-resistant (resistant to a single drug, commonly rifampicin or isoniazid), multi-drug resistant (MDR; resistant to multiple drugs), or extensively drug-resistant (XDR; resistant to isoniazid and rifampicin, a fluoroquinolone, and at least one injectable second-line drug). Treatment guidelines for drug resistant TB vary between settings and are dependent on affordability, drug availability and drug resistance patterns.

1.2.4 Tuberculosis transmission

TB is an airborne disease, spread by droplet nuclei when TB patients with pulmonary disease cough, sneeze or spit. The risk of transmission is dependent on several factors; degree of exposure to an infectious TB case (both duration of exposure

and proximity of exposure), the infectiousness of the TB case (there is variation in the number of infectious quanta produced by patients, e.g. as measured by smear status), and the immune status of the individual being exposed. There are a number of ways TB transmission can be studied; these are described in detail below.

1.2.4.1 Tools for measuring tuberculosis transmission

TB transmission can be studied at the outbreak level by contact tracing and case finding (identifying and screening close contacts of known TB cases for infection and/or disease), at the patient level (by longitudinal testing for infection and/or disease to establish whether the case patient is infectious), or at the population level. Population-level studies can involve longitudinal studies of TST conversion or IGRA positivity, to calculate the risk of infection with TB in a given time period, or collecting epidemiological and/or genomic data for all TB cases in a given population to try and identify chains of transmission. These methods are described in more detail in sections 1.2.4.1.1 and 1.2.4.1.2.

Other methods of understanding transmission include using the rate of TB in non-foreign-born children as a proxy for the rate of recent, within-country transmission,[52] modelling studies,[53, 54] studies of transmission dynamics and infectious quanta production using guinea pigs exposed to hospitalized TB patients,[55, 56] studies of air sharing and the presence of TB in droplet nuclei, and investigating the impact of infection control mechanisms such as improved ventilation and decreased population density (particularly in confined settings, e.g. prisons),[53, 57] and protective mechanisms such as face masks or UV lighting.[58, 59] These fall outside the scope of this thesis.

1.2.4.1.1 Study types

Contact studies can be carried out prospectively or retrospectively. They identify contacts of TB case patients (usually close contacts such as household members or colleagues) and measure the incidence and/or prevalence of LTBI and/or active TB disease. Data is usually collected on the characteristics of the case patient,

and often the characteristics of the contacts, which allows analysis of the incidence or prevalence of TB among their contacts. From this, it is possible to infer factors which may cause or be associated with infectiousness (of the case patient) and susceptibility to infection (of the contact).

Outbreak studies are usually carried out in response to a particular outbreak of TB, usually where it has been identified that a higher number of cases occurred in a particular location or time period than would be expected. Contact between case patients or locations where transmission may have occurred are retrospectively identified, and genotyping of patient samples is often done to establish whether cases could be part of the same transmission chain. The results of outbreak investigations may prompt 'active case finding' of other individuals who may have been infected as part of the same outbreak, by screening contacts of the case patients or people with whom they may have come into contact in a shared environment, e.g. a workplace, hospital or homeless shelter.

Transmission can also be studied at the population level. Many countries, particularly countries with low TB incidence, now routinely genotype samples from notified TB cases. This enables the identification of cases that may belong to the same outbreak; with some limitations, described in section 1.2.4.1.2. This data can be used to initiate outbreak investigations, as described previously, and can also be used to group genomically linked cases into clusters. Epidemiological data, where it is available, can also be used to define clusters. Data on clusters can be used to estimate the proportion of cases that may be due to recent transmission using the 'n minus one' rule (Figure 1.3), to understand which lineages of TB are driving the epidemic, and can also be used to investigate factors associated with transmission.

Figure 1.3: The 'n minus one' rule.

$$\frac{number\ of\ clustered\ cases - number\ of\ clusters}{number\ of\ strain\ typed\ cases}$$

The usefulness of population-level genotyping for estimating clustering is dependent on how complete the sampling of the population is. Many studies are only representative of a limited population and do not capture all TB cases. Even country-level studies where notification of TB cases is mandatory are limited; not all TB cases are notified, notified cases are not all culture positive (particularly extra-pulmonary TB cases), and only culture positive cases can be genotyped. Furthermore, cases may belong to clusters with other patients who are outside the region or country of the study, or who presented with disease before or after the study time period.

Incomplete sampling results in underestimation of the clustering of TB cases; lower sampling fractions particularly result in greater underestimation of the proportion of clustered cases.[60] Incomplete sampling principally affects smaller clusters, which are less likely to be identified when the sampling fraction is low; a smaller number of larger clusters is less susceptible to bias from incomplete sampling.[60] Smaller sampling fractions can cause underestimation of the impact of risk factors associated with clustering;[61] although this is ameliorated by the fact that risk factors for clustering are often risk factors for increased cluster size, which limits the sampling bias.[62] It has also been shown that studies of one year or more, or studies in a limited geographical area, could reliably identify risk factors for clustering provided the total sample size was more than 1,000 cases, the sampling was random (which is often not the case), and the sampling fraction was over 40%.[62]

The molecular methods by which clusters are defined are described in the following section.

1.2.4.1.2 Genotyping tools

Genotyping refers to the investigation of the genetic makeup of an organism - in this case, *MTBC*. Genotyping can examine all or parts of the bacterial genome.

Restriction fragment length polymorphism (RFLP) utilizes the IS*6110* DNA sequence, which is found in most *MTBC* genomes (often with multiple copies), to

distinguish between strains of *MTBC*. A restriction enzyme is used to cut the DNA strands intro fragments at specific sequences in the DNA. These fragments are then sorted by size using gel electrophoresis. A probe which binds to the IS*6110* sequence visualizes the fragments containing this sequence, and the pattern of bands can be compared to a reference standard or patterns from other isolates. TB isolates with the same or very closely related genomes share banding patterns. RFLP has high discriminatory power, but is intensive in terms of the time and laboratory capabilities necessary. Furthermore, some strains of TB contain few copies of IS*6110*, which reduces the discriminatory power of RFLP.[63]

Mycobacterial interspersed repetitive units – variable number tandem repeats (MIRU-VNTR) is a strain typing method which focuses on specific sections of the TB genome (loci) which contain specific DNA sequences which repeat a number of times. The number of repeats varies between different strains of TB. These sequences can be amplified using polymerase chain reaction (PCR), and the number of repeats at each locus enables different strains to be classified. The number of loci used can vary, with discriminatory power increasing as more loci are used. 24-locus MIRU-VNTR is the most discriminatory version used.[64] MIRU-VNTR has comparable resolution to RFLP,[64, 65] is considerably less laborious (allowing higher throughput), and results in a numerical output that is easily comparable between laboratories.

Widespread use of whole-genome sequencing (WGS) is a relatively novel approach to genotyping *MTBC*, which replicates and mutates slowly. WGS involves sequencing the whole genome of the bacteria in the sample, and has the power to detect single nucleotide polymorphisms (SNPs) – single nucleotide changes in the genome. This enables the construction of chains and networks of transmission, based on small evolutionary changes which occur as the organism replicates. WGS data can be used to create phylogenetic trees, showing how samples from an outbreak or population are genetically related to each other. These can then be combined with data on diagnosis and/or symptom onset date, and duration of infectiousness (where known,

estimated by sputum smear testing or dates of TB treatment) to create a model of an epidemic, either to aid understanding of an outbreak or estimate the impact of interventions which may help prevent similar outbreaks.

There has been much debate regarding what the SNP threshold should be, as a cut-off for determining whether two isolates derive from the same source. It has been estimated that the TB genome evolves slowly, at approximately 0.5 SNPs per year.[66-68] SNP thresholds of between 0-10 SNPs have been used to define transmission,[67] supported by an analysis combining WGS and epidemiological data which found that there were no epidemiological links between patients whose isolates had >5 SNPs difference.[68]

WGS has also provided a gold-standard against which other genotyping tools can be evaluated. WGS has demonstrated the discriminatory power to distinguish two separate outbreaks within one MIRU-VNTR cluster,[69] and shown that microevolution within a host can change the MIRU-VNTR strain type.[68] However, MIRU-VNTR is still considered a highly discriminative technique and has many advantages over WGS in terms of cost, throughput and comparability. In particular, strain typing using RFLP or MIRU-VNTR has been routine in many countries for several years,[52, 70-72] giving datasets of a depth and breadth which is not yet possible or practical with WGS.

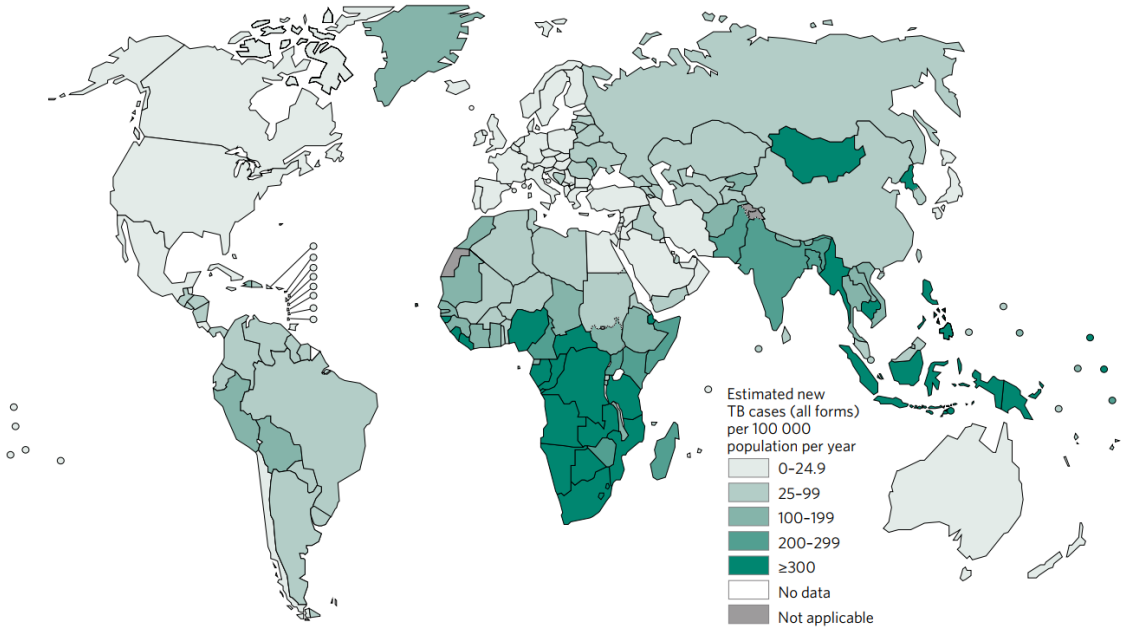
There are some limitations to genotyping. It cannot confirm transmission between two individuals as a shared genome between two TB cases does not prove that one was infected by the other; it is possible that both patients were infected from a common source, or it may be a common genotype. Genotyping also cannot identify the direction of transmission between individuals, although this can sometimes be estimated from dates of diagnosis or symptom onset. Furthermore, many methods of genotyping cannot distinguish between closely related TB isolates. Genotyping can therefore only be definitively used to rule out transmission between individuals, not to confirm it.

1.2.5 Epidemiology of tuberculosis

1.2.5.1 The global burden of tuberculosis

It is estimated that between a quarter and a third of the world's population are infected with latent TB.[45, 73] In 2015, an estimated 10.4 million people fell ill with TB disease, an incidence of approximately 142/100,000 population.[45] Since 2000, the number of incident TB cases has declined from 13.0 million to 10.4 million in 2015 (an average of 1.5% per year). This rate of decline will need to be increased if the WHO's sustainable development goals aim of a 35% reduction by 2020 is to be met.[1] Sixty percent of the global TB burden is concentrated in six countries - India, Indonesia, China, Nigeria, Pakistan, and South Africa - but incidence is high across sub-Saharan Africa and much of south-east Asia (Figure 1.4).

Figure 1.4: Estimated TB incidence rates in 2015.



Source: World Health Organisation [1]

TB is the leading infectious cause of death globally, responsible for 1.8 million deaths in 2015.[1] More than 95% of TB deaths were in low- and middle-income countries, with 86% of TB deaths occurring in Africa or south-east Asia. The case fatality rate (the proportion of TB patients who die from the disease) was 17% in 2015; varying from below 5% in some countries, to over 20% in most of Africa.[1]

The majority of TB disease is pulmonary; 6.1 million incident TB cases were notified to the WHO in 2015 and 15% of these were extra-pulmonary disease.[1] This varies by region, with extra-pulmonary TB representing 8% of TB cases in the western Pacific region and 23% in the eastern Mediterranean region.

Drug resistance is an important problem. In 2015 an estimated 580,000 TB cases were rifampicin- or multi-drug resistant, accounting for 4% of new TB cases and 21% of previously treated cases.[74] Death rates among patients with drug resistance are high; an estimated 250,000 people died of rifampicin resistant or multi-drug resistant TB in 2015.[74]

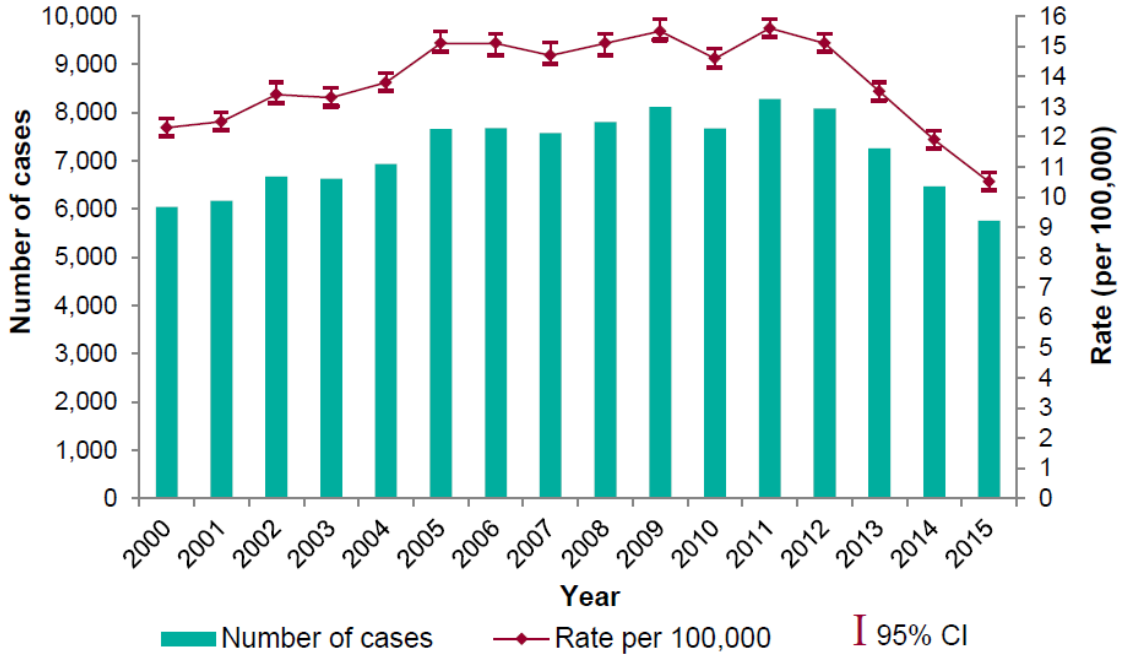
Important risk factors for TB globally include HIV infection (see section 1.3), diabetes, malnutrition, drug or alcohol misuse, and smoking of tobacco or other substances.[75, 76] Other groups that are particularly vulnerable to TB because of difficulties accessing healthcare services include those in poverty, refugees or displaced populations, and people who are imprisoned or work in prisons.[77] Globally, men have a substantially higher risk of TB than women; there were 5.9 million cases of TB in men, 3.5 million in women and 1.0 million in children.[1]

1.2.5.2 Tuberculosis in the UK

TB has been a notifiable disease in the UK since 1912.[78] Public Health England (PHE) receives statutory notifications of all diagnosed TB cases through a national surveillance program (described in detail in section 3.1.2.1).

In 2015, 5,758 TB cases were notified in England, a rate of 10.5 per 100,000 population.[52] TB incidence increased from 2000 to 2011, but has since been declining (Figure 1.5). Outside England, TB incidence was even lower, with 116 TB cases notified in Wales (3.7/100,000), 315 in Scotland (5.9/100,000) and 70 in Northern Ireland (3.9/100,000) in 2015.[79-81]

Figure 1.5: TB case notifications in England, 2000-2015.



Source: Public Health England [52]

TB cases in England are concentrated in London, which accounted for 39% of all cases in 2015. 59% were male and 58% were aged 15-44. Of patients with a recorded place of birth, 73% were born outside the UK; the incidence rate of 73.9/100,000 for foreign-born patients was 15 times higher than that in the UK-born population in 2015 (3.4/100,000).[52] The most common countries of birth for patients with TB in 2015 were India, Pakistan, Somalia, Bangladesh and Nepal, representing a decline in TB patients from sub-Saharan Africa and increases in patients from the Indian sub-continent and eastern Europe since the mid-2000s.[82, 83] This changing epidemiology is partially due to decreased migration from sub-Saharan Africa since the mid-2000s.

The most common site of disease in 2015 was pulmonary, with 53% of cases reporting pulmonary disease (with or without extra-pulmonary disease).[52] 36% reported lymph node TB, and severe forms of TB such as miliary and meningeal TB were reported for 2.8% and 2.3% of patients, respectively. 61% of cases had a sputum smear result reported and 50% of these were sputum smear positive. 95% of sputum smear positive cases were also culture confirmed.[52] Overall, 60% of TB cases were

culture confirmed; 73% of pulmonary cases and 46% of extra-pulmonary cases. The majority of cases were drug-sensitive; of culture confirmed cases 7.4% were resistant to at least one first line drug, 6.9% were isoniazid resistant, 1.4% were rifampicin resistant, and 1.3% had MDR-TB.[52] Rates of treatment completion (defined as completing a full course of treatment within a specified timeframe, usually 12 months) without evidence of relapse [84]) were generally high (85% of drug-sensitive cases had completed treatment at 12 months), but lower among patients prescribed longer treatment regimens, e.g. those with drug-resistant TB (56% completed at 24 months) or who had central nervous system, spinal, miliary or cryptic disseminated TB (67% had completed treatment at their last recorded outcome).[52]

Ten percent of notified TB cases in 2015 had at least one social risk factor (drug misuse [3.3%], alcohol misuse that would prevent adherence to TB treatment [3.3%], imprisonment [3.3%] or homelessness [3.4%]). TB patients with social risk factors represented 21.7% of UK-born TB cases.[52] The number of TB cases with social risk factors has not declined since 2010; consequently the proportion of TB cases with social risk factors has been increasing.

1.3 HIV and tuberculosis co-infection

There is a synergistic relationship between HIV infection and TB disease. Impaired immune function as a result of untreated HIV infection increases the risk of progression from LTBI to active TB.[85, 86] There is also some evidence that TB accelerates the progression of HIV-related disease. Higher rates of AIDS-defining illnesses other than TB have been observed in co-infected patients than HIV-positive patients without TB, even after accounting for CD4 count.[87] Additionally, immune system activation in response to TB has been shown to increase HIV replication, increasing its severity and infectiousness.[88]

1.3.1 The role of HIV in tuberculosis transmission

It is well established that HIV infection increases susceptibility to active TB disease by increasing the rate of progression from LTBI to active disease,[85, 88] resulting in substantially higher incidence rates of TB.[89] However, it is unclear whether HIV infection increases the risk of infection with *MTBC*.[90] HIV infection is also known to increase the risk of recurrent TB.[91] Recurrent TB can be the result of reactivation of LTBI, or re-infection from an exogenous source. A review of TB clustering in TB-endemic countries reported that TB cases with HIV were more likely to be part of a strain type cluster than HIV-negative TB cases; and based on the assumption that PLHIV with TB were less likely to be infectious, concluded that the increase in clustering was a result of more recent infection in PLHIV and not increased reactivation of LTBI.[92] However, the applicability of this conclusion to low-incidence settings, where the majority of TB patients were born abroad and the overall force of infection is much lower, has not been investigated.

Unlike many other risk factors associated disease transmission, HIV may increase susceptibility to TB disease whilst also causing TB patients to be less infectious. Contact studies have shown lower prevalence of TST-positivity and lower rates of TST conversion among contacts of HIV-positive index patients than HIV-negative index patients,[93-95] particularly where the index patients with HIV were immunosuppressed.[96] This may be mediated through a shorter duration of infectiousness due to accelerated TB disease progression, earlier diagnosis [88, 89] and more rapid commencement of TB treatment,[96] and greater mortality.[97, 98] Additionally, lower proportions of cavitary [94, 96] or sputum smear-positive [94, 95, 99] TB, greater likelihood of extra-pulmonary disease,[100] lower sputum bacterial load,[101, 102] and shorter duration of smear positivity [89] or cough [94] may all make HIV co-infected patients less likely to transmit infection than HIV-negative TB patients.

1.3.2 Clinical complexity of co-infected patients

The interaction between TB and HIV is also important clinically, and the two diseases need to be managed together for the best outcomes for patients. Co-infection with HIV and TB can complicate TB diagnosis and treatment, and is generally associated with worse outcomes for patients. Diagnostic tests for LTBI appear less sensitive in HIV-infected individuals, although IGRAs are less affected by low CD4 counts and anergy than TSTs.[103, 104] HIV-positive TB patients often have a lower concentration of bacteria in their sputum, making it harder to detect in a sputum test.[105] Extra-pulmonary TB is also more common among HIV-positive patients, cannot be diagnosed with sputum tests or chest x-rays, and has less specific clinical signs, making diagnosis tricky.[105] HIV co-infection is associated with recurrence of TB,[91, 106] higher rates of other AIDS-defining conditions or opportunistic infections,[87] and higher mortality, particularly at low CD4 counts.[107]

Co-infection also leads to clinical complexity. Immune reconstitution inflammatory syndrome (IRIS) is a condition affecting patients co-infected with TB and HIV, which occurs when pathogen-specific immune responses are restored, usually by a rapid increase in CD4 count. IRIS typically occurs soon after initiation of ART, and can present as the relapse of a previously treated infection, or a new presentation of what was previously sub-clinical disease.[108] This has resulted in debate regarding the optimal timing of ART initiation for co-infected patients, due to concerns regarding drug side effects and adherence; however, general guidance is to start ART as soon as possible after TB treatment, particularly for patients at low CD4 counts.[109, 110]

1.3.3 Epidemiology of HIV-tuberculosis co-infections

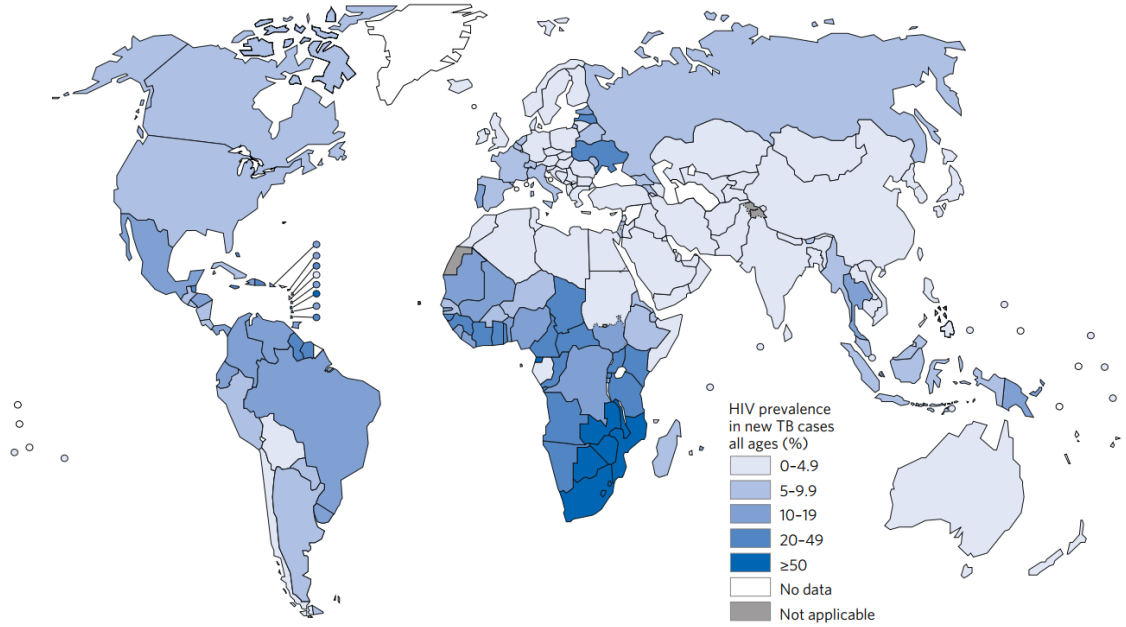
1.3.3.1 Global epidemiology of co-infections

Globally, PLHIV are 17-22 times more likely to develop TB disease than persons not infected with HIV.[111] HIV infection has a significant effect on the epidemiology of TB, increasing the risk of progression from LTBI to active TB disease from a lifetime risk of approximately 10% to an annual risk of around 10%.[112] The

risk of developing TB disease doubles in the year following acquisition of HIV infection, and increases further in following years if HIV remains untreated as immunodeficiency increases.[85, 113] The one-year mortality rate in HIV-positive people is significantly higher in individuals with TB disease, particularly in those with low CD4 counts.[114]

There is considerable geographic overlap in the burden of HIV infection and TB disease globally, particularly in sub-Saharan Africa but also in smaller areas (e.g. London in the UK). Shared risk factors for both diseases also contribute to high rates of co-infection. Both diseases disproportionately affect disadvantaged populations, such as migrants and people who are socioeconomically deprived, homeless, incarcerated or have problems with drug misuse. The elevated risk of TB in these populations occurs through two mechanisms, increased risk of exposure and infection, and an increased risk of progression from LTBI to active TB. Sociodemographic factors such as homelessness, imprisonment, and/or crowded and poorly ventilated living or working conditions increase the risk of transmission by increasing the likelihood of contact with an infectious TB case. Difficulties accessing healthcare can also delay diagnosis, allowing disease to progress further and prolonging the infectious period. The risk of progression from LTBI to active disease is affected by the host immune response; which can also be impaired by other risk factors common to both TB and HIV such as undernutrition, diabetes, smoking and alcohol misuse.[87]

Figure 1.6: Estimated HIV prevalence among TB cases in 2015.



Source: World Health Organisation [1]

In 2015, there were an estimated total of 10.4 million TB cases globally, of which an estimated 1.2 million (12%) were among PLHIV.[115] Within this total, 6.4 million people with TB were notified to national TB programs and reported to the WHO. 3.4 million notified cases (55%) had an HIV test result (this varied by region and was over 80% in Africa and the Americas), which represents an 18-fold increase in HIV testing among TB patients since 2004.[1] Over 500,000 (15%) notified TB cases with known HIV status were HIV-positive. The estimated proportion of TB cases with HIV was highest in Africa (31%), exceeding 50% in some parts of southern Africa (Figure 1.6).[1] The increase in HIV testing of TB patients means that whilst 78% of TB cases with HIV notified in 2015 had initiated ART, it is estimated that overall only a third of co-infected patients globally have started ART; substantially lower than the estimated 53% of all PLHIV who are on ART.[26] This is likely a result of simultaneous diagnoses of TB and HIV, where patients are unaware of their HIV infection until they are diagnosed with TB.[3] The WHO recommends systematic testing for TB among PLHIV, but this is an insufficient strategy to prevent TB if HIV has not already been diagnosed.

TB is the leading cause of death among PLHIV; around 390,000 co-infected patients died in 2015, representing 35% of all HIV-related deaths and 22% of TB-related deaths.[1, 45]

1.3.3.2 HIV co-infection in European countries with low tuberculosis incidence

In countries with low TB incidence, both TB and HIV are at very low levels in the general population, with epidemics concentrated in certain populations. There is considerable overlap between high-risk populations for TB and HIV. Both diseases disproportionately affect migrants (particularly those from countries with high burdens of TB and HIV), people who are socially deprived such as homeless people, people who inject drugs or misuse other substances and people who are imprisoned.

A 2009 systematic review of TB-HIV co-infection in Europe reported that the proportion of TB cases co-infected with HIV varied between 0-15% across European Union (EU) and European Economic Area (EEA) countries, with the highest proportions of co-infection in Portugal (14.6% in 2008), Estonia (9.9% in 2008), Malta (9.4% in 2008) and England and Wales (7.4% in 2005).[116] This was confirmed by an analysis of surveillance data collated by the European Centre for Disease Prevention and Control (ECDC) in 2014 where the prevalence of co-infection was highest in Latvia (19.5%), Malta (17.1%), Portugal (14.7%) and Estonia (10.1%). It is notable that this analysis excluded data from the UK, France, Germany and Poland, which do not report HIV status for TB cases to ECDC, but together account for 40% of TB cases in Europe. This study also excluded data from countries with poor recording of HIV status of TB cases (particularly the Czech Republic, Denmark, Hungary and Iceland).[117]

TB incidence among PLHIV is also substantially higher than in the general population. For example, TB incidence among PLHIV was estimated at 370/100,000 in Germany between 2001-2011 [118] and 375/100,000 in the UK between 1995-2005;[119] these incidence rates were 30-45 times higher than the average incidence in the general population over these time periods of approximately 8/100,000 and 13/100,000, respectively.[120]

Whilst the absolute numbers of people affected by TB-HIV co-infection are relatively few in low-incidence countries, co-infection remains an important clinical problem. Treatment outcome success was significantly lower for co-infected patients than HIV-negative TB patients across Europe in 2013, with only 58% patients successfully completing treatment compared to 84% for HIV-negative patients.[117] The proportion of co-infected patients that died whilst on TB treatment was also significantly higher; 14% of co-infected patients died compared to 5.6% of HIV-negative patients.[117] Rates of death for co-infected patients were substantially higher for patients in eastern Europe than western Europe.[121]

1.3.3.3 Co-infections in the UK

Using the most recent available data, 197/6,209 (3.2%) of adult TB cases in England in 2014 were in people co-infected with HIV.[82] The proportion of new cases of TB in England, Wales and Northern Ireland that were in individuals co-infected with HIV rose from 3.1% in 1999 to a peak of 9.5% between 2005-2006.[122, 123] Over 30% of the increase in the number of new TB cases in adults aged 15-64 years between 1999 and 2003 was estimated to be as a result of HIV infections.[123] Co-infection levels have since decreased, to 4.0% in 2011 and 3.2% in 2014.[82, 124] This has partly been attributed to changing migration patterns, resulting in less TB in black African individuals from countries with a high HIV prevalence and more among individuals from Asian countries with low HIV prevalence.[82]

In the UK, the annual incidence of TB among PLHIV was 1,600/100,000 in 2004 and 1,750/100,000 in 2008,[125, 126] but incidence declined to 438/100,000 by 2011.[125] This decline is partly due to an increase in the number of people living with HIV who are on effective ART, and also due to a decrease in new HIV diagnoses in people originating from sub-Saharan Africa, where the prevalence of both infections is high and late presentation to health services is common.[125] Until 2015, TB was the second most common AIDS-defining illness among PLHIV in the UK.[10, 125]

1.4 Summary and objectives

Co-infection with HIV and TB is an important clinical and public health problem. Despite low levels of both HIV and TB in many high-income countries, both diseases are more prevalent in hard-to-reach sub-populations, making it difficult to eradicate these diseases. Timely diagnosis of HIV can be challenging, which leaves opportunities for diseases such as TB to manifest as immune function declines, causing significant morbidity and mortality. In contrast to high-burden settings, where HIV infection contributed substantially to the modern TB epidemic, the extent of the role of HIV infection in TB transmission in low-incidence settings is less clear. The aims of this thesis are therefore as follows:

1. To review the literature on the risk factors for developing TB for people with HIV in settings with low TB incidence (Chapter 2).
2. To link national TB and HIV surveillance datasets to identify co-infected patients in England, Wales and Northern Ireland (Chapter 3).
3. To describe the epidemiology of TB-HIV co-infections in the UK and examine risk factors for developing TB for people living with HIV (Chapter 4).
4. To describe trends in HIV co-infection among TB cases and investigate risk factors for co-infection (Chapter 5).
5. To examine whether HIV is a risk factor for TB transmission in the UK (Chapter 6).

2 Tuberculosis infection and disease in people living with HIV in countries with low tuberculosis incidence: A systematic review

2.1 Introduction

TB-HIV co-infection remains a persistent public health problem in countries with low TB incidence, which is defined by the WHO as less than 10 cases per 100,000 population.[87] In 2012, an estimated 155,000 people fell ill with TB in low-incidence countries and 10,000 people died.[127] TB in low-incidence countries mostly affects vulnerable populations, such as recent migrants (particularly those from countries with high TB incidence), people who are homeless, incarcerated, or have problems with substance abuse, as well as PLHIV.[127] There is considerable overlap between these groups, who are often referred to as 'hard to reach' by healthcare services as poverty and social deprivation make accessing care difficult.[128]

Accordingly, priority areas of the WHO's action framework for TB in low-incidence countries include addressing TB in the most vulnerable and hard-to-reach groups, screening for latent and active TB in high-risk groups and ensuring continued surveillance, evaluation and case-based data management.[87] Universal screening of PLHIV for active TB is recommended, but is not implemented in all low-incidence countries.[87] Universal screening for LTBI is also recommended for PLHIV and other high-risk groups, but is also not widely implemented. Other interventions - such as the offer and recommendation of a HIV test to all people newly diagnosed with TB, contact tracing and active case finding - focus more on finding and promptly treating new cases, and preventing further transmission, than on preventing TB occurring as a result of LTBI reactivation.

Efforts to control TB in people with HIV are hampered by the significant proportion of PLHIV who are undiagnosed, diagnosed late, or not diagnosed until they present to healthcare service with other opportunistic infections such as TB.[129]

Managing TB in PLHIV is also more clinically complex, as a result of side effects of treatment, drug-drug interactions, and often a more severe presentation of disease.[130] Preventing TB in PLHIV is therefore an important goal. Pressure on resources, even in countries with low TB incidence, necessitates the targeting of preventative activities such as screening for HIV, LTBI and active TB. It is therefore critical to have a clear insight into risk factors for TB disease in PLHIV in countries with low TB incidence. This review provides an up-to-date synthesis of such evidence for both LTBI and TB among PLHIV.

2.2 Methods

A systematic search of PubMed, EMBASE and Web of Science was done on 4th May 2017 for articles on risk factors for LTBI or active TB disease among PLHIV in countries with low TB incidence.

Low TB incidence was defined as per the WHO definition, as ≤10 cases per 100,000 population,[87] in 2015, and included the following countries: American Samoa, Andorra, Antigua and Barbuda, Australia, Austria, Barbados, Belgium, Bermuda, Bonaire, Saint Eustace and Saba, British Virgin Islands, Canada, Cook Islands, Cuba, Curacao, Cyprus, Czechia, Denmark, Finland, France, Germany, Greece, Grenada, Hungary, Iceland, Ireland, Israel, Italy, Jamaica, Jordan, Luxembourg, Malta, Monaco, Montserrat, the Netherlands, New Zealand, Niue, Norway, Oman, Puerto Rico, Saint Kitts and Nevis, Saint Lucia, Saint Vincent and the Grenadines, San Marino, the Seychelles, Sint Maarten, Slovakia, Slovenia, Sweden, Switzerland, Tokelau, Turks and Caicos Islands, United Arab Emirates, United Kingdom of Great Britain and Northern Ireland, United States of America, US Virgin Islands, Wallis and Futuna Islands, West Bank and the Gaza Strip. The search terms included variations of ‘tuberculosis’, ‘HIV’ and the low-incidence countries; full details of the search strategy and results are shown in Appendix 1. Studies presenting data on multi-country cohorts were excluded if any country contributing patients to the cohort

did not have low TB incidence, unless data was presented separately for each country. No limits were placed on year or language of publication.

The search results were de-duplicated and screening was done by title, abstract and then full-text by two reviewers (Joanne Winter and Aishatu Adamu), with 10% overlap. Review articles, case reports, and original research articles where the outcome was not latent or active TB in a population of PLHIV were excluded. Any disagreements between the reviewers were resolved by consensus. The reference lists of included articles and relevant review articles were screened for additional papers.

The following data were extracted into a form developed in Excel (Microsoft 2010): the outcome(s) studied (LTBI and/or active TB), publication details (author, year and country), details of the study (the time period covered, location, design, and the population of PLHIV that made up the study population), the prevalence of LTBI and the test used to detect it (for studies examining LTBI), the incidence of active TB during the study (for studies examining active TB), the risk factors the study reported on, and the associations between these risk factors and the outcome(s).

A narrative synthesis of the findings is presented. The review was registered on the PROSPERO international prospective register of systematic reviews (CRD42017069544).

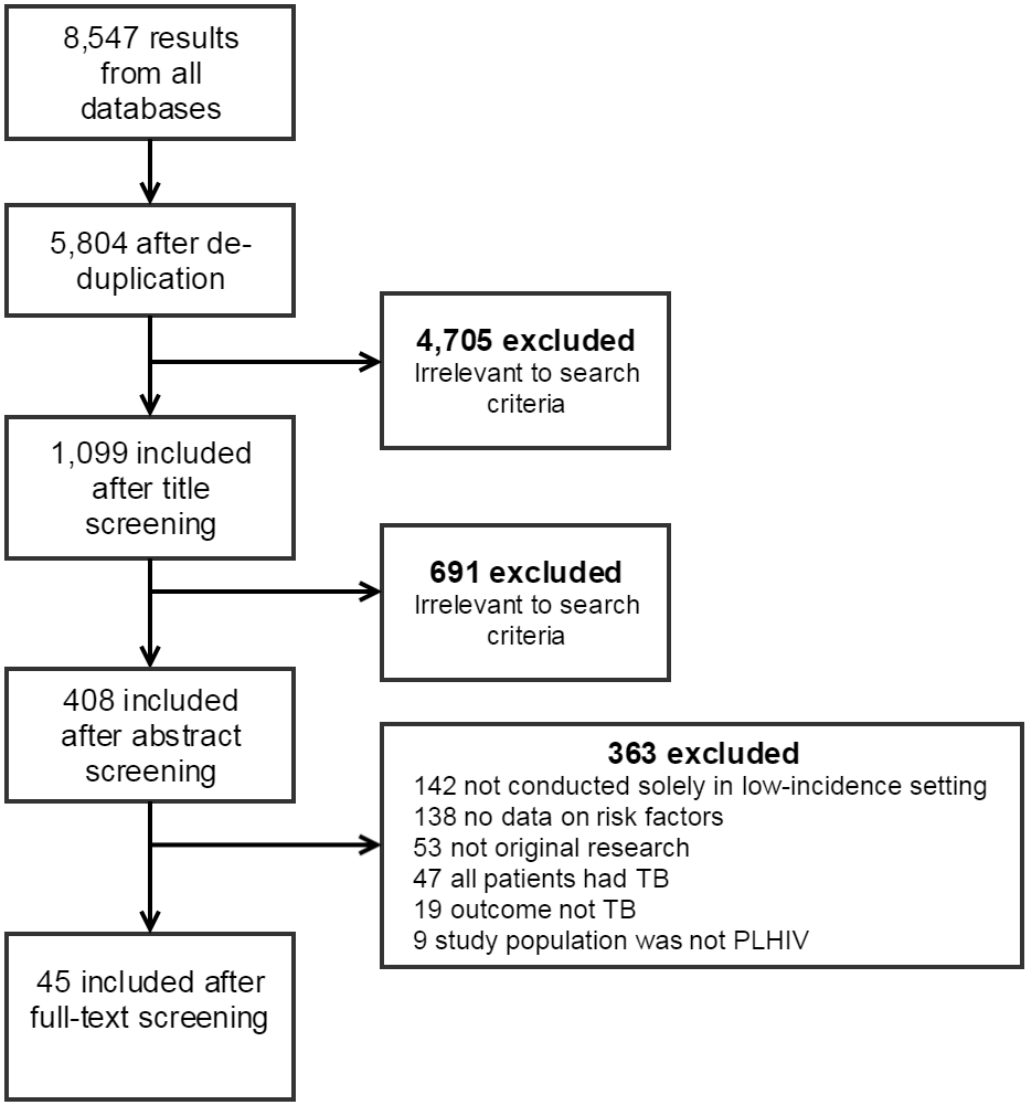
2.3 Results

2.3.1 Search results and included studies

After de-duplication, the search yielded 5,804 results; full search results are shown in Appendix 1. After title, abstract and full-text screening, 45 studies were included (Figure 2.1). Initial reviewer agreement was 99.3%, resolved to 100% after comparing the discrepant papers against the inclusion criteria. Ten studies which reported risk factor data across a mixture of high- and low-burden countries were excluded.

Of the 45 included studies (summarised in Table 2.1), 13 provided data on risk factors for LTBI,[104, 131-146] 28 contained data on risk factors for active TB,[118, 119, 143-172] and 4 studies contained data on both.[143-146] Of the included studies, 34 were cohort studies (17 prospective,[118, 119, 131, 134, 139, 140, 143, 144, 147, 148, 150, 152, 153, 156, 163, 169, 172] 12 retrospective,[133, 146, 149, 151, 155, 159, 164-167, 170, 171] 5 not stated [136, 154, 160-162]), 7 were cross-sectional (5 collected data prospectively,[104, 132, 137, 138, 142] 2 retrospectively [135, 141]). Two studies were case-control studies nested within cohort studies[157, 168] and 3 were outbreak investigations,[145, 156, 158] one of which included a retrospective case-control analysis.[158]

Figure 2.1: Flow chart of the screening process and included studies.



PLHIV: people living with HIV, TB: tuberculosis.

Table 2.1: Studies included in the review.

Author, year, reference	Outcomes studied	Study period	Location	Study design	Study population	LTBI prevalence and test used	TB incidence (per 100,000 PY)	Risk factors examined
Bourgarit, 2015,[131]	LTBI	2009-2011	France	prospective longitudinal cohort	HIV-1 - infected patients	11.8% (TST + QFT-GIT/T-SPOT.TB), 14.7% (TST only)	14.5% over two years of follow-up	Age, sex, country of birth/residence, CD4 count, viral load, BCG vaccination, history of TB, history of LTBI, contact with TB, group/institutional living
Brock, 2006,[104]	LTBI	2004	Copenhagen, Denmark	prospective cross-sectional	HIV-positive patients	4.6% (QFT-GIT)	N/A	age, sex, country of birth, year of HIV diagnosis, AIDS-defining illness, previous TB diagnosis and treatment, country of birth, exposure to smear-positive TB contact, injecting drug use, alcohol misuse, diabetes
Cheallaigh, 2013,[132]	LTBI	2008-2010	Dublin, Ireland	prospective cross-sectional	HIV-1-infected patients	18% (QFT-GIT), 11% (T-SPOT.TB), 10% (TST).	N/A	sex, age, country of birth, TB exposure (contact with TB, high-incidence country of birth, or occupational exposure), drug misuse, homelessness, imprisonment, ART, CD4 count, viral load.
Doshi, 2012,[133]	LTBI	1995-1996	New York City, USA	retrospective cohort	HIV-infected adults	Prevalence 10.2% (TST), incidence 5.3% (1.77/100PY) (TST conversion)	N/A	age, sex, ethnicity, country of birth, CD4 count, ever homeless, group living, viral load, alcohol misuse, drug misuse, imprisonment
French, 2006,[134]	LTBI	1994-2002	USA	prospective longitudinal cohort	HIV-infected and uninfected women	0.8% conversion PY (TST)	N/A	CD4, viral load, time on ART.

Author, year, reference	Outcomes studied	Study period	Location	Study design	Study population	LTBI prevalence and test used	TB incidence (per 100,000 PY)	Risk factors examined
Gampper, 1998,[135]	LTBI	1992-1994	USA	retrospective cross-sectional	HIV-infected individuals	9.1% (TST)	N/A	ethnicity, sex, route of HIV infection.
Girardi, 2002,[136]	LTBI	1997-2000	Italy	cohort	TST-negative HIV patients on HAART	5.4% TST conversion over 3 years	N/A	age, sex, drug misuse, delayed-type hypersensitivity result, AIDS-defining illness, viral load, CD4 count.
Kall, 2012,[137]	LTBI	2006-2009	London, UK	prospective cross-sectional	HIV patients	13.3% (T-SPOT.TB)	No active TB cases during the follow-up period	age, sex, ethnicity, country of birth and TB incidence, CD4 count, ART.
Luetkemeyer, 2007,[138]	LTBI	Not stated	San Francisco, USA	prospective cross-sectional	HIV-positive adults	9.3% (TST) 8.5% (QFT-GIT)	N/A	age, sex, ethnicity, homelessness, injecting drug use, CD4 count, viral load, ART, TST placed, TST result, prior active TB, BCG, country of birth, contact with TB cases.
Mofenson, 1995,[139]	LTBI	1992-1993	USA and Puerto Rico	prospective longitudinal cohort	HIV-infected women	6% (standard TST definition of 5mm+) or 9% (using 2mm+ definition)	N/A	Age, ethnicity, pregnancy, years of education, medicaid health insurance, low income, drug misuse, TB contact, hospitalisation, imprisonment, homeless.
Pullar, 2014,[140]	LTBI	2009-2010	Norway	prospective cohort	HIV-infected adults	26% (QFT-GIT), 25% (TSPOT.TB), 24% (TST)	No active TB cases over 24 months follow-up	Age, sex, country of birth, years resident in Norway, previous AIDS-defining illness, time since HIV diagnosis, CD4 count, ART, TB contact, visit to TB-endemic country, previous TB disease.

Author, year, reference	Outcomes studied	Study period	Location	Study design	Study population	LTBI prevalence and test used	TB incidence (per 100,000 PY)	Risk factors examined
Schulte, 2002,[141]	LTBI	1995-1996	Miami, USA	retrospective cross-sectional	HIV-infected pregnant women	31.9% (TST)	N/A	Age, country of birth, prenatal care in first trimester, unknown HIV at first prenatal visit, ADI, drug misuse, TB exposure.
Stephan, 2008,[142]	LTBI	2006-2007	Frankfurt, Germany	prospective cross-sectional	HIV-positive adult outpatients	12% (TST), 24% (QFT-GIT), 24% (TSPOT.TB)	No active TB cases during the follow-up period	age, sex, ethnicity, country of birth, route of HIV infection, prior AIDS diagnosis, CD4 count, viral load, ART, time since HIV diagnosis, prior TB, imprisonment, social aid, homelessness/sheltered living
Abgrall, 2010,[147]	Active TB disease	1997-2008	France	prospective cohort	HIV-1 or HIV-2 infected patients	N/A	400/100,000	age, sex, country of birth, duration of follow-up, CD4 count, viral load, ART, region of care, prior AIDS, heterosexual route of HIV infection, period of study
Antonucci, 1995,[148]	Active TB disease	1990-1991	Italy	prospective cohort	HIV-infected adults	N/A	2250/100,000	Age, sex, HIV transmission category, place of residence, history of TB, ART, CDC clinical class, CD4, DTH
Dooley, 1992,[149]	Active TB disease	1987-1989	San Juan, Puerto Rico	retrospective cohort	HIV patients admitted in HIV unit	N/A	35,700/100,000 (TB exposed), 3,200/100,000 (TB unexposed)	TB exposure, age, sex, injecting drug use, prior AIDS.

Author, year, reference	Outcomes studied	Study period	Location	Study design	Study population	LTBI prevalence and test used	TB incidence (per 100,000 PY)	Risk factors examined
Girardi, 2000,[150]	Active TB disease	1995-1996	Italy	prospective observational cohort	HIV-infected individuals	N/A	790/100,000	age, sex, country of birth, place of residence, route of HIV infection, history of TB, recent TB contact, TST result, CD4 count, clinical stage, previous isoniazid preventive therapy, ART
Grant, 2009,[119]	Active TB disease	1996-2006	UK	prospective observational cohort	HIV infected adults 16+ years in CHIC centres in UK	N/A	328/100,000	age, sex, route of HIV infection, ethnicity, CD4 count, time on ART, viral load, calendar year.
Gupta, 2015,[151]	Active TB disease	2007-2011	England, Wales and Northern Ireland	retrospective observational cohort	HIV-infected adults on ART	N/A	670/100,000	sex, age, ethnicity, route of HIV infection, CD4 count, viral load.
Hasse, 2014,[152]	Active TB disease	1992-2012	Switzerland	prospective observational cohort	HIV infected adults 16+ years from SHCS	N/A	76/100,000 (currently only ART), 320/100,000 (no previous ART)	age, CD4 count, cumulative co-trimoxazole use, region of origin, injecting drug use
Jones, 2000,[153]	Active TB disease	1996-1998	USA and Puerto Rico	prospective observational cohort	HIV infected adults and adolescents 13+ years	N/A	500/100,000	age, sex, ethnicity, country of birth, route of HIV transmission, CD4 count, ART, isoniazid preventive therapy, clinical disease
Karo, 2014,[118]	Active TB disease	2001-2011	Germany	prospective longitudinal cohort	HIV-infected patients	N/A	370/100,000	sex, country of birth, ART, viral load, CD4 count, route of HIV infection, age, year, co-morbidities, follow-up time
Manavi, 2016,[154]	Active TB disease	2011-2015	Birmingham, England	observational cohort	HIV-infected patients	N/A	281/100,000	age, sex, ethnicity, country of birth, ART, CD4 count, viral load

Author, year, reference	Outcomes studied	Study period	Location	Study design	Study population	LTBI prevalence and test used	TB incidence (per 100,000 PY)	Risk factors examined
Maniewski, 2016,[155]	Active TB disease	2005-2012	Brussels, Belgium	retrospective cohort	HIV patients	N/A	Overall incidence not stated. Incidence declined from 2,900/100,000 to 300/100,000 over study period.	Age, sex, ethnicity, CD4 count, viral load, duration of follow-up, ART
Markowitz, 1997,[156]	Active TB disease	1990-1994	USA	prospective cohort	HIV infected individuals	N/A	700/100,000	sex, ethnicity, route of HIV infection, region, education, CD4 count, TST result
Miguez-Burbano, 2003,[157]	Active TB disease	1998-2000	USA	case control nested in a cohort	HIV-infected individuals	N/A	2,316/100,000	CD4 count, smoking.
Mohle-Boetani, 2002,[158]	Active TB disease	1995	California, USA	outbreak investigation including retrospective case control	HIV-infected male prison inmates	N/A	21,100/100,000	CD4 count, close contact with TB disease, time in day room, TV in cell, prior LTBI treatment, prior TB infection or disease, anergy, school attendance, work, TB inmate in next cell.
Mor, 2013,[159]	Active TB disease	1983-2010	Israel	retrospective cohort	PLHWA	N/A	690/100,000	age, sex, route of HIV infection, year of HIV diagnosis, time from migration to HIV diagnosis.
Moro, 2000,[160]	Active TB disease	1992-1994	Milan, Italy	observational cohort	HIV-infected patients	N/A	1,060/100,000	age, sex, CD4 count.
Pettit, 2016,[161]	Active TB disease	1998-2011	USA and Canada	observational cohort	HIV-1 infected ART-naïve individuals	N/A	65/100,000	ART, age, sex, ethnicity, country of birth, injecting drug use, CD4 count, viral load, duration of follow-up.
Pettit, 2011,[162]	Active TB disease	1998-2008	Nashville, USA	observational cohort	HIV-infected persons	N/A	165/100,000	ART, sex, age, ethnicity, country of birth, route of HIV infection, CD4 count, viral load.

Author, year, reference	Outcomes studied	Study period	Location	Study design	Study population	LTBI prevalence and test used	TB incidence (per 100,000 PY)	Risk factors examined
Pullar, 2014,[163]	Active TB disease	2009-2011	Norway	prospective longitudinal cohort	HIV-infected adults	N/A	No active TB.	age, sex, time since migration, years since HIV infection, CD4 count, ART, LTBI treatment.
Rice, 2013,[164]	Active TB disease	2002-2010	England and Wales	retrospective observational cohort	HIV-infected heterosexual adults	N/A	1,900/100,000	sex, age, ethnicity, CD4 count, country of birth, country of HIV infection, ART.
Rubinstien, 1996,[165]	Active TB disease	1984-1992	Connecticut, USA	retrospective cohort	HIV-seropositive injecting drug users	N/A	600/100,000	LTBI infection, isoniazid preventive therapy.
Sackoff, 2001,[166]	Active TB disease	1995-1997	New York City, USA	retrospective cohort	HIV-infected individuals	N/A	530/100,000	TST result, preventive therapy.
Sterling, 2011,[167]	Active TB disease	1995-2009	USA and Canada	retrospective cohort	HIV-1 infected adults on HAART	N/A	71/100,000	age, sex, ethnicity, ART, CD4 count, viral load, injecting drug use.
Sudre, 1996,[168]	Active TB disease	1981-1994	Switzerland	case control nested in a cohort	HIV-infected individuals	N/A	3.4% of cohort had TB. Incidence not reported.	age, sex, country of birth, route of HIV infection, calendar year, CD4 count, duration of follow-up.
Taarnhoj, 2011,[169]	Active TB disease	1995-2007	Denmark	prospective longitudinal cohort	HIV-1 infected adults (>15 years) living in Denmark when diagnosed with HIV	N/A	820/100,000	sex, age, year, country of birth, CD4 count, viral load, route of HIV infection, alcohol misuse, ART.

Author, year, reference	Outcomes studied	Study period	Location	Study design	Study population	LTBI prevalence and test used	TB incidence (per 100,000 PY)	Risk factors examined
Thomas, 2000,[170]	Active TB disease	1977-1995	New York City, USA	retrospective cohort	HIV-infected children	N/A	610/100,000	year of birth, sex, ethnicity, mother's route of HIV infection, level of immunosuppression (based on CD4 counts)
Trieu, 2010,[171]	Active TB disease	2001-2005	New York City, USA	retrospective cohort	HIV infected individuals	N/A	151/100,000	Age, sex, ethnicity, country of birth
Turkova, 2015,[172]	Active TB disease	1996-2014	UK and Ireland	prospective cohort	HIV infected children	N/A	196/100,000	sex, age, country of birth, ethnicity, region, clinical stage, calendar period, ART, WHO stage, viral load
Aichelburg, 2009,[143]	LTBI and active TB disease	2006-2008	Vienna, Italy	prospective longitudinal cohort	HIV-1 - infected patients	5.3% (QFT-GIT)	840/100,000	LTBI: age, sex, ethnicity, country of birth, CD4 count, ART, viral load, prior TB, AIDS-defining illness, route of HIV infection. TB disease: country of birth, symptoms, CD4, viral load, ART, TST result, AIDS-defining illness
Brassard, 2009,[146]	LTBI and active TB disease	1988-2007	Montreal, Canada	retrospective cohort	HIV-infected adults	14.1% (TST)	181/100,000	LTBI: CD4 count, country of birth, route of HIV infection, ART. TB disease: age, sex, country of birth, time from HIV diagnosis to first clinic visit, HIV infection route, ART, CD4 count, LTBI treatment.

Author, year, reference	Outcomes studied	Study period	Location	Study design	Study population	LTBI prevalence and test used	TB incidence (per 100,000 PY)	Risk factors examined
Elzi, 2007,[144]	LTBI and active TB disease	1996-2006	Switzerland	prospective observational cohort	HIV-infected individuals	9.4% (TST)	220/100,000	Active TB: route of HIV infection, age, region of origin, isoniazid preventive therapy, ART, viral load, sex, CD4 count. LTBI: country of origin, sex, route of HIV infection, age, CD4 count, viral load, ART.
McLaughlin, 2003,[145]	LTBI and active TB disease	1999	South Carolina, USA	longitudinal outbreak investigation	HIV-infected male prison inmates exposed to Mycobacterium tuberculosis	55.1% (TST)	30/323 developed TB within 1 month, 32 within 2 years.	Days exposed to source patients, CD4 count, proximity to source patient, viral load, ART.

AIDS: acquired immunodeficiency syndrome, ART: anti-retroviral therapy, BCG: Bacillus Calmette-Guérin, CDC: Centers for Disease Prevention and Control, DTH: delayed-type hypersensitivity, LTBI: latent tuberculosis infection, QFT-GIT: quantiFERON-gold in-tube, TB: tuberculosis, TST: tuberculin skin test.

The studies spanned the period 1977-2014. One study examining both latent and active TB reported no cases of active TB during the follow-up period and so no data on risk factors for active TB were available.[163]

2.3.2 Risk factors for latent tuberculosis infection

The documented risk factors for LTBI were divided into two categories: sociodemographic and clinical (Table 2.2; shown in more detail in Appendices 2 and 3).

Sociodemographic factors

The literature included data on the following socioeconomic risk factors for LTBI: age, ethnicity, country of birth or residence, sex, pregnancy, route of HIV infection, drug misuse, alcohol misuse, homelessness, imprisonment, level of education, income and occupation.

In terms of ethnicity, LTBI was more common among black, Asian and Hispanic study participants than white participants in low TB incidence countries.[133, 135, 137, 139, 143] These studies were all undertaken in countries where people are predominantly white, except one partially conducted in Puerto Rico,[139] which is primarily Hispanic. Only two studies reported no difference in LTBI risk by ethnicity.[138, 142] Country of birth or residence was a strong predictor of LTBI, which was positively correlated with TB incidence in country of birth or long-term residence [104, 131, 132, 137, 138, 142, 143, 163] and was higher in participants born abroad,[141, 142, 173] particularly those from African countries [173] and medium-risk areas.[133, 144, 146] LTBI risk decreased with time since migration for foreign-born participants.[163] Visiting a TB-endemic country was not associated with increased risk of LTBI.[163]

Table 2.2: Summary of risk factors for LTBI among people living with HIV.

Risk factor for LTBI	Summary of results
Sociodemographic factors	
Age	Greater risk with increasing age: 1 study [139]
	QFT-GIT-positive participants were significantly younger than QFT-GIT-negative participants: 1 study [104]
	No association: 11 studies [131-133, 136-138, 141-144, 163]
Ethnicity	Greater risk for black participants than white participants: 4 studies [135, 137, 139, 143]
	Greater risk for Asian participants than white participants: 2 studies [133, 137]
	Greater risk for Hispanic participants than white participants: 1 study [135]
	No association: 2 studies [138, 142]
Country of birth or residence	Greater risk of LTBI in participants who were born abroad: 4 studies [133, 141, 143, 144]
	Greater risk of LTBI in participants born in Africa: 3 studies [143, 144, 146]
	Greater risk for participants who were born or long-term residents in countries with high TB incidence: 7 studies [104, 131, 132, 137, 138, 142, 163]
	Risk of TB decreased with time since migration: 1 study [163]
	No association with visiting a TB-endemic country: 1 study [163]
Sex	Greater risk for women: 2 studies [131, 143]
	Greater risk for men: 1 study [144]
	No association: 9 studies [104, 132, 133, 135-138, 142, 163]
	Single-sex studies: 4 studies [134, 139, 141, 145]
Pregnancy	No association: 1 study [139]
	All participants were pregnant: 1 study [141]
Route of HIV infection	Higher rates of TST positivity among people who inject drugs and/or have a partner who injects drugs: 2 studies [135, 144]
	Higher rates of TST positivity among people who acquired HIV through heterosexual sex than by sex between men: 1 study [144]
	Significant, but not described, differences in risk by infection route: 1 study [146]
	No association: 2 studies [142, 143]
Drug misuse	Higher rates of illicit drug use among TST-positive participants than TST-negative participants: 1 study [139]
	No association between injecting or recreational drug use and LTBI: 8 studies [104, 132, 133, 136, 138, 141-143]
Alcohol misuse	No association: 2 studies [104, 133]

Risk factor for LTBI	Summary of results
Homelessness	Greater risk among those who were homeless, or lived in group/institutional living or a homeless shelter: 3 studies [131, 133, 139]
	No association: 3 studies [132, 138, 142]
Imprisonment	No association: 4 studies [132, 133, 139, 142]
	All participants were in prison: 1 study [145]
Level of education	No association: 1 study [139]
Income	No association between income and LTBI: 1 study [139]
	No association between receiving public social aid and LTBI: 1 study [142]
Occupation	No association: 1 study [132]
Clinical factors	
CD4 count and DTH result	Higher rates of LTBI at higher CD4 counts: 7 studies [133, 136, 142-144, 146, 163]
	No association between CD4 count and LTBI: 7 studies [104, 131, 132, 134, 137, 138, 145]
	No association between a positive DTH reaction and LTBI: 1 study [136]
Viral load	Greater risk with high viral load: 1 study [143]
	Lower risk with viral load: 1 study [144]
	No association: 9 studies [104, 131-134, 136, 138, 142, 145]
Anti-retroviral therapy	Lower risk of LTBI for participants on ART: 3 studies [134, 143, 144]
	No association: 8 studies [104, 132, 137, 138, 142, 145, 146, 163]
	All patients were on ART: 1 study [136]
BCG vaccination	No association: 2 studies [131, 138]
Previous TB disease	Greater risk of LTBI for patients with prior active TB disease: 6 studies [104, 131, 138, 142, 143, 163]
Latent tuberculosis infection	Greater risk of current LTBI for patients with a history of LTBI: 2 studies [131, 138]
Previous treatment for LTBI	Participants with LTBI were more likely to have been previously treated for LTBI: 1 study [138]
Contact with TB, or other close exposure to TB	Greater risk for those living or working in a homeless shelter, prison or drug rehabilitation unit: 1 study [138]
	Greater risk for participants with a contact with TB or smear-positive TB: 5 studies [104, 131, 139, 141, 163]
	Greater risk for prisoners closer to the source patient of an outbreak within a prison: 1 study [145]
	No association: 2 studies [132, 138]
AIDS-defining illness	Higher risk of LTBI for patients with prior AIDS: 1 study [142]
	No association: 4 studies[104, 136, 141, 163]

Risk factor for LTBI	Summary of results
Other non-infectious co-morbidities	No association between diabetes and LTBI: 1 study [104]
Time since HIV diagnosis or duration of follow-up time	No association: 4 studies [104, 141, 142, 163]

AIDS: acquired immunodeficiency syndrome, ART: anti-retroviral therapy, BCG: Bacillus Calmette-Guérin, DTH: delayed-type hypersensitivity, LTBI: latent tuberculosis infection, QFT-GIT: quantiFERON-gold in-tube, TB: tuberculosis, TST: tuberculin skin test.

No studies reported a definitive association between age and LTBI, although the majority of studies were restricted to adults, and one was restricted to children.[172] Ten studies reported no association between sex and LTBI and four were restricted to either men or women; however two studies reported men were more likely to be interferon-gamma release assay (IGRA)-positive [131, 143] and one older study found that women were less likely to be TST-positive.[144] One study reported no association between pregnancy and TST positivity among women.[139] HIV acquisition through injecting drug use, heterosexual sex or an unknown route was associated with higher prevalence of LTBI than MSM, and having a partner who injected drugs was also associated with increased risk.[135, 139, 144] However, two other studies found no association between route of HIV infection and LTBI,[142, 143] and eight other studies found no association between injecting or recreational drug use and LTBI.[104, 132, 133, 136, 138, 141-143].

Collective living in a group or institution was positively associated with IGRA-positivity,[131] and TST conversion rates were higher among people who were homeless or had ever lived in a shelter.[133, 139] However, three other studies found no association between homelessness or living in a shelter and LTBI.[132, 138, 142] Living or working in a homeless shelter, prison or drug rehabilitation unit was associated with higher rates of TST positivity but not QuantiFERON-TB Gold In-Tube (QFT-GIT) positivity.[138] However, in other studies, imprisonment was not associated with TST positivity,[139] TST conversion,[133, 142] or QFT-GIT or T-SPOT.TB positivity.[132, 142] One study of TST conversion rates during an outbreak of TB in a

prison found that rates of TST conversion were not significantly different between the dormitory of the source patient and other dormitories over the same time period.[145] Neither alcohol misuse nor years of education were associated with LTBI. Income was not associated with LTBI.[139, 142] One study found no association between occupational exposure and LTBI, although 'occupational exposure' was not defined.[132]

Clinical factors

The available literature included data on the following risk factors: CD4 count and DTH result, viral load, anti-retroviral therapy, BCG vaccination, previous TB disease, LTBI, previous treatment for LTBI, contacts with TB, AIDS-defining illnesses, other non-infectious co-morbidities and time since HIV diagnosis.

Higher CD4 count was associated with higher rates of TST positivity [134, 144, 146] and TST conversion,[136] whilst an increase in CD4 count from the baseline measurement was associated with higher odds of TST conversion compared to participants whose CD4 count had decreased.[133] However, three other studies found no significant association.[134, 142, 145] Nadir CD4 count >350cells/µl was associated with higher rates of TST and QFT-GIT positivity in two studies,[138, 163] and higher CD4 counts were associated with QFT-GIT positivity.[132, 142, 143] Other studies reported no association between CD4 count and LTBI.[104, 131, 132, 137, 142] Only one study reported higher median viral load among participants with a positive QFT-GIT result,[143] all others reported no significant association between viral load and TST positivity,[134, 138, 142, 144, 145] TST conversion,[133, 136] or IGRA-positivity.[104, 131, 132, 138, 142] The majority of studies found no significant association between ART and LTBI diagnosed by TST,[138, 142, 145, 146, 163] QFT-GIT,[104, 132, 138, 142, 163] or T-SPOT.TB.[132, 137, 142] There was no significant association between prior TB disease and TST positivity,[138] but all studies which used IGRAs to diagnose LTBI found higher rates of IGRA positivity among patients with previous active TB disease than patients who had not previously had TB.[104,

131, 138, 142, 143, 163] A history of LTBI was associated with IGRA positivity [131] and prior TST positivity was associated with current TST positivity and QFT-GIT positivity.[138]

Exposure to TB disease was an important risk factor for having LTBI. Eight studies reported data on exposure and only two reported no association with LTBI.[132, 138] Having a contact with active TB was associated with higher rates of IGRA-positivity[104, 131, 163] and TST positivity,[141] whilst another study found that more TST-positive participants had a close contact with TB disease than TST-negative participants.[139] The odds of TST conversion during a TB outbreak in a prison were increased for patients on the same side of the dormitory as the source patient and patients who were exposed for a longer time period.[145]

BCG vaccination did not appear to be associated with LTBI among PLHIV; both studies which examined BCG as a risk factor utilised both IGRAs and the TST and found no association using either test.[131, 138] Prior AIDS was not associated with LTBI,[104, 136, 141, 163] although one study found that a higher proportion of TST-positive participants had a prior AIDS diagnosis than TST-negative participants.[142] Diabetes was not associated with a positive QFT-GIT result.[104] Time since HIV diagnosis was not significantly associated with LTBI risk.[104, 142, 163]

2.3.3 Risk factors for tuberculosis disease

The documented risk factors for TB disease were also divided into sociodemographic and clinical factors (Table 2.3; shown in more detail in Appendices 4 and 5).

Sociodemographic factors

Data were reported on the following sociodemographic risk factors for active TB: age, ethnicity, country of birth or residence, region of country of birth, sex, route of HIV infection, drug misuse, alcohol misuse, homelessness, imprisonment and level of education.

Black or black African ethnicity was consistently associated with higher risk of active TB than for people of white ethnicity.[119, 145, 151, 155, 161, 162, 164, 167, 170, 171] South Asian [151, 164], Hispanic [167, 170, 171] and 'other' [119, 151, 167, 170] ethnicities were also associated with higher rates of TB, although black Caribbean ethnicity was not.[164] Being born in Asia,[152, 169] Africa [146, 152, 169] or sub-Saharan Africa,[118, 144, 146, 147, 154] or south America [152] was associated with a higher risk of TB than being born in the low-incidence countries in which these studies were conducted (Canada, Denmark, France, Germany and Switzerland). Five studies reporting country of birth as foreign-born or not all found a higher risk of TB for foreign-born patients.[153, 161, 162, 171, 172] Risk of TB increased with the TB incidence in country of birth,[154] and with increasing time interval from entry to the study country to HIV diagnosis.[159]

There was generally no association between age and risk of TB, although the only study which included children reported a significantly lower risk for children aged 0-5 years than the baseline of 5-10 years.[172] Data on the risk of TB associated with sex were mixed. Two studies reported a lower risk of TB for men than women,[146, 170] but several others reported higher rates of TB for men than women.[153, 159, 162, 164, 165, 169]

Table 2.3: Summary of risk factors for active TB disease among people living with HIV.

Risk factor for active TB	Summary of results
Sociodemographic factors	
Age	Risk increases with age: 2 studies [152, 159]
	Risk decreases with age: 1 study [146]
	Risk was highest in the middle age categories: 3 studies [151, 171, 172]
	No association: 17 studies [118, 119, 144, 147-150, 153-155, 160-162, 164, 167-169]
Ethnicity	Greater risk for black or black African participants than white participants: 10 studies [119, 145, 151, 155, 161, 162, 164, 167, 170, 171]
	Greater risk for Asian participants than white participants: 3 studies [151, 164, 171]
	Greater risk for Hispanic participants: 2 studies [167, 171]
	No association: 4 studies [153, 154, 156, 172]
Country of birth or residence	Greater risk among participants who were born or acquired HIV abroad: 8 studies [146, 147, 153, 161, 162, 164, 171, 172]
	Greater risk among participants from Africa: 5 studies [118, 144, 147, 152, 169]
	Greater risk among participants from Asia: 2 studies [152, 169]
	Greater risk among participants from south America: 1 study [152]
	Greater risk for participants from countries with higher TB incidence: 2 studies [154, 168]
	Risk of TB increased with time since arrival for foreign-born participants: 1 study [159]
	No association: 2 studies [143, 150]
Region	Risk of TB varied by region of country: 1 study [147]
	No association: 4 studies [148, 150, 156, 172]
Sex	Greater risk for men: 5 studies [153, 159, 162, 164, 165]
	Greater risk for women: 2 studies [146, 170]
	No association: 15 studies [118, 144, 148-151, 155, 156, 160, 161, 167-169, 171, 172]
	Risk presented grouped with route of HIV infection: 2 studies [119, 147]
	Single-sex studies: 1 study [158]

Risk factor for active TB	Summary of results
Route of HIV infection	Greater risk for people who acquired HIV through heterosexual sex than from sex between men: 7 studies [119, 147, 151, 153, 159, 162, 169]
	Greater risk for people who acquired HIV infection through injecting drug use than from sex between men: 5 studies [147, 151, 153, 159, 169]
	Greater risk for people who acquired HIV through mother-to-child transmission than from sex between men: 1 study [151]
	Greater risk among children with mothers who acquired HIV through injecting drug use: 1 study [170]
	No association: 7 studies [118, 144, 146, 148, 150, 156, 168]
	All participants were heterosexual: 1 study [164]
Drug misuse	Greater risk for participants who inject drugs: 3 studies [152, 161, 167]
	Greater risk for participants who used tobacco long-term: 1 study [157]
	No association: 1 study [149]
	All participants had a history of prior or current drug use: 2 studies [157, 165]
Alcohol misuse	No association: 1 study [169]
Imprisonment	All participants were in prison: 2 studies [145, 158]
Level of education	No association: 1 study [156]
Clinical factors	
Year of diagnosis	Greater risk among those who entered the cohort later: 1 study [147]
	Greater risk of TB if diagnosing with HIV during the ART era than prior to ART being introduced: 1 study [159]
	No association: 3 studies [119, 169, 172]
CD4 count	Greater risk at lower CD4 count: 22 studies [118, 119, 143, 144, 147, 148, 150-156, 158, 160-162, 164, 167-170]
	No association: 3 studies [145, 146, 157]
Viral load	Greater risk with higher viral load: 9 studies [118, 119, 143, 144, 147, 151, 154, 161, 162]
	No association: 6 studies [145, 153, 155, 167, 169, 172]
Anti-retroviral therapy	Lower risk for patients on ART: 10 studies [118, 143, 144, 147, 150, 153, 161, 162, 164, 169]
	Lower risk with increasing time on ART: 4 studies [119, 147, 161, 162]
	No association: 7 studies [145, 146, 148, 154, 155, 167, 172]
	All patients were on ART: 1 study [151]
Previous TB disease	Greater risk for patients with a history of TB: 1 study [148]
	No association: 1 study [158]
LTBI infection and anergy	Greater risk for patients with known LTBI: 6 studies [144, 148, 150, 156, 165, 166]
	Greater risk for anergic patients: 1 study [148] No association: 1 study [158]

Risk factor for active TB	Summary of results
Previous treatment for LTBI	Lower risk for patients who had been treated for LTBI: 4 studies [144, 146, 165, 166]
	No association: 2 studies [153, 158]
Contact with TB or other close exposure to TB	Greater risk for patients who had been exposed to active TB: 3 studies [145, 149, 158]
AIDS-defining illness or other opportunistic infections	Greater risk for patients with prior AIDS: 2 studies [147, 153]
	Greater risk for patients with more clinical symptoms of AIDS: 1 study [143]
	No association: 4 studies [148-150, 172]
	Lower risk for patients taking prophylactic co-trimoxazole: 1 study [152]
Time since HIV diagnosis or duration of follow-up time	Decreasing risk over time: 2 studies [118, 147]
	Greater risk for patients with more follow-up time: 1 study [168]
	No association: 1 study [155]

AIDS: acquired immunodeficiency syndrome, ART: anti-retroviral therapy, BCG: Bacillus Calmette-Guérin, LTBI: latent tuberculosis infection, QFT-GIT: QuantiFERON-gold in-tube, TB: tuberculosis, TST: tuberculin skin test.

Three studies reported no association between route of HIV infection and risk of tuberculosis.[148, 150, 168] Other studies reported that MSM had lower risk of TB than people who acquired HIV infection through heterosexual sex or other routes of infection,[118, 119, 146, 151, 153, 159, 169] injecting drug use [151, 153, 156, 159, 169] or mother-to-child transmission.[151] One study of children reported HIV infection route for the mother, and reported that children with TB were more likely to have mothers who acquired HIV through injecting drug use than through heterosexual sex or other routes of transmission, compared to HIV-infected children without TB.[170] Injecting drug use was associated with increased risk of TB,[152, 161, 167] as was long term tobacco use.[157] The risk of TB was not associated with alcohol misuse [174] or level of education.[156]

Clinical factors

Data were reported on the following clinical risk factors for active TB: year of HIV diagnosis, CD4 count, viral load, anti-retroviral therapy, previous TB disease, LTBI,

previous treatment for LTBI, exposure to TB, AIDS-defining illnesses and time since HIV diagnosis.

Lower CD4 counts were overwhelmingly associated with greater risk of developing TB.[118, 119, 144, 147, 148, 150, 152-154, 156, 164, 167-169] Studies reported that patients with active TB had significantly lower median CD4 counts [143, 155, 161, 162] or CD4 percentage[160, 162] during the study period, although lower nadir CD4 count was not associated with TB.[143] Four studies reported no association between CD4 count and risk of TB.[145, 146, 157, 158] Only one study examined differences in risk associated with CD4 counts >500 cells/µl, finding that compared to CD4 count >700 cells/µl, the risk of TB was only lower for patients with CD4 counts <500, or for patients with unknown CD4 count.[151]

Fifteen studies reported data on the relationship between viral load and TB disease. Viral loads of ≥40,[154] ≥50,[119, 151] and ≥10,000 [118, 147] were associated with increased risk of TB in adjusted analyses, whilst other studies demonstrated that TB patients had significantly higher HIV-1 RNA levels than PLHIV without TB.[143, 144, 161, 162] Six studies reported no association between active TB and viral load.[145, 153, 155, 167, 169, 172] The risk of TB was lower for patients who were on ART,[118, 144, 147, 150, 153, 164, 169] and further decreased for patients who had been on ART for more than 6 months[147, 161, 162] or more than two years.[119] Six studies reported no significant association between active TB and ART [145, 146, 148, 154, 155, 167] or duration of ART.[167] Two studies reported higher risk of TB in the first six months [147] and first year [118] of follow-up. The decrease in risk of TB over time was less steep for migrants than for non-migrants.[147]

The CDC classifies HIV infection into clinical stages according to criteria based on CD4 count <200 cells/µl and symptomatic conditions (including tuberculosis). Four studies reported that having AIDS (as defined by the CDC, but excluding TB) was not significantly associated with risk of TB;[148-150, 172] two of these studies had adjusted

for ART.[150, 172] One study reported a higher risk of TB among patients with clinical AIDS,[153] whilst another found that prior AIDS was associated with an increased risk of TB for non-migrant patients, but there was no significant association for migrants.[147]

Three studies reported data on exposure to a TB patient and risk of TB. A study in an HIV unit of a hospital reported a greatly increased risk of TB among exposed patients than unexposed patients.[149] In two case-control studies in prisons, TB patients were more likely to have had contact with a TB case and spent more time in the communal day room than control patients,[158] and patients on the same side of the dormitory as the source patient were more likely to have TB.[145]

Anergic patients and patients with a positive TST result had significantly higher risk of TB than TST-negative and/or non-anergic patients.[144, 148, 150, 156, 165, 166] Only one study reported no TB cases among TST-negative participants.[166] Prior treatment of LTBI was associated with lower risk of developing active TB in four studies,[144, 146, 165, 166] and longer duration of preventive therapy was more protective.[166] However, two other studies found no association.[153, 158] Prior active TB disease was associated with a higher risk of TB in one study,[148] but a second study reported no association.[158]

2.4 Discussion

2.4.1 Summary of results

Sociodemographic risk factors were largely consistent for LTBI and active TB disease among PLHIV in countries with low TB incidence. For example, risks were greater for people of black ethnicity (particularly black African) but also elevated for Asian and Hispanic study participants, compared to white participants. Rates of LTBI and TB disease were higher among foreign-born patients, irrespective of their country of birth, but were highest among patients from medium and high risk areas (i.e. sub-Saharan Africa) and correlated with TB incidence in country of birth. The risk of testing

positive for LTBI decreased with time since migration. The risk of active TB decreased over the duration of the studies, although one study demonstrated that the risk decreased less for foreign-born patients.

2.4.2 Strengths and limitations of the study design and findings

The systematic search strategy enabled a comprehensive review of the literature. Due to the high number of records found by the search, 10% of results were screened by both reviewers. Between-reviewer agreement was high and the discrepant articles were generally excluded after reviewing them against the inclusion criteria. Accordingly, the search strategy demonstrated a high sensitivity.

For the purposes of the inclusion criteria, countries were defined as 'low incidence' based on the TB incidence of TB in 2015; this was a pragmatic choice as many studies spanned several years and TB incidence changed over time. Consequently, studies may have been included from periods where the TB incidence was above 10/100,000 population, or excluded studies conducted where TB incidence was less than 10/100,000 prior to 2015 but not in 2015. This is particularly relevant for small countries, where a small change in the number of cases reported could result in a substantial fluctuation in incidence. Multi-country studies were excluded if not all the countries had low TB incidence and data were not presented stratified by country, which led to the exclusion of some large studies of active TB incidence among PLHIV. However, the findings of these studies were consistent with those of this review; the HIV-CAUSAL study, the EuroSIDA cohort and the ART Cohort Collaboration reported that higher TB incidence was associated with low CD4 count, high viral load, HIV acquisition by injecting drug use or heterosexual sex (compared to MSM), and being born in Africa.[175-178]

The papers included in this study were of varying quality. The majority were large observational cohorts with data collected prospectively or routinely. Several were representative of whole countries, particular populations within countries, or specific

cities. Some studies, particularly older studies, had only small numbers of participants or small numbers of participants developing TB, which limited their ability to detect associations particularly in multivariable analyses.[139, 156-158] The findings of these studies were in line with those of the larger and more rigorous studies; however, there were some risk factors which were only examined in these studies. These findings (that there was no association between pregnancy and TST positivity;[139] whilst the risk of active TB was associated with smoking tobacco,[157] but not with level of education [156]) should therefore be interpreted with caution.

The magnitude of associations between the various risk factors and LTBI or active TB varied substantially due to the considerable heterogeneity in the design and statistical analyses of the included studies, particularly variation in how variables were categorised, which category was used as the baseline, and which other variables (if any) were adjusted for. There were no clear differences in the directions of associations between risk factors and the outcome in studies which conducted multivariable analyses and those which only reported univariable associations; accordingly, this does not appear to be a source of bias.

2.4.3 Sociodemographic risk factors

Both LTBI and active TB were more likely to occur in study participants who had acquired their HIV infection through heterosexual sex, injecting drug use or mother-to-child transmission than by sex between men. It is likely that routes of HIV infection other than sex between men are confounders for exposure to TB, which was an important risk factor for both LTBI and active TB, but one which few studies examined directly. Similarly, few studies prospectively examined the risk of LTBI and active TB disease among PLHIV with socioeconomic risk factors, this would be a useful subject for further research so that testing and preventive strategies can be appropriately targeted. Mixed results were reported regarding the significance of the association between being homeless or living in a shelter and LTBI, but working in a homeless

shelter was also associated with greater risk of LTBI.[138] Proximity to persons with TB was also important; close contact with a TB case increased the risk of LTBI and active TB, and a study set in a prison demonstrated that closer proximity was associated with increased risk. Known LTBI was also associated with increased risk of TB disease, although this was lessened by preventive therapy.

Globally, men are disproportionately affected by TB and women disproportionately affected by HIV,[1, 26] however global estimates of TB-HIV co-infection have not been stratified by sex.[1] In this review, data on the effect of sex on risk of TB among PLHIV was mixed. Of the nine studies which reported no association between sex and risk of LTBI, none adjusted for other factors such as route of HIV infection, which may have biased associations of risk of TB towards women; people who acquire HIV through heterosexual sex usually have higher rates of TB than MSM, as black African ethnicity and birth in countries with high TB incidence can be confounders for route of HIV infection. Consequently, it is likely that these studies underestimated the relative risk of TB among men with HIV, compared to women. Of the seven studies which did report differences in risk of TB among PLHIV, five reported that the risk was higher among men than women, in line with the higher risk globally of TB among men than women. The higher risk among men may be attributable to several factors, including greater exposure (from more social mixing) and a higher prevalence of smoking, alcohol abuse, imprisonment and silicosis among men.[179, 180] There may also be immunological or physiological differences which mean men are more susceptible than women.[179] There is also a larger gap between TB prevalence and reporting of TB cases among men than women,[1] suggesting that men face greater difficulties accessing care for TB. If this is also the case for accessing HIV care, men with HIV may be at greater risk of developing TB than women with HIV.

The risk of both LTBI and active TB was greatest among PLHIV from ethnic minorities (particularly black African) and those born in countries with high TB incidence; both of which are proxy measures for exposure to TB. To decrease TB in

these populations, prompt identification and treatment of both LTBI and active TB are necessary. Clinicians should be particularly aware of the risks of TB in these high-risk groups so that any patients from these populations presenting with symptoms compatible with TB can be diagnosed promptly. More active screening of high-risk populations such as black Africans, those born in high-incidence countries and PWID for LTBI, and treating LTBI were appropriate, could significantly reduce the TB burden in these populations. Earlier diagnosis of HIV would also enable earlier initiation of ART, increase CD4 counts and contribute to TB prevention. As the groups of PLHIV who are most at risk for developing TB also have a high risk of HIV, concurrent screening for HIV, LTBI and active TB could be an effective measure for preventing TB and diagnosing HIV infection sooner.

One study observed that the risk of TB among migrants decreased with increasing time since entry to the country.[159] This may be attributable to a lower risk of new infection with *MTBC* following migration from a high- to low-incidence country, and also to the increased risk of TB among people who were infected with HIV abroad;[164] as the likelihood of HIV infection remaining undiagnosed decreases with time, the risk of TB decreases as patients become aware of their HIV infection and initiate ART. Pre-entry screening of migrants from countries with high TB incidence for pulmonary TB is now required for migrants moving to the UK, Australia, Canada, New Zealand and the USA, and other countries could consider adopting these guidelines to reduce the incidence of active TB. However, this cannot prevent reactivation of LTBI from causing TB disease after migration, or TB among migrants who arrive via irregular routes. Promoting screening for HIV infection and LTBI to migrants from countries with high TB incidence, particularly those who have recently entered the country, could be an effective measure to diagnose HIV sooner and prevent LTBI from developing into active disease.

2.4.4 Clinical risk factors

There were differences in the clinical factors that were associated with LTBI and active TB disease. There was no consistent association between CD4 count or ART and LTBI; however high CD4 count and initiation of ART were both strongly protective against active TB disease. High viral load was also associated with increased risk of TB.

Low CD4 count was not associated with an increased risk of LTBI among PLHIV, and some studies reported higher rates of LTBI in participants with high CD4 counts. This suggests that tests for LTBI are not sufficiently sensitive at low CD4 counts, something which is well documented for TSTs but less so for IGRAs. This lack of sensitivity is because a limitation of tests for LTBI is that they measure the immune response to tuberculosis infection, rather than directly detecting the presence of *MTBC*, and therefore can be impeded when a person with HIV is immunosuppressed. This difficulty diagnosing LTBI at low CD4 count means that it is not established as to whether HIV increases the risk of becoming infected with TB (as opposed to progression to active disease, the evidence for which is definitive). However, as a substantial proportion of HIV infections are diagnosed late, and LTBI tests appear less sensitive at low CD4 counts, simultaneous screening for HIV and LTBI would allow a negative TST or IGRA to be assessed in the context of CD4 count. Patients newly diagnosed with HIV could then be re-tested for LTBI once they have initiated ART and their CD4 count has increased.

A study of PWID with known dates of HIV seroconversion also reported low CD4 count as a strong predictor of TB disease. Compared to the first three years of HIV infection, in this population the risk of TB was around three times higher 4-6 years after HIV infection and five times higher by nine years after HIV infection, although the risk was not significantly different in the intermediate period.[181] This contrasts with the results presented in this review where the risk of TB was highest immediately following HIV diagnosis. This can probably be explained by the fact that many PLHIV

are not diagnosed with HIV immediately following infection, and half of diagnosed PLHIV across Europe are diagnosed late (CD4 count ≤350 cells/μl).[129]

Prior treatment of LTBI was associated with a significantly reduced risk of developing TB in 4 of the 6 studies which examined it as a risk factor; supporting the results of a recent meta-analysis which found that treating LTBI reduced the risk of active TB disease, and also found no conclusive evidence that treatment efficacy was reduced for PLHIV.[182] One study identified in this chapter, which found no association, was of an outbreak within a prison.[158] As TB disease in this study was predominantly being driven by recently acquired infection, LTBI treatment would not be expected to be protective, as LTBI treatment prevents reactivation of LTBI but would not prevent a new infection from occurring. The other study which did not report a significant association did not report the number of participants who had received prior treatment for LTBI, although they stated that only one person who was prescribed preventive therapy developed active TB.[153] Consequently, it is possible that the risk of TB was decreased for PLHIV who had been treated for LTBI, but that the study lacked power to detect the association; or that the relationship was confounded by factors which influenced which participants were treated for LTBI, such as LTBI test result and CD4 count.

Two studies examined the relationship between prior active TB disease and the risk of active TB; an observational cohort reported that prior TB was associated with a higher risk of TB, even after adjusting for LTBI (diagnosed by TST);[148] however a study of an outbreak resulting from a single source patient in a prison found no association between prior TB disease and risk of TB.[145] This suggests that prior TB increases the risk of active TB either because it is a marker for other risk factors relating to risk of exposure to TB or progression from LTBI to active TB. These factors could include higher background rates of TB, low CD4 count, unsuccessful treatment of prior disease or development of drug resistance during treatment of the previous disease.

2.4.5 Conclusions

To conclude, it is important to note that risk factors for TB in PLHIV in low TB incidence countries can be divided into two categories; those which are unalterable (but useful for screening purposes) and those which are alterable. Important risk factors in the former category include black African ethnicity, birth in a country with high TB incidence (particularly sub-Saharan Africa) and HIV acquisition through injecting drug use or heterosexual sex; whilst alterable risk factors include CD4 count, viral load and initiation of ART. More screening for HIV, TB, and particularly LTBI needs to be done in high-risk groups of PLHIV such as PWID and recent migrants from countries with high TB incidence. Simultaneously screening for HIV and LTBI among migrants, particularly those who have recently migrated, and promoting LTBI treatment to eligible PLHIV could be effective mechanisms for reducing TB among people with HIV.

3 Record linkage between tuberculosis and HIV surveillance data to identify co-infected individuals

3.1 Introduction

3.1.1 Surveillance data for use in epidemiological analyses

Public health surveillance is defined by the WHO as '*the continuous, systematic collection, analysis and interpretation of health-related data needed for the planning, implementation and evaluation of public health practice*'.[183] Routine collection of electronic health records provides an opportunity to undertake epidemiological analyses; however, linking records from separate datasets is frequently necessary to maximise the usefulness of such studies. For example, laboratory results can be linked to clinical data,[184, 185] and databases of different infections can be linked to identify patients with co-morbidities or who have died.[123, 164, 186]

Global TB surveillance is coordinated by the WHO; the 194 member states report annually on the number of TB case notifications, demographic factors (age and sex) and clinical factors (site of disease, HIV status [when available], drug resistance and treatment outcomes).[1] Countries in the EU report data through The European Surveillance System (TESSy) managed by the ECDC. TB incidence is estimated from notifications of TB cases; however notification data are of varying quality, there can be under-reporting (where TB is undiagnosed or not notified), over-reporting (if there is duplication of case notifications) or misclassification (e.g., recurrent TB may be misclassified as a new case). Estimates of incidence are calculated at the country level and reported with uncertainty bounds, adjusting the case notification rate using data from surveys quantifying over- and under-reporting, TB prevalence surveys, mortality data, and expert opinion.[187]

Similarly, global surveillance on HIV is coordinated jointly by the WHO and the Joint United Nations Programme on HIV/AIDS (UNAIDS).[188] Since 2000, HIV surveillance strategies have been tailored to the needs of each country; with the aims

of concentrating resources where they are most useful, collecting data from key populations at risk of HIV, and tracking changes in the epidemic over time. In countries with generalised HIV epidemics (HIV prevalence >1% in the general population), surveillance systems aim to monitor HIV infection and risk behaviours in the general population. In concentrated epidemics (HIV prevalence >5% in sub-populations) and low-level epidemics (HIV prevalence >1% in sub-populations), surveillance systems concentrate on monitoring infections in those populations, and high-risk behaviours in the general population.[189] As with TB, data on HIV and AIDS case notifications are reported by each country to the WHO/UNAIDS, through TESSy for EU countries.[190] TESSy collects data on demographic variables (including age, sex, country of birth and nationality) and clinical factors (diagnosis date, CD4 counts, viral load, ART, probable route of HIV infection and date and cause of death [if relevant]). Even within the EU the quality and coverage of surveillance systems are varied; therefore statistical adjustments are made to account for reporting delays which improves the precision of estimates of recent trends in HIV epidemiology.[190]

3.1.2 Tuberculosis and HIV surveillance data collection in the UK

In the UK, surveillance data on TB and HIV are collected separately, by the Respiratory and HIV & Sexually Transmitted Infections departments, respectively, at PHE.[191, 192] UK TB surveillance systems do not collect HIV data, because personally identifiable information (PII) is recorded and HIV status is considered too sensitive to include in such a dataset. The UK's HIV surveillance systems record data on TB as an AIDS-defining illness, but this field only contains a subset of TB diagnoses (where a clinical decision has been made that the patient has AIDS) and does not record date of TB diagnosis or any clinical data relating to TB disease. Therefore, to identify all co-infected individuals and carry out detailed epidemiological analyses, it is necessary to link the two surveillance datasets.

3.1.2.1 Tuberculosis surveillance

TB is a notifiable disease in the UK and has been since 1912.[78] Statutory notifications of all new cases of active TB disease are made by staff at TB clinics through the Enhanced Tuberculosis Surveillance system (ETS, which includes the London TB Register) in England, Wales and Northern Ireland, and through Enhanced Surveillance of Mycobacterial Infections (ESMI) in Scotland (Figure 3.1). Reporting of cases is considered to be comprehensive, particularly for patients treated by the National Health Service (NHS); there may be underreporting of TB patients who are seen in private clinics, but this is estimated to be a negligible number of cases (personal communication, Public Health England).

The information reported includes notification details (patient identifiers and notification date), and demographic, clinical and microbiological data. Information on treatment outcomes is collected 12 and 24 months from the treatment start date, or the notification date if treatment start date is not available. Information on isolates identified as part of the *M. tuberculosis* complex are also collected by ETS, including culture confirmation of species, drug sensitivity results, and some clinical and demographic information.

3.1.2.2 HIV surveillance

Prior to 2015 in England, Wales and Northern Ireland, data on all individuals aged 15 or older living with a diagnosed HIV infection who accessed care at NHS services was collected through the Survey of Prevalent HIV Infections Diagnosed (SOPHID). Information on new diagnoses of HIV infection, AIDS and death, from clinicians and microbiologists in England, Wales and Northern Ireland was collected through the HIV & AIDS New Diagnoses Database (HANDD). Scottish data is collected separately by Public Health Scotland (Figure 3.1). The Recent Infection Testing Algorithm (RITA) programme monitors the number of patients who had recently acquired HIV infection at the time of diagnosis in England, Wales and Northern Ireland. The CD4 Surveillance Scheme monitors immunosuppression in adults living with HIV in

England, Wales and Northern Ireland by collecting longitudinal data on the CD4 count of HIV patients.

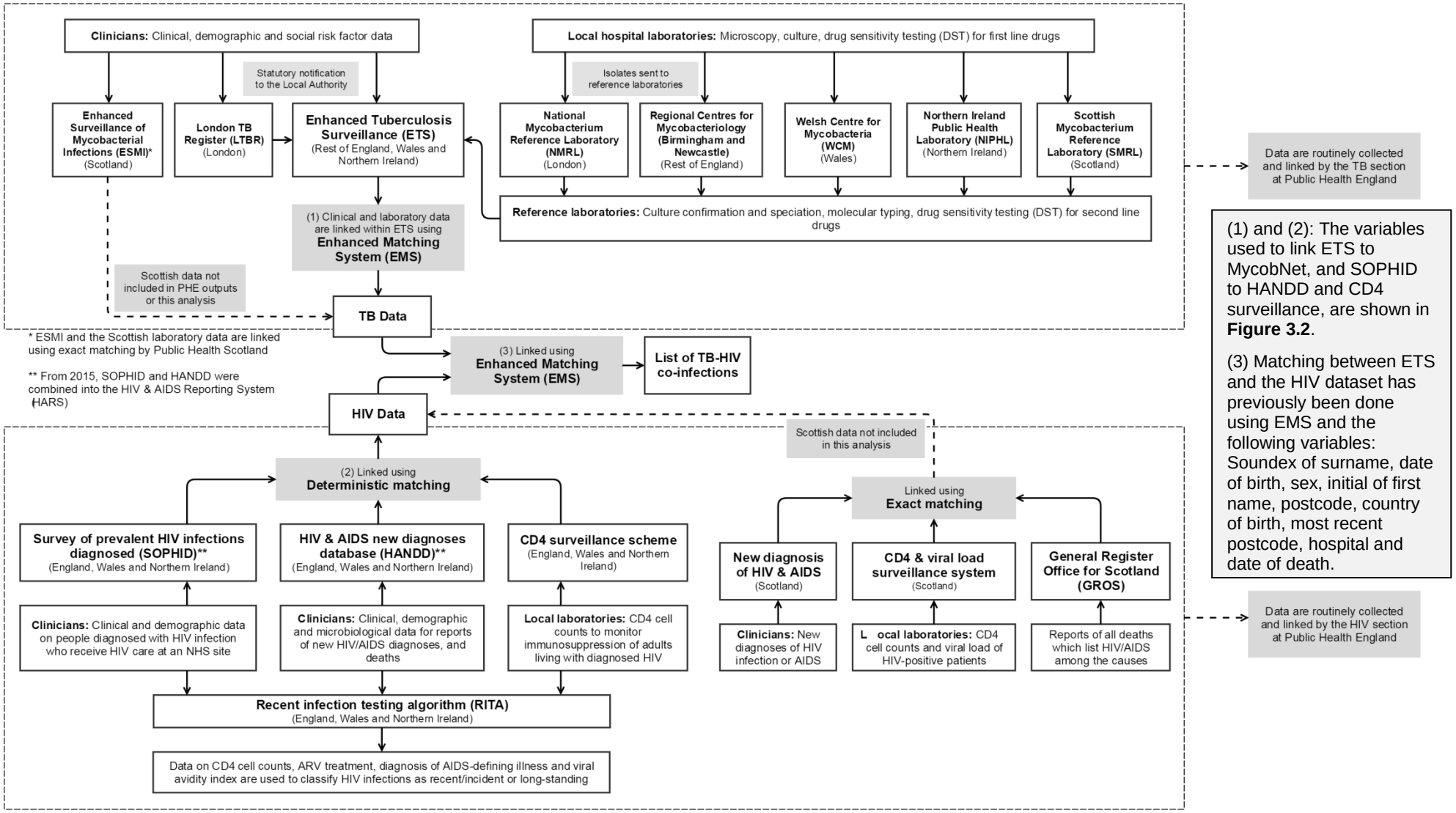
Whilst HIV infection is not a notifiable disease in the UK, the funding that NHS healthcare facilities receive is dependent on the reporting of new diagnoses of HIV and patients seen for HIV care to PHE; correspondingly the datasets are considered to be indicative of all patients diagnosed with HIV, or seen for HIV care. However, the national surveillance datasets cannot include the estimated 13% of HIV infections in the UK that were undiagnosed in 2016.[10]

From 2015, HIV surveillance data in England has been collected through the HIV and AIDS Reporting System (HARS), which collects detailed patient demographic details, clinical and risk factor data, including data about clinic attendance, and HIV diagnosis and treatment. HARS collects data on all patients who have been confirmed as HIV antibody positive (Figure 3.1).

3.1.2.3 Data protection and patient consent

All databases contain minimal patient identifiers, are under strict data protection and are held securely at PHE.[191] Patient consent for data linkage and the subsequent epidemiological analyses was not required, as PHE has authority under the Health and Social Care Act 2012 to hold and analyse national surveillance data for public health and research purposes.

Figure 3.1: The flow of data in TB and HIV surveillance systems in the UK, until 2014.



(1) and (2): The variables used to link ETS to MycobNet, and SOPHID to HANDD and CD4 surveillance, are shown in **Figure 3.2**.

(3) Matching between ETS and the HIV dataset has previously been done using EMS and the following variables: Soundex of surname, date of birth, sex, initial of first name, postcode, country of birth, most recent postcode, hospital and date of death.

3.1.3 Methods for record linkage

There are three commonly used methods of matching records between datasets; exact matching, deterministic matching and probabilistic linkage. Exact matching requires all records in both datasets to have a unique identifying variable, such as NHS number. Probabilistic and deterministic methods of record linkage are both dependent on the quality and completeness of the original data. They are described in more detail below.

3.1.3.1 Probabilistic matching

Probabilistic record linkage utilises the probabilities of agreement and disagreement between variables to determine whether or not two or more records are a match; *i.e.* whether they belong to the same person.[184] Probabilistic linkage can account for errors and omissions of data in variables used in the matching process and records can be linked without the need for a unique identifier such as NHS number. However, its main disadvantage is the need to define a threshold score below which records are deemed to not be matches, or undertake a significant amount of manual review above and below the threshold value. Manual review introduces subjectivity which limits the repeatability of the process, particularly as those doing the manual review are normally removed from the clinical situation and therefore do not have any personal insight into whether or not two records are a match. This may bias the results of the matching process by introducing non-random error. In addition, manual review is very time-consuming, and therefore costly.

3.1.3.2 Deterministic Matching

Deterministic record linkage is the process of linking records based on rules of agreement. Deterministic matching algorithms are often structured hierarchically. For example, records may initially be matched based on the agreement of several key fields such as NHS number, first name and surname, address, and date of birth/death. If unsuccessful, this can be followed by less restrictive criteria, such as a combination of parts of the above data in combination with demographic information such as ethnicity or country of birth.

Deterministic matching has many advantages; it is straightforward to understand and easily repeatable, with no time-consuming manual review. However, it is limited in that there is no way to weight matches, and the rigid 'rules' which define whether records match or not mean that any small discrepancies or errors in the data will result in matches being missed.

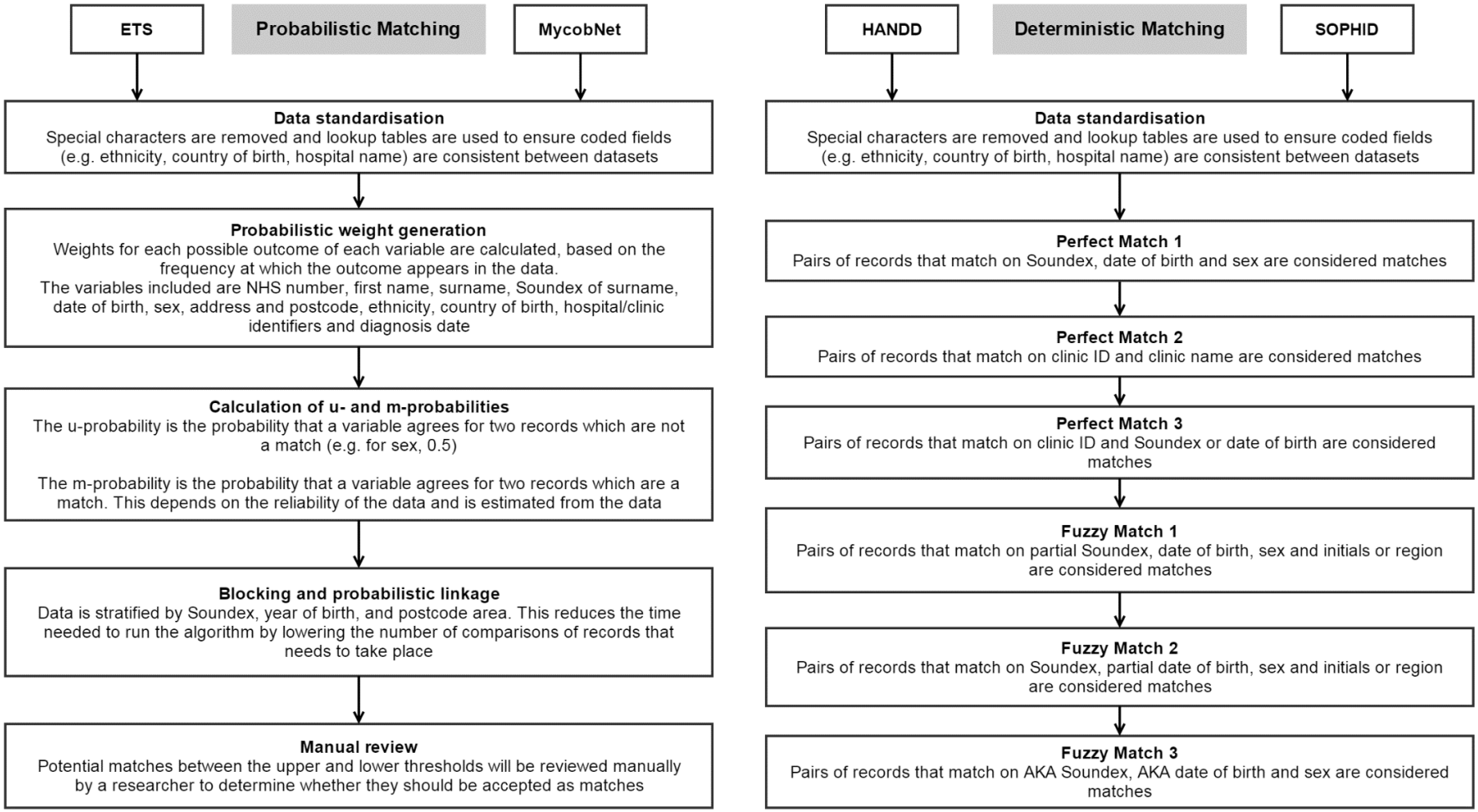
3.1.3.3 Considerations when choosing a method of record linkage

It is important to consider the purpose of linking records when deciding which method of record linkage to use. If any clinical decisions are to be made on the basis of the linked data, then the specificity of the methods used is the most important consideration. For surveillance purposes, precise estimates of the burden of number of patients and the use of a consistent methodology are both crucial. For epidemiological analyses, it is more important that the dataset is representative of the underlying population. Purely deterministic methods of record linkage can be inappropriate for epidemiological analyses, as patients missing data may be systematically different from patients with complete data and this could bias the dataset. As an example, hard-to-reach patients may be more likely to have missing data, and therefore may be less likely to be linked if a purely deterministic algorithm was used for record linkage. This could lead to underestimation of the prevalence of co-infection among these patients.

3.1.4 Record linkage at Public Health England

At PHE, both the TB and HIV sections use record linkage to match records from different data sources to produce their final datasets. The methods are described in more detail below, and summarised in Figure 3.2.

Figure 3.2: The steps involved in probabilistic matching (using EMS) to link TB datasets and deterministic matching to link HIV datasets, at Public Health England



AKA: also known as, EMS: enhanced matching ETS: enhanced tuberculosis surveillance, HANDD: HIV and AIDS new diagnoses database, NHS: National Health Service, SOPHID: survey of prevalent HIV infections diagnosed.

3.1.4.1 Tuberculosis record linkage

The Enhanced Matching System (EMS) is a probabilistic matching algorithm developed by the Tuberculosis section at PHE.[184] EMS was developed primarily to link notifications of TB cases to reports of culture-positive isolates from TB reference laboratories, due to low recording rates of unique identifiers (specifically, National Health Service [NHS] number, a unique 10-digit patient identifier assigned to each patient registered with the NHS).

EMS generates weights for matches that reflect the probability of agreement and disagreement between variables. The variables utilised can vary depending on which are available, but routine linkage between ETS and laboratory data utilises patient identifiers such as names, date of birth/death, sex, address variables, ethnicity, country of birth, hospital/clinic identifiers, and diagnosis dates. Where a unique identifier such as NHS number is available, this is also included. Agreement weights (the probability that two matching records are in fact a match) depend on the commonality of specific variable values (for example, rare surnames will be weighted more strongly than very common ones) and data quality, whilst disagreement weights reflect the probability of random agreement of two variables. The weights for each variable are summed to give an overall 'match score' for each possible pairing of records. A threshold score, at which there is a 99% probability that two records belong to the same person, is calculated. Pairs of records with a score greater than this threshold are accepted as matches and assumed to belong to the same patient, whilst pairings below the threshold are rejected. Manual review of a limited number of records above and below the threshold value can be done to refine the process.

EMS has been validated against a gold-standard identifier (NHS number) using the subset of ETS data for which NHS number was available, demonstrating a sensitivity of up to 99.5% and specificity of up to 100.0% when probabilistic matching was combined with manual review.[184]

3.1.4.2 HIV record linkage

Internal matching within the HIV surveillance system between SOPHID and HANDD was done using a hierarchical deterministic matching system which first attempts to link records using combinations of Soundex code (a four-digit characterisation of a surname which cannot be back-translated),[193] date of birth, sex, clinical ID number and clinic name. Records not successfully linked using these criteria were then matched, where possible, using combinations of Soundex (or part of), date of birth (or part of), sex, initials, region and 'Also Known As' ('AKA') Soundex (where people have changed their name or use more than one name) or 'AKA' date of birth (where it has been verified that the date of birth on previous records was incorrect).

3.1.4.3 Previous record linkage between the tuberculosis and HIV datasets

Until 2011, matching the TB and HIV datasets was done using the probabilistic system EMS,[98, 164] using the following fields: Soundex of surname, date of birth, sex, initial of first name, postcode, country of birth, most recent postcode, hospital, date of death. However, a number of issues make linking the TB and HIV datasets more difficult than the internal matching of different datasets within the TB or HIV surveillance systems.

The lack of a unique identifier, such as NHS number, for even a subset of the HIV surveillance data has prevented validation of the methods used to match records between the HIV and TB datasets against a gold-standard.

Even without manual review, EMS has very high sensitivity (99.3%) and specificity (99.9%) when matching ETS data to other TB surveillance data. However, a sensitivity analysis demonstrated that omission of address variables from the matching process resulted in a decreased sensitivity of 97.1%,[184] although specificity remained high at 100.0%. For TB-HIV matching the sensitivity of EMS is likely to be even lower as there are fewer variables which can be used. HIV surveillance data includes little PII (no NHS numbers, names or addresses) due to the confidentiality requirements of HIV surveillance and clinical data recording. Individuals are identified

by a Soundex code, the initial of their forename and their postcode. This limits the ability of EMS to successfully link the two datasets.

The HIV surveillance system collects other fields that are useful in matching, e.g. AKA Soundex, AKA date of birth and postcode over time (longitudinal data is collected on people living with HIV), which cannot be used in probabilistic matching algorithms for logistical reasons relating to the amount of computational power necessary.

When linking records within the HIV or TB surveillance systems, the majority of individuals would be expected to appear in both datasets. This simplifies the matching process as it is more likely that two records belong to the same person if they are reasonably similar across key fields. However, when matching between the TB and HIV datasets expect a much lower proportion of overlap would be expected, reflecting the proportion of co-infected individuals (previously around 5-10%).[123] Identifying individuals expected to be in both the HIV and TB datasets is difficult; ETS does not collect data on HIV status (the only HIV variable collected being the offer of a HIV test), and data on TB as an AIDS-defining illness in the HIV data is not always well completed.

Some fields that are useful in internal matching, such as clinic ID, cannot be used for TB-HIV matching as many patients receive HIV and TB treatment at different locations. Diagnostic dates are also of limited use because TB and HIV diagnoses are not always made concurrently. Other problems arise due to the demographics of the populations with TB and HIV. Patients in both populations can be very transient, so postcodes are not necessarily consistent even over short periods of time, and some individuals have no fixed abode and therefore no postcode. Postcodes are therefore a useful field for confirming records belong to the same person, but disagreeing postcodes should not be used to discount a match.

3.1.5 Summary and objectives

TB and HIV surveillance data are collected separately in the UK, and the two datasets therefore need to be linked in order to identify co-infected patients and conduct epidemiological analyses. The lack of a unique identifier between the TB and HIV surveillance datasets means that exact matching is not possible for UK surveillance data. There were three options for linking the two datasets; probabilistic matching, deterministic matching, or a combination of the two. Given the lack of an effective, accurate and repeatable method, a primary aim of this PhD was to investigate potential alternatives to the use of EMS and develop a more efficient and repeatable methodology.

3.2 Methods

3.2.1 Data

For the period 2001-2011, the TB and HIV datasets collated by PHE had previously been linked using EMS.[184] To determine the most efficient, accurate and repeatable methodology for future matching between the TB-HIV datasets, the two raw datasets used for matching 2001-2011 data were retrieved. These datasets were used to make comparisons between the results of the previous probabilistic linkage using EMS (i.e. the list of HIV and TB records matched by the algorithm) and the results of the alternative methods that were investigated; a deterministic matching algorithm and subsequently a hybrid method which utilised both probabilistic and deterministic techniques.

3.2.2 Investigation of deterministic matching

The first objective was to determine if deterministic matching was a reasonable alternative to EMS for linking records between the TB and HIV datasets. A deterministic algorithm was developed using expertise from both database managers and epidemiologists within the HIV and TB sections at PHE to identify potential criteria for defining a match. The criteria used various combinations of the following fields: Soundex of surname, date of birth, sex, initial of first name, ethnicity, country of birth,

postcode and date of death. Partial fields (initial of Soundex, day/month/year of birth and postal district) and alternative fields (AKA Soundex and AKA date of birth) were also used. These criteria are shown in Table 3.1.

Table 3.1: Hierarchy of the matching criteria developed for the deterministic algorithm.

Criteria	Soundex (AKA)	Date of birth (AKA)	Year of birth	Sex (can be unknown)	Initial	Initial of Soundex	Country of birth	Ethnicity	Postcode	Postal district	Date of death
1	y	y		y	y				y		
1a	(y)	y		y	y				y		
1b	y	(y)		y	y				y		
1c	(y)	(y)		y	y				y		
2	y	y		y	y						
3	y		y	y					y		
3a	(y)		y	y					y		
4	y			y							y
4a	(y)			y							y
5	y	y		(y)			y	y		y	
5a	(y)	y		(y)			y	y		y	
5b	y	(y)		(y)			y	y		y	
5c	(y)	(y)		(y)			y	y		y	
6	y	y		y							
6a	(y)	y		y							
6b	y	(y)		y							
6c	(y)	(y)		y							
7	y		y	y	y					y	
7a	(y)		y	y	y					y	
8		y		y	y	y	y				
8b		(y)		y	y	y	y				

AKA: also known as. Brackets indicate that the criteria was an alternative criteria (AKA Soundex or AKA sex), or that it was considered a match if the item was unknown in one dataset.

As HIV surveillance data does not include NHS number, there was no unique identifier available for even a subset of records, meaning that the matching criteria could not be validated against a gold-standard identifier. Comparisons were made between the number of matches found and the number of matches missed (approximations of sensitivity and specificity) between the different techniques. Where matches were found using both methods, the average (mean) and distribution of the probabilistic score was calculated for each of the matching criteria. This gave an insight of the approximate range of scores found by each deterministic criterion. This was used to refine the matching criteria.

3.2.3 Development of hybrid method

A 'hybrid' matching method was developed, using a combination of probabilistic and deterministic methods. Two thresholds of match score were defined for the purposes of accepting/rejecting probabilistic matches. The upper threshold was the probabilistic score at which there was a 99% probability that two matched records belonged to the same person;[184] the lower threshold was the score at which the probability was 50%. Matches with a probabilistic score above the upper threshold were accepted and matches below the lower threshold were rejected. Deterministic criteria were used in place of manual review to decide whether matches with scores between the upper and lower thresholds should be accepted or not.

Where matches were found by both probabilistic and deterministic matching methods, the average probabilistic score for matches found by each deterministic criterion was calculated. Deterministic criteria with an average score below the upper threshold were excluded, except for criteria which included an AKA Soundex or AKA date of birth, which would be expected to have significantly lower scores. This is because AKA Soundex and AKA date of birth are not included in the probabilistic matching algorithm, and so agreement on these fields cannot increase the probabilistic score of a match. Additionally, Soundex and date of birth are two of the most heavily weighted fields in the probabilistic algorithm, meaning that disagreement on these

fields (because they agree on AKA Soundex/date of birth) are strongly weighted negatively.

The average scores of each criterion were used in combination with expertise from PHE epidemiologists to determine the order in which the criteria should be used, in order to create a hierarchical algorithm. If a record was not matched using the first criterion, matches were attempted using subsequent criteria in the order shown in Table 3.1.

Additional criteria were added to refine the final list of matches and exclude matches which could not belong to the same person, based on diagnosis dates of TB, HIV infection and AIDS and dates of death from both datasets. HIV records with TB as an AIDS-defining illness were given a 'TB flag', which was used to identify HIV records that were expected to match to the TB data.

As part of the algorithm development process, a limited amount of manual review was done to evaluate the final combination of probabilistic cut-off scores and deterministic criteria, particularly to investigate matches that were ranked highly by probabilistic or deterministic methods but missed by the other.

3.3 Results

3.3.1 Deterministic matching algorithm

The final deterministic matching criteria were based on combinations of all variables that were common to both the TB and HIV datasets: Soundex of surname, sex, date of birth, initial of first name, country of birth, ethnicity, postcode and date of death (Table 3.1).

Deterministic matching alone was not considered to be accurate enough for TB-HIV matching, as many more matches (n=1,084, Figure 3.3) were found using the deterministic algorithm than were found using EMS, which were not considered close enough matches upon manual inspection. However, investigating the criteria that these additional matches were selected on (Table 3.2) demonstrated that many were very

strong matches; particularly the 185 pairs of records which had matched on Soundex, date of birth, sex, initial and postcode (criterion 1). It was evident that a number of records which clearly matched strongly were not being detected by EMS. As a result, the use of a hybrid method incorporating both probabilistic and deterministic techniques was investigated.

Figure 3.3: Comparison of matching results from EMS (with manual review of borderline matches) and the deterministic matching algorithm.

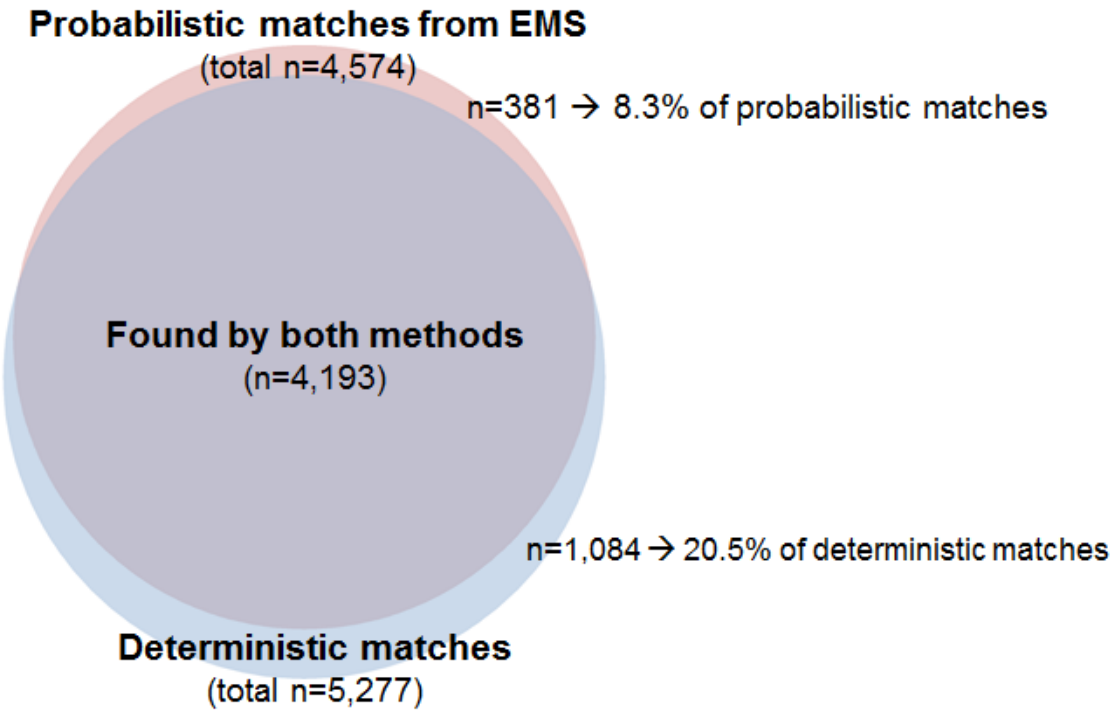


Table 3.2: The deterministic matching criteria of the 1084 additional matches found by the deterministic matching algorithm.

Criterion	n	%	Cumulative %
1	185	17.1	17.1
1a	25	2.3	19.4
2	124	11.4	30.8
3	122	11.3	42.1
3a	18	1.7	43.7
4	28	2.6	46.3
4a	1	0.1	46.4
6	430	39.7	86.1
6a	48	4.4	90.5
7	41	3.8	94.3
7a	5	0.5	94.7
8	57	5.3	100.0
Total	1,084	100.0	

n: number. Full explanations of each criterion are shown in Table 3.1

Of the 2,248 HIV patients with a report of TB as an AIDS-defining illness between 2001 and 2011, 1,534 (68.2%) were linked to a TB case in ETS using the deterministic matching algorithm. This was slightly higher than the 1,471 (65.4%) which were linked to a TB case using EMS with manual review.

3.3.2 Description of a hybrid matching method

The number and percentage of matches which were found using each criterion, and the mean probabilistic score of the matches, are shown in Table 3.3. All criteria had a mean probabilistic score of greater than the upper threshold of 27.88, except for those using AKA Soundex or AKA date of birth.

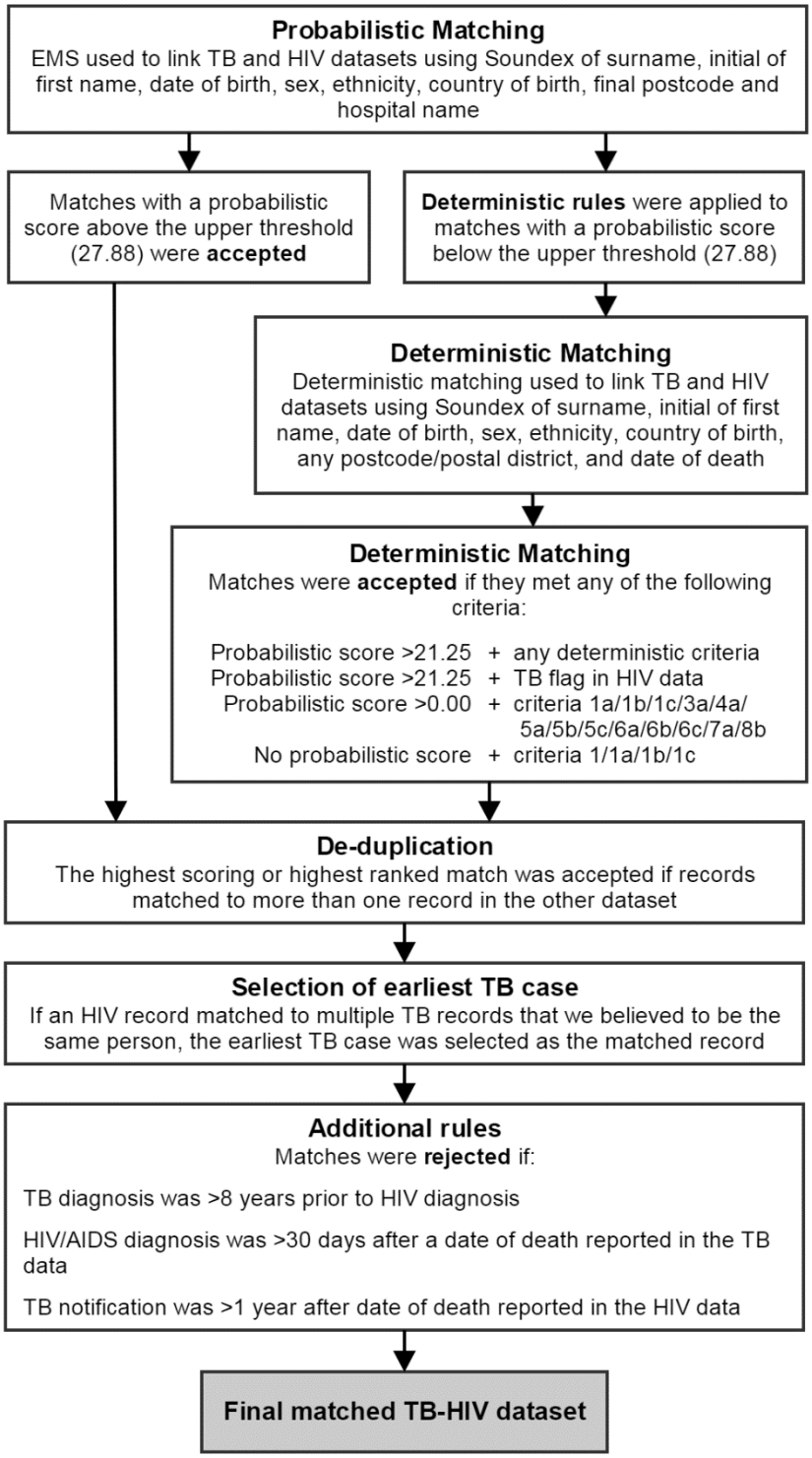
Table 3.3: The number and percentage of matches found by each of the deterministic matching criteria used in the hybrid matching algorithm, and the average probabilistic score for each criterion, using data for the period 2001-2011.

Criteria	Number of matches		Mean probabilistic score
	n	%	
1	2336	51.8	47.7
1a	179	4.0	33.6
1b	0	0.0	-*
1c	0	0.0	-*
2	1006	22.3	34.6
3	433	9.6	43.8
3a	31	0.7	27.7
4	50	1.1	35.5
4a	4	0.1	17.9
5	4	0.1	41.3
5a	0	0.0	-*
5b	0	0.0	-*
5c	0	0.0	-*
6	286	6.3	32.4
6a	70	1.6	20.1
6b	0	0.0	-*
6c	0	0.0	-*
7	41	0.9	32.5
7a	6	0.1	11.0
8	63	1.4	38.7
8b	0	0.0	-*
Total	4509	92.9	

*** could not be calculated. n: number. 382/4,891 (7.1%) of all matches were found by probabilistic matching only and therefore no match type was available**

The criteria for accepting pairs of records into the final dataset as matches were based on a combination of probabilistic score and deterministic criteria (Figure 3.4). Pairs of records with a probabilistic score above the 99% probability threshold (27.88) were accepted as matches. Pairs of records with a match score between 21.25 (the lower threshold, at which the probability of two records being a match is 50%) and 27.88 were accepted as matches if they were also matched deterministically, or if the HIV record was flagged as having TB as an AIDS-defining illness. Pairs with a score below 21.25 were accepted only if they matched on a criteria utilising AKA Soundex or AKA date of birth (criteria 1a/1b/1c/3a/4a/5a/5b/5c/6a/6b/6c/7a/8b). Records matched on criteria 1/1a/1b/1c were considered ‘perfect’ matches and accepted regardless of whether they had been matched probabilistically or not.

Figure 3.4: Flow chart of the hybrid matching process.



27.88 was the score at which there was a 99% probability that two matched records belonged to the same person.
21.25 was the score at which there was a 50% probability that two matched records belonged to the same person.
Deterministic matching criteria are detailed in Table 3.3.

Where records matched to more than one record in the other dataset, the highest scoring or highest ranked match was accepted. If an HIV record matched to multiple TB records believed to belong to the same person then the earliest TB case was used as the match; recurrent cases of TB were not matched. Matches were discounted if the TB diagnosis was more than 8 years before the HIV diagnosis, if HIV or AIDS diagnosis date was >30 days after the date of death reported in the TB data, or if the TB notification was more than one year after the date of death reported in the HIV data (Figure 3.4).

As part of the algorithm development process, hybrid records were manually reviewed if they met any of the following criteria: probabilistic score >50 and not deterministically matched, probabilistic score <21.25 and matched using criteria 1 or 1a, probabilistic score <15 and had a TB flag in the HIV record. All of these records were considered acceptable matches and no adaptations were made to the algorithm as a result.

3.3.3 Comparison of the hybrid method to EMS, using data for the period 2001-2011

Using data from the period 2001-2011, the hybrid matching method found 4,891 matches, compared to 4,574 using EMS with manual review of borderline cases (Figure 3.5). 4,308 matches were found by both the hybrid method and EMS, representing 94.2% of the matches found probabilistically and 88.1% of the matches found using the hybrid method. Of the additional 583 matches found using the hybrid method, 29% matched on the highest criteria (1/1a: Soundex/AKA Soundex, date of birth, sex, initial of first name and postcode) whilst 60% matched on the top three criteria and a further 28% matched on criteria 6 (Soundex, date of birth and sex) (Table 3.4). Criteria 6 is considered a ‘perfect match’ in internal HIV matching, but is not specific enough alone for TB-HIV matching as many of the matches found using this criteria had low probabilistic scores due to disagreement on other fields such as ethnicity and country of birth. Of the 266 ‘not matched’ records from the probabilistic

matching results, 142 were where multiple TB records had been matched to a single HIV record, and therefore were not considered as having been missed by the new method.

Of the 2,248 HIV patients who had TB reported as an AIDS-defining illness, 1,580 (70.2%) were linked to a TB case in ETS using the hybrid method. This was higher than the 1,534 (68.2%) and 1,471 (65.4%) found by the deterministic and probabilistic methods individually. Of the 668 patients with TB as an AIDS-defining illness who were not linked to ETS, 20 had probabilistic scores below the lower threshold and the remaining 648 (97.0%) were not linked at all using the probabilistic algorithm.

Figure 3.5: Comparison of matching results from the hybrid method, and probabilistic matching (EMS, with manual review of borderline matches).

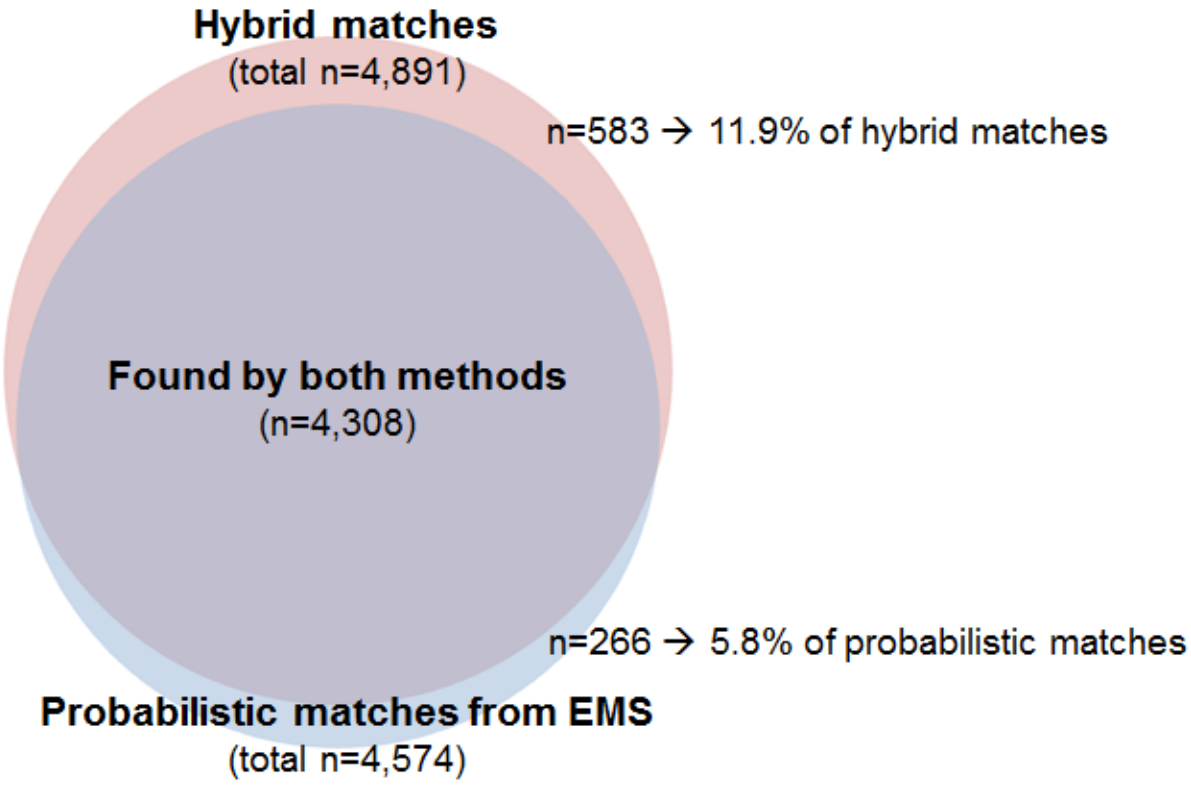


Table 3.4: Distribution of the deterministic match types of 471/583 additional matches found using the hybrid method.

Match type	n	%	Cumulative %
1	122	25.9	25.9
1a	15	3.2	29.1
2	82	17.4	46.5
3	51	10.8	57.3
3a	14	3.0	60.3
4	19	4.0	64.3
4a	1	0.2	64.5
6	85	18.1	82.6
6a	48	10.2	92.8
7	22	4.7	97.5
7a	5	1.1	98.5
8	7	1.5	100.0
Total	471	100.0	

n: number. Only 471/583 additional matches found using the hybrid method had a match type available, others were accepted at lower probabilistic scores because they had a TB flag.

3.3.4 Matching results for the period 2000-2014

Following the development of the hybrid method of record linkage, individuals co-infected with TB and HIV were identified from all TB cases notified to ETS from 2000 to 2014 and all patients reported to HARS as having been seen for HIV care between 2000 and 2014. Both datasets of patients included adults (≤15 years) in England, Wales and Northern Ireland only.

For this period, there were 112,660 TB cases and 124,588 HIV patients. The hybrid matching method resulted in 5,810 matches. Of these, 5,737 (98.7%) matched probabilistically, with a mean match weight of 43.8 (standard deviation 12.8). 5,386 (92.7%) were matched using deterministic criteria. The distribution of the hierarchical deterministic criteria used for the deterministic matches is shown in Table 3.5.

Table 3.5 Distribution of the match types of the 5,386 deterministic matches identified for the period 2000-2014.

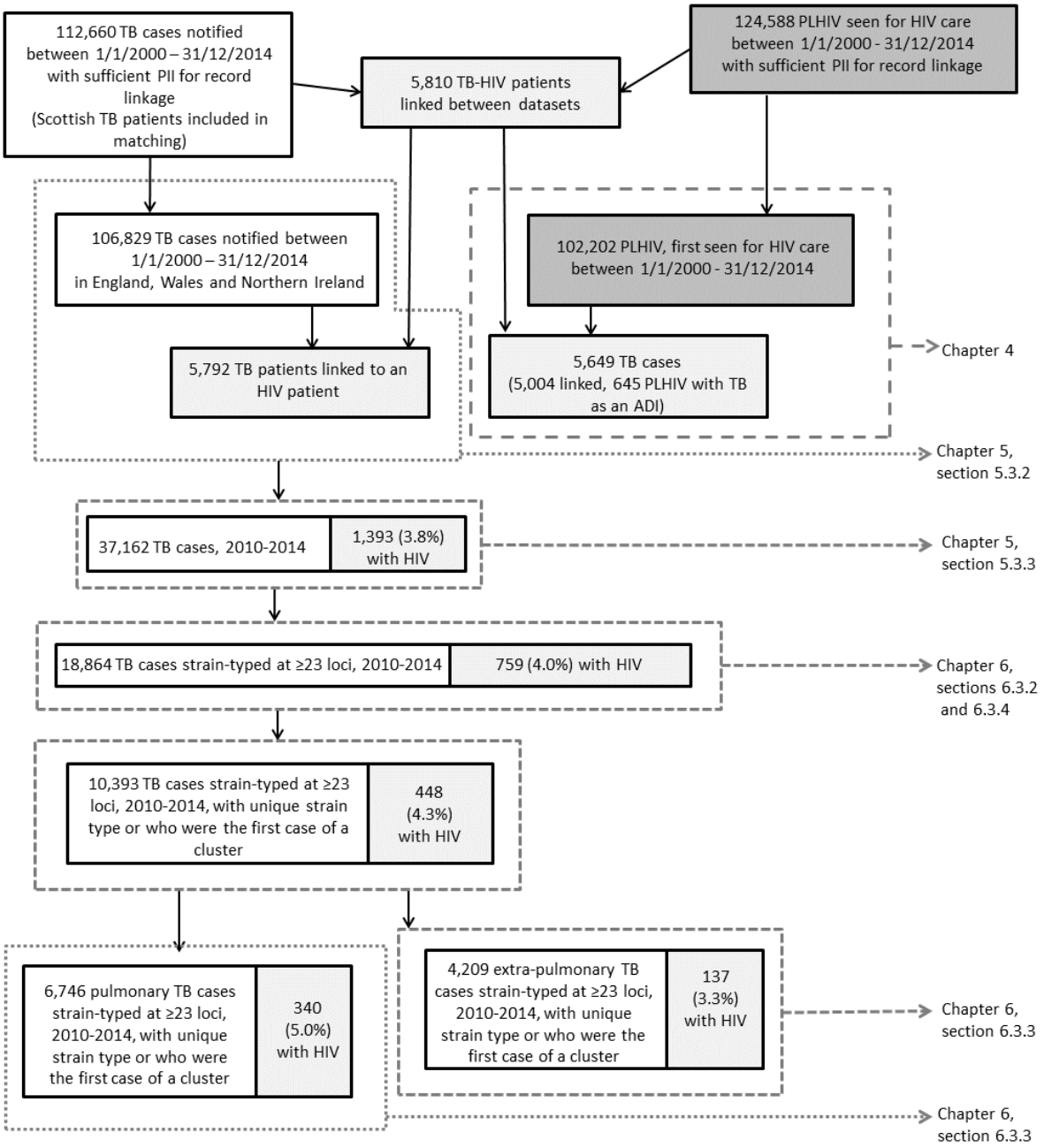
Match type	n	%	Cumulative %
1	3,304	61.3	61.3
1a	113	2.1	63.4
1b	60	1.1	64.6
1c	6	0.1	64.7
2	1,173	21.8	86.5
3	311	5.8	92.2
3a	25	0.5	92.7
4	33	0.6	93.3
4a	13	0.2	93.5
6	93	1.7	95.3
6a	60	1.1	96.4
6b	54	1.0	97.4
6c	1	0.0	97.4
7	26	0.5	97.9
7a	3	0.1	97.9
8	108	2.0	99.9
8b	3	0.1	100.0
Total	5,386	100.0	

n: number.

Additionally, of the 3,079 HIV patients reported to HARS with TB as an AIDS-defining illness, 695 (22.6%) were not linked to a record in ETS. These were included in the analysis presented in Chapter 4, but excluded from the analyses in Chapters 5 and 6 as there was no data on the characteristics of their TB disease (Figure 3.6).

The data used in each chapter of this thesis are shown in Figure 3.6.

Figure 3.6: Flow chart of the datasets linked in this chapter, and the records used in each thesis chapter.



The numbers of cases in this diagram only includes adults.

3.4 Discussion

3.4.1 Summary

The aim of this chapter was to investigate and develop an efficient and repeatable alternative to using the probabilistic algorithm of EMS for linking the TB and HIV surveillance datasets collated by PHE. A deterministic algorithm was developed and compared to existing probabilistic methods. The large number of additional matches found using the deterministic algorithm, some of which had low probabilistic scores and disagreed on other fields not included in the criteria used, indicated that deterministic matching alone was not discriminative enough to be used in place of EMS. A hybrid method of record linkage was the optimal solution, maintaining the ability to weight matches (using probabilistic methods) whilst using a set of consistent rules (the hierarchy of deterministic criteria) to accept or reject intermediate matches.

3.4.2 Strengths, limitations and implications of the work

The hybrid method provides a number of advantages over probabilistic or deterministic methods alone. The main advantage is the elimination of the manual review which is a necessary part of probabilistic matching using EMS; resulting in a much quicker, more consistent and repeatable method, with less potential to introduce systematic bias. Whilst the development of the algorithm involved manual review to confirm that the final selection of matching criteria were appropriate, this manual review will not be necessary as part of the routine use of the hybrid method. This is advantageous because where co-infection with TB and HIV is used as the outcome of epidemiological analyses, systematic misclassification of the outcome has the potential to significantly bias the results of any analyses due to over- or under-estimation of the number of co-infected cases.

One limitation of this work was that there was no gold-standard identifier to validate the method against. EMS was previously validated using a subset of the ETS data for which NHS number was available,[184] however the HIV dataset does not

contain any unique identifiers for patients. However, this limitation would exist regardless of the method used for linking TB and HIV records, and the deterministic algorithm used for internal matching within the HIV surveillance system has not been validated against a gold-standard for the same reason. The probabilistic algorithm EMS was validated using two TB surveillance datasets and several variables of PII which are not available when matching between the TB and HIV datasets. The sensitivity of EMS decreased from 99.5% to 97.1% when address variables were omitted, and it is reasonable to assume that the sensitivity would decrease further when other variables such as first name, surname and hospital names also cannot be used, as is the case for TB-HIV matching.

Another limitation was that there were 695 individuals with HIV reported to have TB as an AIDS-defining illness, who were not linked to a record in the TB surveillance dataset. This may be because of combinations of missing data, names having been changed but this not being reported, names which do not convert well to Soundex codes (particularly very short or very long names), or common combinations of name, ethnicity and date of birth; all of which could result in a low probabilistic matching score and therefore the match being missed. It is also possible that some of these were misdiagnosed cases of TB, which were reported to ETS but then later de-notified, or that this is an indication that there is under-reporting of TB cases. The hybrid matching method resulted in a higher proportion of the HIV patients with TB as an AIDS-defining illness being linked to a TB case in ETS (70.2%) than either the probabilistic or deterministic approaches alone (65.4% and 68.2% respectively), and this proportion was even higher (77.4%) when the datasets were matched for the period 2000-2014. However, it was not clear whether the remaining 22.6% remained unmatched as a result of the limitations of the matching algorithms or the limitations of the surveillance data collection. Future work to improve surveillance data collection or linkage could include an audit of these patients' clinical records to determine whether their TB was notified to ETS, and if so, why their records were not linked.

When developing the deterministic criteria and the hybrid method, the results were evaluated by making comparisons with the results from matching done using EMS. However, it is important to consider that the previous results are not a gold-standard; some matches may have been incorrectly matched and others missed. For this reason, the sensitivity and specificity of individual deterministic criteria could not be calculated. Future work could include recruiting a subset of patients, gaining patient consent to collect PII (including NHS number) and then using this data to validate the algorithm. Alternatively, the use of latent class models (which are often developed for evaluating new diagnostic tests) could be investigated.[194]

The hierarchical nature of the hybrid matching method means that sensitivity analyses can be done during epidemiological analyses to evaluate the impact of including 'less certain' matches in the analysis. This will not only inform on the reliability of the epidemiological results, but could also be used to further develop and refine the matching method by identifying specific groups which may be more or less likely to be matched on a low probabilistic score or a less stringent deterministic criterion. The results of the sensitivity analyses are described in sections 4.3.5 and 5.3.2.2.

3.4.3 Conclusions

Data linkage is an important epidemiological tool, and needs to be accurate so that the resulting datasets are free from bias. The hybrid matching method developed in this chapter is a considerable improvement on the use of EMS in the case of TB-HIV data linkage. The elimination of the time-consuming manual review means that the overall process is considerably quicker, more repeatable and less subject to bias, whilst the use of deterministic criteria in addition to probabilistic matching facilitated the finding of a number of very strong matches which EMS had failed to identify.

4 Sociodemographic and clinical risk factors for developing tuberculosis disease among people living with diagnosed HIV

4.1 Introduction

4.1.1 HIV and tuberculosis co-infection in the UK

In 2015, an estimated 101,200 people were living with HIV in the UK,[10] a prevalence of 0.16%.[195] 6,095 PLHIV were newly diagnosed in 2015 and 39% of these were diagnosed late, with a CD4 count of less than 350 cells/µl. PLHIV are known to be at a higher risk of developing TB than people without HIV.

In 2015, there were just over 6,000 TB cases notified in the UK.[52, 79-81] As HIV infection is a chronic, lifelong condition, which is known to increase the risk of developing TB; a substantial proportion of these would be expected to occur in PLHIV. The epidemiology of both TB and HIV in the UK was described previously in Chapter 1.

Previous estimates of the incidence of TB among PLHIV in the UK have varied as a result of differing methodologies. Studies using national surveillance data reported a TB incidence of 1,900 per 100,000 person-years (PY) among people who acquired their HIV infection through heterosexual sex from 2002-2010,[164] and 669/100,000 PY for all PLHIV from 2007-2011.[151] However, neither of these studies specified a minimum follow-up period, and began the follow-up period on the date of HIV diagnosis. They therefore included patients who were diagnosed simultaneously with their TB and HIV infections, leading to inflated estimates of incidence as a result of including a substantial number of TB cases without accounting for the time they spent at risk of disease. A study of TB incidence among PLHIV in an observational cohort reported a much lower incidence of 328/100,000 PY from 1996-2005,[119] having excluded TB cases where patients were diagnosed with TB and HIV within 91 days of each other. The difference between these estimates highlights the importance of

separating prevalent TB cases (present at the time of HIV diagnosis) from incident TB cases (where the disease actually occurs during the follow-up period) when calculating incidence rates. Additionally, evaluating risk factors separately for prevalent and incident cases is important, because different interventions are required to prevent them.

4.1.2 Known risk factors for co-infection

A series of risk factors for active TB disease for PLHIV in countries with low TB incidence have been documented and were described in detail in Chapter 2. A summary of the evidence is presented below.

Two large, high-quality prospective cohort studies in France and Germany reported that TB incidence decreased with time since HIV diagnosis.[118, 147] Similarly, a UK study of people who acquired HIV heterosexually demonstrated a high proportion (54%) of TB diagnoses were made simultaneously with HIV diagnosis;[164] and most TB diagnosis among people with diagnosed HIV were in the first two years after diagnosis, although this study did not adjust for the duration of follow-up.

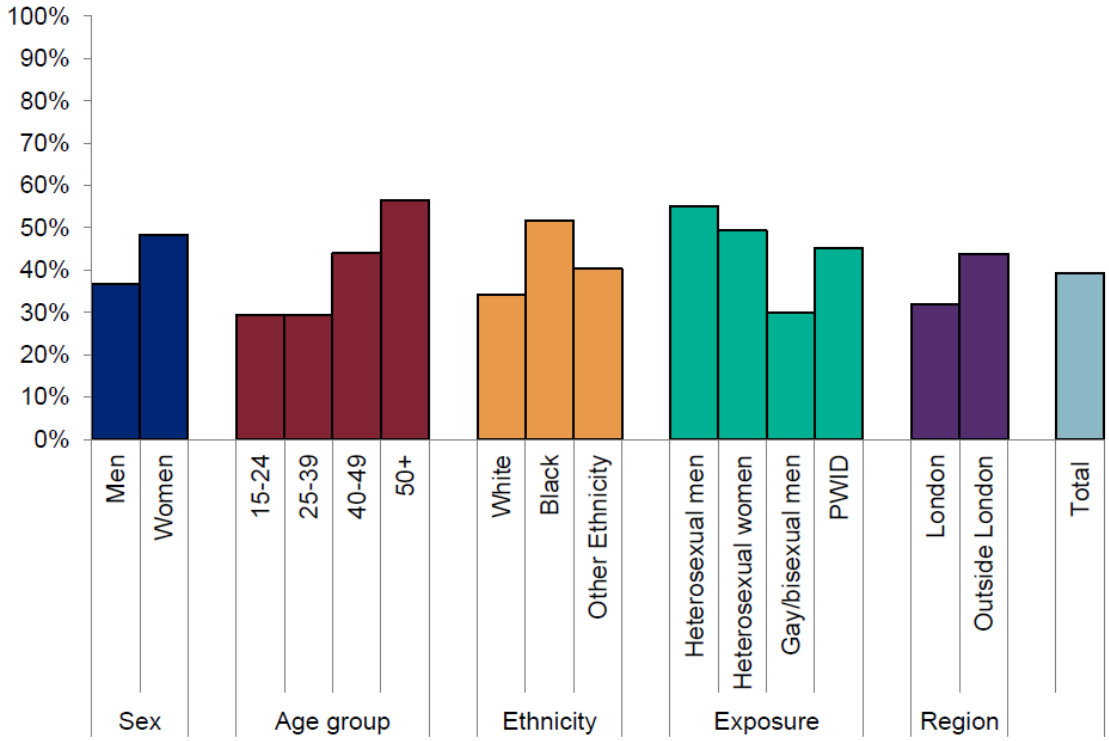
Injecting drug use was associated with TB disease for PLHIV in large observational cohort studies in Switzerland, Canada and the USA.[152, 161, 167] Several other large cohort studies in the UK, USA, Canada, Denmark and Israel reported that PLHIV who acquired their HIV infection through heterosexual sex and/or people who inject drugs had substantially higher incidence rates of TB than MSM.[119, 146, 151, 153, 156, 159, 169] However, of the two studies in the UK, one was restricted to PLHIV on ART [151] and the other did not examine injecting drug use as a risk factor for TB.[119]

Some studies suggested a lower risk of TB for men than women in the USA and Canada,[146, 153, 170] but studies in the UK and Israel which adjusted for route of HIV infection (and the substantial number of MSM in most populations of PLHIV who are at low risk of TB) reported a higher risk of TB for men than women.[119, 159, 164] Being

of black African or south Asian ethnicity was also associated with higher TB incidence than white ethnicity in three UK studies,[119, 151, 164] although none adjusted for TB incidence in country of birth, which is related to black African ethnicity. Studies conducted in other low-incidence countries found that TB incidence was consistently higher among foreign-born patients than those born in the countries in which the study was conducted,[153, 161, 162] particularly among patients born in sub-Saharan Africa or other countries with a high incidence of TB.[118, 144, 147, 169]

In the UK, the proportion of PLHIV that are diagnosed late varies for different sub-populations (Figure 4.1). Multiple cohort studies reported that late diagnosis of HIV infection and low CD4 count were important risk factors for TB disease; particularly for CD4 counts below 500 cells/µl,[118, 119, 144, 147, 148, 150, 152-154, 156, 164, 167-169] though risks have not been shown to be substantially different between the strata of 500-699 cells/µl and >700 cells/µl.[151]

Figure 4.1: Proportion of adults diagnosed late (with a CD4 cell count of less than 350 cells/µl) by demographic in the UK in 2015.



Source: Public Health England [10]

The evidence for viral load as an independent risk factor for TB is unclear. Several studies in the UK, France and Germany have reported associations between higher viral loads and TB.[118, 119, 147, 151, 154] Other studies in the USA, Canada and Denmark reported no association between viral load and risk of TB after adjusting for CD4 count and anti-retroviral therapy;[145, 153, 167, 169] however, adjusting for all three of these variables (which are causally related and highly collinear) can make it difficult to determine which variable is causing the effect.

Similarly, initiating ART [118, 144, 153, 164, 169] and increased duration of time on ART [119, 161, 162] were generally associated with lower rates of TB, although some studies found no significant association after adjusting for CD4 count and/or viral load.[145, 148, 154, 155, 167]

The risk of TB was significantly higher for patients with known LTBI,[144, 148, 150, 156, 165, 166] although this was decreased for patients who had taken preventive therapy.[144, 146, 165, 166] Age was not generally associated with risk of TB for PLHIV.

4.1.3 Summary and objectives

To conclude, there is a substantial body of literature on risk factors for developing TB for PLHIV in countries with low-incidence settings. However, previous studies in the UK have left a number of gaps in the literature through not considering HIV acquisition through injecting drug use as a risk factor for TB,[119] or restricting the studies to heterosexuals or those who have initiated ART.[164, 196] Additionally, none of these studies adjusted for the incidence of TB in the country of birth of patients born abroad, which may confound the association between ethnicity and risk of TB for PLHIV.

The aims of this chapter were to use HIV and TB national surveillance datasets to describe the incidence and epidemiology of TB-HIV co-infections in England, Wales and Northern Ireland over a 15-year period (2000 to 2014), to examine the relationship

between the timing of TB and HIV diagnoses for different sub-populations, and quantify clinical and demographic risk factors for TB across the whole population of PLHIV.

4.2 Methods

4.2.1 Study population

The study population was adults aged 15 years or older notified to PHE's HARS (described in section 3.1.2.2), who first presented with HIV to health services in England, Wales or Northern Ireland from 1st January 2000 to 31st December 2014.

4.2.2 Outcome: tuberculosis disease diagnosed from 2000-2014

HIV surveillance data collected in HARS were linked to TB surveillance data collected in ETS, as described in Chapter 3. HIV patients were considered to have TB disease if they were linked to a culture-confirmed or presumptive (clinical and radiological signs, including a response to specific therapy) diagnosis of TB reported to ETS, or if they had TB reported to HARS as an AIDS-defining illness.

Incident TB was defined as TB disease notified to ETS, or reported to HARS as a new AIDS-defining illness, that was diagnosed >91 days after HIV diagnosis. To differentiate patients who were not aware of their HIV infection prior to their TB diagnosis, TB cases diagnosed within 91 days of HIV diagnosis were considered simultaneous diagnoses. TB cases diagnosed >91 days before HIV were considered existing disease. A 91 day threshold for defining simultaneous diagnoses was a pragmatic choice to account for delays in diagnosis and reporting, and to exclude ART-induced unmasking immune reconstitution inflammatory syndrome.

4.2.3 Exposure variables

A series of demographic and clinical exposure variables, which were routinely collected in HARS or derived from other variables, were deemed of interest. Demographic factors included age at HIV diagnosis, sex, ethnicity, country of birth, TB incidence in country of birth, route of HIV infection, year of HIV diagnosis, and index of

multiple deprivation [IMD] decile. IMD scores represent relative levels of deprivation of income, employment, health, education, housing and services, crime and living environment for small areas in England and Wales, and are ranked into deciles where one is the most deprived and ten is the least deprived.[197, 198] The most recent IMD data for each country between 2000 and 2014 were used; 2010 for England and 2014 for Wales.

Composite variables were created combining ethnicity and country of birth or sex and infection route due to mutually exclusive combinations (e.g. being a woman and a MSM is impossible) and known associations. As a proxy for TB exposure, countries of birth outside the UK were grouped by TB incidence; 'high incidence' was defined as >40 cases/100,000 adult population in 2013.

Clinical risk factors included viral load, CD4 count and whether the patient was receiving ART. Longitudinal data was available for all three of these risk factors. Data were available for every CD4 count taken for each patient. Data on ART included the date the patient first initiated ART, and annual data on whether they were on ART at their last clinic visit. Data on viral load were available at diagnosis, and for the last clinic visit each year.

4.2.4 Statistical Analysis

4.2.4.1 Descriptive analysis

Data were analysed in Stata version 13.1. The proportion of the population of PLHIV diagnosed with TB and the relationship between the timing of HIV and TB diagnoses (simultaneous, TB first, or HIV first), were described both for the whole population and stratified by each exposure variable. Nelson-Aalen cumulative hazard plots were used to assess the cumulative risk of TB disease over time, whilst accounting for censoring of data due to shorter durations of follow-up.

4.2.4.2 Poisson regression modelling of risk factors for developing tuberculosis after HIV diagnosis

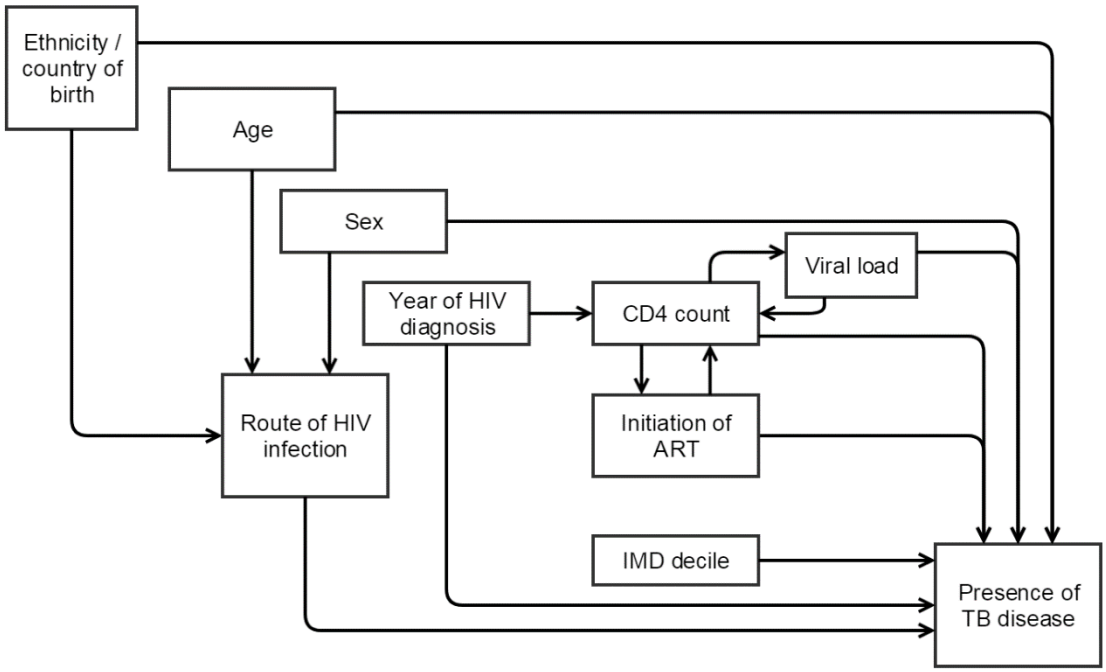
To investigate risk factors for developing TB following HIV diagnosis, incidence rates of TB per 100,000 PY follow-up were calculated. Individuals diagnosed with TB ≤91 days after HIV diagnosis were excluded. Cox regression was precluded as the data did not satisfy the proportional hazards assumption for key variables such as route of HIV infection, and therefore incidence rate ratios were estimated using univariable and multivariable Poisson regression models, offset by follow-up time. Follow-up began 92 days from date of HIV diagnosis or date of first presentation to UK health services, and ended on the date of TB diagnosis, death, or 31/12/2014, whichever was earliest.

CD4 count and ART initiation were included in the Poisson regression model as time-updated covariates. Incidence rates for different CD4 strata were calculated using the number of days from each CD4 count to the date of the next CD4 count for each patient. As data on ART is only reported annually from the most recent clinic visit for each patient, it was not possible to accurately determine when patients discontinued or reinitiated ART. Therefore, TB incidence was compared between ART-naïve patients and patients who had initiated ART, rather than patients who were on ART and patients who were not on ART. Patients were considered ART-naïve from the date of their HIV diagnosis until the date they first initiated ART, and to have initiated ART from that date until the end of their follow-up period. In a post-hoc analysis of patients who had initiated ART, the proportion of patients who were not on ART at their last clinic visit during the study period was examined for patients who developed TB and patients who remained TB-free. The proportion of patients who initiated ART and the time from the most recent clinic visit to the end of the study period were compared for different stratum of risk factors.

Potential confounders and effect modifiers were prospectively identified [199] and a causal framework was built (Figure 4.2). As there was not a single ‘main’ exposure variable, there were no confounders in the traditional sense, and therefore

the multivariable model was informed by the causal inference framework defined *a priori*. From this framework, it was determined that viral load should be excluded from the multivariable model because of the potential for causal loops between viral load and CD4 count, which could not be adequately accounted for in the data available as there were data on all performed CD4 counts but only one viral load measurement per year for each patient. Linearity (of age, CD4 count and year of HIV diagnosis) and statistical interactions (between ART status and CD4 count) were assessed using likelihood-ratio tests. Statistical interactions were considered significant at $P<0.05$; this was considered appropriate due to the large size of the dataset. All stated confidence intervals were two-sided 95% confidence intervals. Patients missing data on one or more variables were excluded. To assess the likely impact of missing data, the distributions of age, sex, route of HIV infection, CD4 count and ethnicity/country of birth were compared for cases with missing vs. complete data on infection route, CD4 count, IMD score and country of birth.

Figure 4.2: Causal diagram of potential risk factors for developing TB disease among people diagnosed with HIV.



ART: anti-retroviral therapy, IMD: index of multiple deprivation, TB: tuberculosis. Arrows show the direction of effect

4.2.4.3 Sensitivity analyses

Sensitivity analyses were planned as follows. (1) Investigating the impact of using a 6-month threshold (182 days) for simultaneous diagnosis, rather than a 91-day threshold. (2) Examining the impact of excluding weaker matches between HARS and ETS, excluding co-infected patients in the lowest quartile of probabilistic matching score, and excluding patients who were matched using the three weakest deterministic matching criteria. (3) Excluding patients who acquired HIV infection through mother-to-child transmission, as the dataset only contained adults and so individuals infected through this route could be missing 15 years of follow-up.

4.3 Results

4.3.1 Description of co-infected patients

Between 2000 and 2014, 102,202 adults were newly diagnosed with HIV, among whom 5,649 (6%) had TB. 3,103 (55%) were simultaneously diagnosed with TB and HIV, 2,187 (39%) developed TB after more than 91 days and 359 (6%) were diagnosed with TB first (Table 4.1).

Of people with TB who acquired HIV infection through heterosexual sex, over half were diagnosed simultaneously with TB and HIV; 60% for men and 54% for women. In contrast, more TB cases in MSM and PWID were diagnosed more than 91 days after diagnosis of HIV infection (51% and 54%, respectively). The proportion of TB cases occurring after HIV diagnosis was highest in white, UK-born individuals (179/359, 48%) and those born in low TB incidence countries (116/245, 47%); these two groups comprise 38% of the cohort.

There was no substantial difference in the age, sex, ethnicity/country of birth, route of HIV infection or CD4 count of patients with missing data on any of the following variables: route of HIV infection, CD4 count, IMD decile and country of birth. Patients with missing route of infection were less likely to be diagnosed with TB; however, there were no substantial differences for patients missing data on any other variable.

Table 4.1: TB diagnoses in people notified with HIV from 2000 to 2014 in England, Wales and Northern Ireland, and the incidence rates of TB in people who were diagnosed with TB more than 91 days following HIV diagnosis.

	HIV cases	TB cases						
		Total	Prior to HIV diagnosis	Simultaneous with HIV diagnosis	Following HIV diagnosis			
	n (column %)	n (row %)	n (row %)	n (row %)	n (row %)	PY follow-up	Incidence rate* (95% CI)	Incidence rate after 365 days from HIV diagnosis* (95% CI)
Total	102,202	5,649 (5.5)	359 (6)	3,103 (55)	2,187 (39)	635,591	344 (330 - 359)	247 (234 - 260)
Route of HIV infection								
MSM	35,879 (35.1)	462 (1.3)	31 (7)	195 (42)	236 (51)	212,844	111 (98 - 126)	86 (74 - 100)
Heterosexual men	18,738 (18.3)	2,013 (10.7)	127 (6)	1,205 (60)	681 (34)	113,802	598 (555 - 645)	402 (365 - 443)
Heterosexual women	30,489 (29.8)	2,815 (9.2)	167 (6)	1,520 (54)	1,128 (40)	201,644	559 (528 - 593)	404 (376 - 434)
Men who inject drugs	1,453 (1.4)	132 (9.1)	5 (4)	55 (42)	72 (55)	8,216	876 (696 - 1,104)	660 (499 - 873)
Women who inject drugs	532 (0.5)	35 (6.6)	1 (3)	15 (43)	19 (54)	3,138	605 (365 - 945)	526 (295 - 868)
Blood/Tissue transfer	505 (0.5)	58 (11.5)	6 (10)	31 (53)	21 (36)	2,928	717 (468 - 1,100)	527 (288 - 883)
Mother-to-child	253 (0.2)	15 (5.9)	1 (7)	4 (27)	10 (67)	863	1,159 (556 - 2,131)	836 (307 - 1,819)
Unknown≠	14,353 (14.0)	119 (0.8)	21 (18)	78 (66)	20 (17)	92,155	22 (14 - 34)	13 (7 - 24)
Ethnicity/Country of birth								
White, UK-born	27,320 (26.7)	359 (1.3)	24 (7)	161 (45)	174 (48)	160,488	108 (93 - 126)	84 (70 - 100)
Black African, UK-born	947 (0.9)	51 (5.4)	6 (12)	25 (49)	20 (39)	5,556	360 (232 - 558)	260 (151 - 448)
Other ethnicity, UK-born	2,687 (2.6)	72 (2.7)	6 (8)	26 (36)	40 (56)	14,948	268 (196 - 365)	217 (151 - 313)
Ethnicity unknown, UK-born	403 (0.4)	3 (0.7)	0 (0)	3 (100)	0 (0)	544	0 (0 - 678)‡	0 (0 - 876)‡
Born in low-TB incidence country	11,551 (11.3)	245 (2.1)	11 (4)	118 (48)	116 (47)	65,376	177 (148 - 213)	125 (99 - 157)

	HIV cases	TB cases						
		Total	Prior to HIV diagnosis	Simultaneous with HIV diagnosis	Following HIV diagnosis			
	n (column %)	n (row %)	n (row %)	n (row %)	n (row %)	PY follow-up	Incidence rate* (95% CI)	Incidence rate after 365 days from HIV diagnosis* (95% CI)
White, born in high-TB incidence country	7,461 (7.3)	126 (1.7)	4 (3)	71 (56)	51 (40)	47,593	107 (81 - 141)	84 (61 - 116)
Black African, born in high-TB incidence country	35,035 (34.3)	3,877 (11.1)	223 (6)	2,142 (55)	1,512 (39)	234,853	644 (612 - 677)	454 (426 - 483)
Other ethnicity, born in high-TB incidence country	6,756 (6.6)	518 (7.7)	52 (10)	311 (60)	155 (30)	35,614	435 (372 - 509)	290 (236 - 356)
Ethnicity unknown, born in high-TB incidence country	1,140 (1.1)	13 (1.1)	2 (15)	10 (77)	1 (8)	7,556	13 (0 - 74)‡	15 (0 - 81)
White, country of birth unknown	3,065 (3.0)	52 (1.7)	4 (8)	31 (60)	17 (33)	23,968	71 (41 - 114)	54 (28 - 95)
Other ethnicity, country of birth unknown	4,226 (4.1)	300 (7.1)	23 (8)	181 (60)	96 (32)	33,093	290 (237 - 354)	210 (164 - 268)
Both Unknown≠	1,611 (1.6)	33 (2.0)	4 (12)	24 (73)	5 (15)	6,002	83 (27 - 194)	39 (5 - 141)
Age at HIV diagnosis (years)								
15-24	11,513 (11.3)	437 (3.8)	25 (6)	173 (40)	239 (55)	73,647	325 (286 - 368)	260 (224 - 302)
25-34	38,910 (38.1)	2,227 (5.7)	129 (6)	1,121 (50)	977 (44)	261,955	373 (350 - 397)	280 (260 - 302)
35-44	31,894 (31.2)	1,944 (6.1)	133 (7)	1,147 (59)	664 (34)	199,946	332 (308 - 358)	232 (211 - 255)
45-64	18,357 (18.0)	973 (5.3)	64 (7)	619 (64)	290 (30)	93,708	309 (276 - 347)	183 (156 - 214)
65+	1,479 (1.4)	68 (4.6)	8 (12)	43 (63)	17 (25)	5,764	295 (172 - 472)	99 (172 - 472)
CD4 count at HIV diagnosis† (incidence rates are calculated for time-updated CD4)								
≥500	20,153 (19.7)	381 (1.9)	30 (8)	88 (23)	263 (69)	187,994	139 (123 - 157)	122 (106 - 139)
350-499	14,801 (14.5)	455 (3.1)	34 (7)	133 (29)	288 (63)	114,505	259 (231 - 290)	270 (241 - 304)
200-349	16,282 (15.9)	861 (5.3)	61 (7)	388 (45)	412 (48)	81,579	527 (480 - 579)	454 (407 - 506)

	HIV cases	TB cases						
		Total	Prior to HIV diagnosis	Simultaneous with HIV diagnosis	Following HIV diagnosis			
	n (column %)	n (row %)	n (row %)	n (row %)	n (row %)	PY follow-up	Incidence rate* (95% CI)	Incidence rate after 365 days from HIV diagnosis* (95% CI)
100-199	9,514 (9.3)	1,039 (10.9)	79 (8)	613 (59)	347 (33)	24,933	1,356 (1,219 - 1,508)	785 (673 - 916)
50-99	5,039 (4.9)	718 (14.2)	35 (5)	525 (73)	158 (22)	6,247	2,209 (1,870 - 2,610)	1,072 (817 - 1,407)
0-49	8,731 (8.5)	1,241 (14.2)	63 (5)	956 (77)	222 (18)	5,166	2,788 (2,368 - 3,282)	891 (648 - 1,224)
Unknown≠	27,682 (27.1)	954 (3.4)	57 (6)	400 (42)	497 (52)	-	-	-
Viral load at diagnosis (copies/ml)								
≤200	13,951 (13.7)	580 (4.2)	51 (9)	311 (54)	218 (38)	63,098	345 (303 - 395)	227 (190 - 270)
>200	58,824 (57.6)	3,735 (6.3)	229 (6)	2,050 (55)	1,456 (39)	339,621	428 (407 - 451)	305 (286 - 325)
Unknown≠	29,427 (28.8)	1,334 (4.5)	79 (6)	742 (56)	513 (38)	232,872	221 (202 - 241)	170 (153 - 188)
Ever started ART (incidence rates calculated for time-updated ART)§								
No	32,207 (31.5)	809 (2.5)	-	-	1336§	261,662	511 (484 - 539)	337 (314 - 362)
Yes	69,995 (68.5)	4,840 (6.9)	-	-	851§	373,929	228 (213 - 243)	188 (174 - 203)

	HIV cases	TB cases						
		Total	Prior to HIV diagnosis	Simultaneous with HIV diagnosis	Following HIV diagnosis			
	n (column %)	n (row %)	n (row %)	n (row %)	n (row %)	PY follow-up	Incidence rate* (95% CI)	Incidence rate after 365 days from HIV diagnosis* (95% CI)
IMD decile								
1	13,498 (13.2)	900 (6.7)	64 (7)	470 (52)	366 (41)	75,516	485 (437 - 537)	343 (301 - 390)
2	15,075 (14.8)	920 (6.1)	66 (7)	510 (55)	344 (37)	86,339	398 (358 - 443)	286 (251 - 327)
3	12,746 (12.5)	688 (5.4)	53 (8)	385 (56)	250 (36)	72,760	344 (304 - 389)	247 (212 - 288)
4	9,150 (9.0)	474 (5.2)	29 (6)	273 (58)	172 (36)	52,758	326 (281 - 379)	222 (183 - 268)
5	6,732 (6.6)	336 (5.0)	22 (7)	191 (57)	123 (37)	37,961	324 (272 - 387)	235 (189 - 293)
6	5,233 (5.1)	253 (4.8)	18 (7)	134 (53)	101 (40)	29,630	341 (280 - 414)	238 (186 - 304)
7	3,870 (3.8)	164 (4.2)	10 (6)	89 (54)	65 (40)	21,596	301 (236 - 384)	233 (174 - 312)
8	3,304 (3.2)	140 (4.2)	6 (4)	83 (59)	51 (36)	17,934	290 (221 - 381)	207 (147 - 291)
9	2,809 (2.7)	110 (3.9)	7 (6)	64 (58)	39 (35)	15,846	246 (180 - 337)	163 (108 - 245)
10	2,217 (2.2)	97 (4.4)	3 (3)	52 (54)	42 (43)	11,925	352 (260 - 477)	274 (190 - 394)
Unknown≠	27,568 (27.0)	1,567 (5.7)	81 (5)	852 (54)	634 (40)	213,326	297 (274 - 321)	217 (197 - 238)

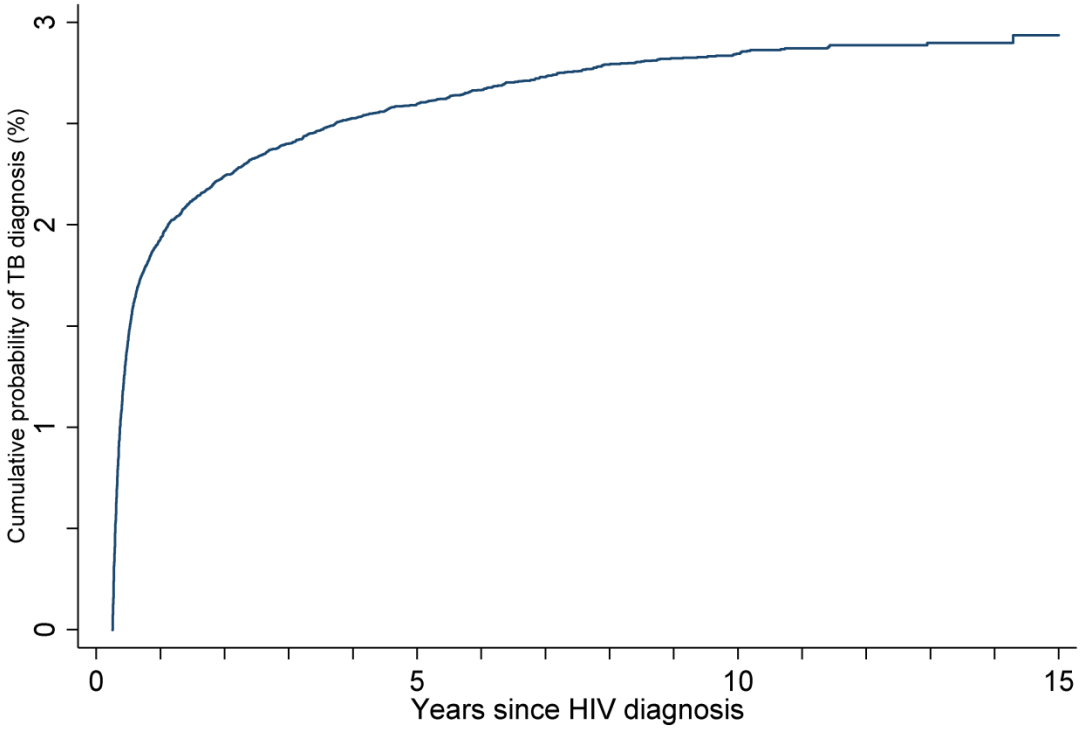
*** Incidence is given per 100,000 population aged ≥15 years, per year. † Incidence rates are calculated for time-updated CD4 count. ≠ Unknown strata includes both unknown and missing data. ‡One-sided, 97.5% CI. §Of the 5,649 PLHIV who got TB, 809 never initiated ART. However, of the 2,187 who got TB >91 days after their HIV infection, 1,336 had not initiated ART at the time of their TB diagnosis. ART: anti-retroviral therapy, CI: confidence interval, IMD: index of multiple deprivation (only available for patients in England and Wales), MSM: men who have sex with men, n: number, PWID: people who inject drugs, PY: person-years, TB: tuberculosis.**

4.3.2 Incidence of tuberculosis following HIV diagnosis

95,003 adults were TB-free 92 days after presenting for HIV care, with a total of 635,591 PY follow-up. Median age at HIV diagnosis was 34 years (inter-quartile range [IQR] 28-42) and median CD4 count was 340 cells/µl (IQR 170-527). 95% of patients who were linked to CD4 surveillance data had more than one CD4 count; the median number of counts was 14.

Overall TB incidence following HIV diagnosis was 344/100,000 PY (95% confidence interval [CI]: 330-359, Table 4.1). The risk of developing TB was highest in the year following HIV diagnosis and then decreased (Figure 4.3); after the first year the incidence of TB was 247/100,000 PY [234-260/100,000] (Table 4.1).

Figure 4.3: Cumulative hazard plot of the probability of developing TB from 91 days after HIV diagnosis for all people diagnosed with HIV in England, Wales and Northern Ireland from 2000 to 2014.



Incidence was high in PWID (men 876/100,000 [696-1,104/100,000]; women 605/100,000 [386-949/100,000]) and heterosexuals (men 598/100,000 [555-645/100,000], women 559 [528-593/100,000]), particularly compared with MSM

(111/100,000 [98-126/100,000]). Incidence among PLHIV who acquired their HIV infection through blood or tissue transfer or mother-to-child transmission was also high, although the number of person-years of follow-up for both of these groups was low (2,928 and 863 PY respectively, Table 4.1).

PLHIV born in countries with high TB incidence had substantially higher incidence rates of TB compared to people born in the UK or other low-incidence countries. TB incidence in people born in the UK was 108/100,000 [93-126/100,000] for white people, 360/100,000 [232-558/100,000] for people of black African ethnicity and 268/100,000 [196-365/100,000] for people of other ethnicities. In comparison, TB incidence in people born in countries with high TB incidence was 644/100,000 [612-677/100,000] for black Africans, 435/100,000 [372-509/100,000] for other ethnicities and 107/100,000 [81-141/100,000] for white people.

Stratifying by both route of HIV infection (Figure 4.4) and ethnicity (Figure 4.5), the largest differences in cumulative probability of TB diagnosis occurred in the first two years following HIV diagnosis; the rate of diagnosis remained relatively constant across all groups thereafter. The risk of TB was greatest for men who injected drugs and people who acquired HIV heterosexually, but also high among people who were infected with HIV through blood/tissue transfer or through mother-to-child transmission, although the numbers of patients in the latter categories were much lower. Men who injected drugs also had highest cumulative risk of TB over the study period (Figure 4.4).

Figure 4.4: Cumulative hazard plot of the probability of developing TB from 91 days after HIV diagnosis for people diagnosed with HIV in England, Wales and Northern Ireland from 2000 to 2014, stratified by route of HIV infection.

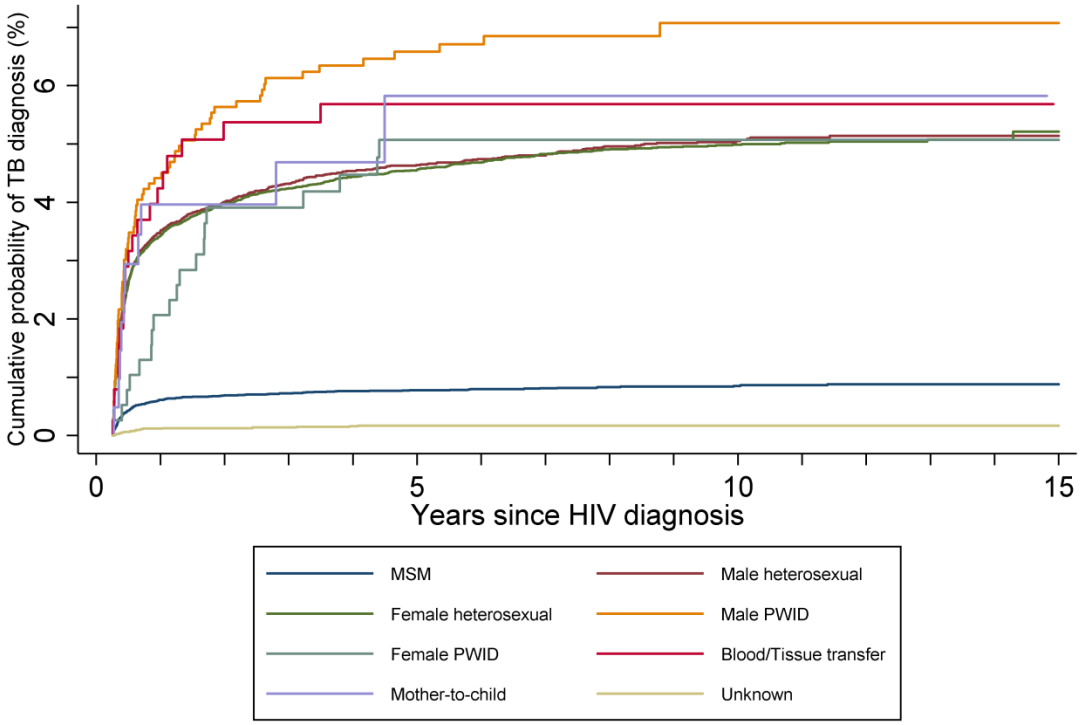
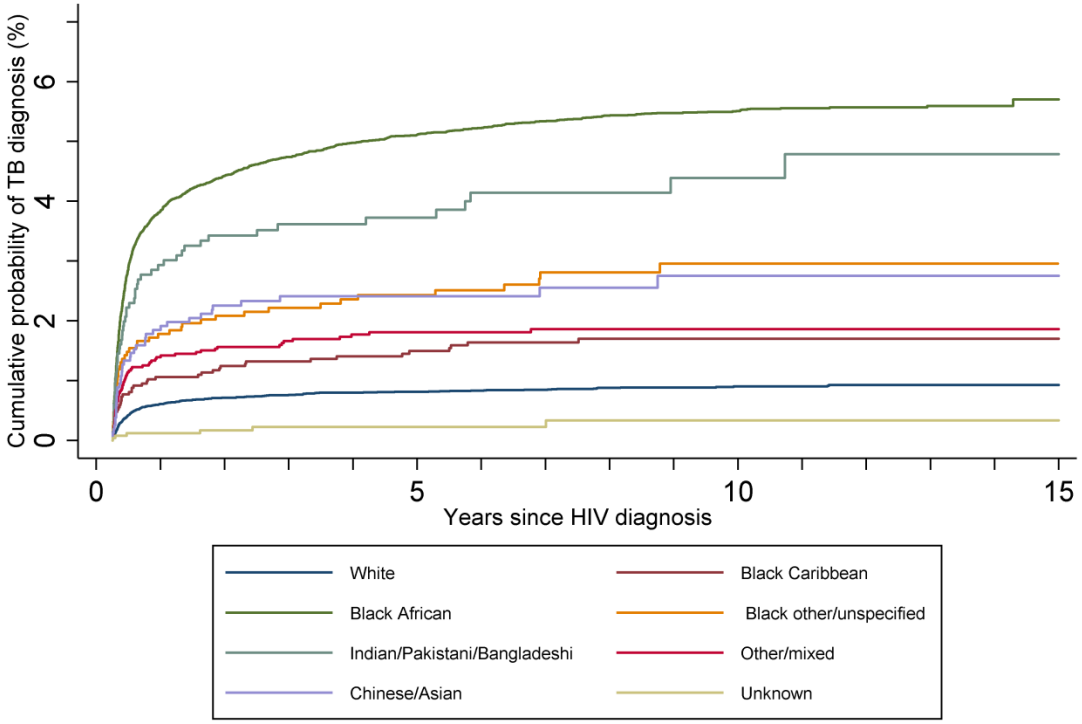


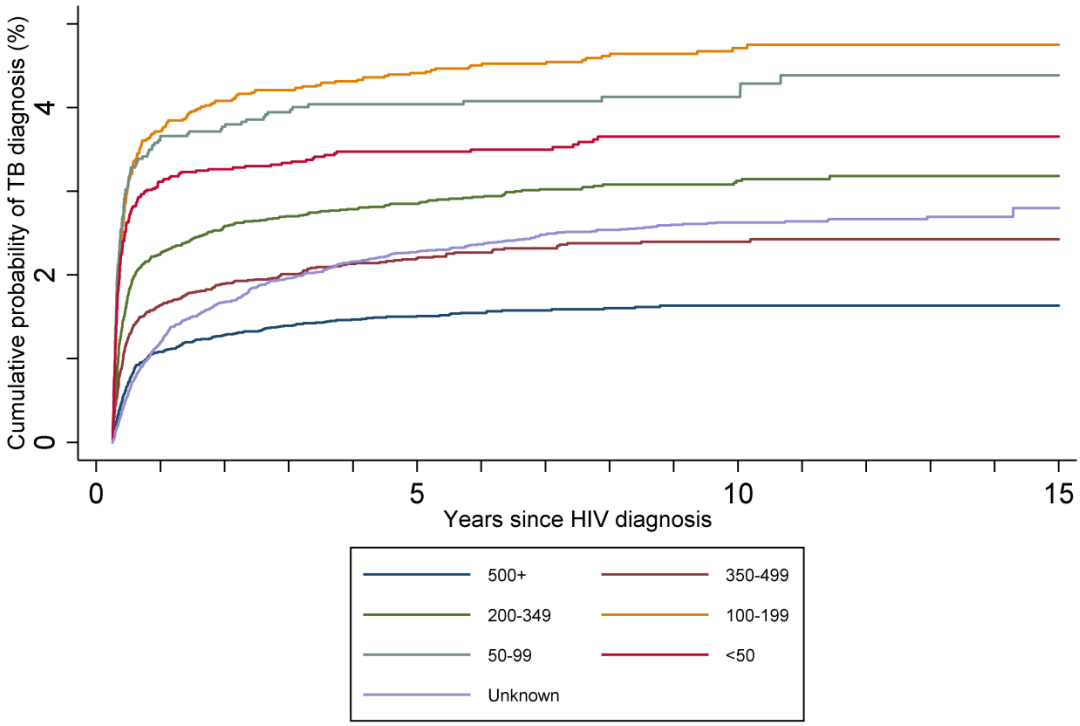
Figure 4.5: Cumulative hazard plot of the probability of developing TB from 91 days after HIV diagnosis for people diagnosed with HIV in England, Wales and Northern Ireland from 2000 to 2014, stratified by ethnicity.



TB incidence was highest among PLHIV who were aged 25-34 years at the time of their HIV diagnosis (373/100,000 [350-397/100,000], and decreased with age at HIV diagnosis to 295/100,000 [172-472/100,000] for people aged over 65 years when diagnosed with HIV.

TB incidence increased with decreasing time-updated CD4 count, from 139/100,000 (123-157/100,000) for those with CD4 count ≥500 cells/μl to 2,788/100,000 (2,368-3,282/100,000) for those with CD4 count <50 cells/μl. Irrespective of CD4 count at diagnosis, the risk of TB was greatest in the first year following diagnosis, but the cumulative risk of TB was inversely associated with CD4 count at diagnosis (Figure 4.6). The only exception to this was for patients who had a CD4 count of <50 cells/μl at HIV diagnosis, who had a lower cumulative risk of TB than those with CD4 counts between 50 and 199 cells/μl.

Figure 4.6: Cumulative hazard plot of the probability of developing TB from 91 days after HIV diagnosis for people diagnosed with HIV in England, Wales and Northern Ireland from 2000 to 2014, stratified by CD4 count at HIV diagnosis.



TB incidence was 511/100,000 (484-539/100,000) in people who had never received ART (26% of all PY) compared to 228/100,000 (213-243/100,000) in people who had (74% of PY). TB incidence was higher for PWID who had never initiated ART (1,478/100,000 [95% CI 1,157-1,888/100,000] than for black Africans from high-TB incidence countries who had never initiated ART (991/100,000 [929-1,058/100,000]) although incidence rates following ART initiation were similar in both groups (384/100,000 [264-560/100,000] for PWID versus 421 [389-456/100,000] for black Africans).

TB incidence was highest in those living in areas of England and Wales with the lowest decile of IMD score (485/100,000 [437-537/100,000]) and decreased with increasing IMD decile.

4.3.3 Factors associated with developing tuberculosis disease

62,684 individuals with complete case data and a TB-free follow-up period of >91 days following HIV diagnosis were included in the time-to-event analysis. There were a total of 414,714 PY of follow-up (median follow-up 7.1 years, IQR 3.6-10.4), during which there were 1,591 TB diagnoses (Table 4.2). The median duration of follow-up was 7.3 years (IQR 3.9-10.4) for patients who did not develop TB, whilst patients who did develop TB did so in a median of 0.2 years (IQR 0.1-0.5). Black African patients born in high-TB countries had a slightly longer median follow-up period of 8.2 (4.7-11.0) years, compared to 6.3 (3.1-9.8) for MSM and 6.5 (3.4-9.8) for PWID, as black Africans were more likely to be diagnosed earlier in the study period than PWID or MSM.

Table 4.2: Univariable and multivariable incidence rate ratios from Poisson regression of factors associated with incident TB disease (more than 91 days after HIV diagnosis) among PLHIV in England, Wales and Northern Ireland from 2000 to 2014.

	TB cases	PY	Univariable IRR (95% CI)	Multivariable IRR (95% CI)
Route of HIV infection				
MSM	184	172,708	1.00 (P<0.001)	1.00 (P<0.001)
Male heterosexual	474	82,460	5.40 (4.55 - 6.40)	1.70 (1.38 - 2.10)
Female heterosexual	837	148,391	5.29 (4.51 - 6.21)	1.86 (1.51 - 2.29)
Male PWID	61	5,895	9.71 (7.27 - 12.97)	5.47 (4.07 - 7.35)
Female PWID	16	2,514	5.97 (3.58 - 9.95)	4.59 (2.75 - 7.67)
Blood/Tissue transfer	14	2,251	5.84 (3.39 - 10.05)	2.70 (1.55 - 4.71)
Mother-to-child	5	494	9.51 (3.91 - 23.11)	2.80 (1.13 - 6.97)
Ethnicity/Country of birth				
White, UK-born	134	127,453	1.00 (P<0.001)	1.00 (P<0.001)
Black African, UK-born	13	4,317	2.86 (1.62 - 5.06)	1.97 (1.10 - 3.51)
Other ethnicity, UK-born	31	12,040	2.45 (1.66 - 3.62)	1.92 (1.29 - 2.84)
Ethnicity unknown, UK-born	0	252	†	†
Born in low-TB incidence country	98	53,647	1.74 (1.34 - 2.25)	1.33 (1.02 - 1.73)
White, born in high-TB incidence country	38	12,606	2.87 (2.00 - 4.11)	2.19 (1.53 - 3.15)
Black African, born in high-TB incidence country	1,093	148,017	7.02 (5.87 - 8.40)	4.27 (3.42 - 5.33)
Other ethnicity, born in high-TB incidence country	105	22,219	4.50 (3.48 - 5.80)	3.36 (2.57 - 4.39)
Ethnicity unknown, born in high-TB incidence country	1	323	2.95 (0.41 - 21.07)	1.35 (0.19 - 9.71)
White, country of birth unknown	12	15,491	0.74 (0.41 - 1.33)	0.52 (0.29 - 0.94)
Other ethnicity, country of birth unknown	66	18,348	3.42 (2.55 - 4.59)	1.60 (1.17 - 2.20)
CD4 count (time-updated, cells/µl)				
≥500	259	185,719	1.00 (P<0.001)	*
350-499	293	113,185	1.86 (1.57 - 2.19)	
200-349	427	80,443	3.81 (3.26 - 4.44)	
100-199	332	24,367	9.77 (8.30 - 11.49)	
50-99	137	6,093	16.12 (13.11 - 19.83)	
0-49	143	4,905	20.90 (17.04 - 25.64)	
Ever initiated ART (time-updated)				
No	928	107,477	1.00 (P<0.001)	*
Yes	663	307,237	0.25 (0.23 - 0.28)	
Viral load at HIV diagnosis (copies/ml)				
≤200	154	43,347	1.00 (P=0.006)	-
>200	1,063	261,249	1.15 (0.97 - 1.36)	

	TB cases	PY	Univariable IRR (95% CI)	Multivariable IRR (95% CI)
Age at HIV diagnosis				
15-24	169	48,805	0.95 (0.79 - 1.13)	0.92 (0.77 - 1.10)
25-34	714	170,957	1.14 (1.02 - 1.28)	1.06 (0.94 - 1.19)
35-44	477	130,441	1.00 (P<0.001)	1.00 (P=0.332)
45-64	220	61,028	0.99 (0.84 - 1.16)	1.11 (0.95 - 1.31)
≥65	11	3,484	0.86 (0.47 - 1.57)	0.92 (0.51 - 1.68)
Year of HIV diagnosis				
(for each year increase from 2000)	1,591	414,714	0.98 (0.97 - 1.00)	1.02 (1.00 - 1.04)
			P=0.036	P=0.014
IMD decile (England and Wales only)				
1	264	51,685	1.00 (P<0.001)	-
2	269	63,391	0.83 (0.70 - 0.98)	
3	193	54,955	0.69 (0.57 - 0.83)	
4	127	38,159	0.65 (0.53 - 0.81)	
5	83	26,725	0.61 (0.48 - 0.78)	
6	78	20,986	0.73 (0.57 - 0.94)	
7	47	15,254	0.60 (0.44 - 0.82)	
8	38	12,644	0.59 (0.42 - 0.83)	
9	24	10,743	0.44 (0.29 - 0.66)	
10	32	8,326	0.75 (0.52 - 1.09)	

62,684 PLHIV were included in this analysis; 32,319 were excluded from the model due to missing data on ethnicity and country of birth, route of HIV infection, CD4 count or age at HIV diagnosis. Viral load was not included in the multivariable model due to collinearity with CD4 count and ART status. *Interaction present between time-updated CD4 count and time-updated ART status, see Table 4.4 and Table 4.3. †Not calculated as numerator was zero. ART: anti-retroviral therapy, CI: confidence interval, IMD: index of multiple deprivation, MSM: men who have sex with men, PWID: people who inject drugs, PY: person years, IRR: incidence rate ratio, TB: tuberculosis.

4.3.3.1 Univariable results

In univariable analyses, the incidence rate ratio (IRR) for TB in PLHIV was significantly higher for all other routes of HIV infection than for MSM. The IRR was 9.71 [7.27-12.97] for men who inject drugs, 5.97 [3.58-9.95] for women who inject drugs, 5.40 [4.55-6.40] for heterosexual men, 5.29 [4.51-6.21] for heterosexual women and 9.51 [3.91-23.11] for people infected with HIV by mother-to-child transmission (Table 4.2).

Compared to white, UK-born people, TB incidence was higher for black Africans (IRR 2.86 [1.62-5.06]) and people of other ethnicities (2.45 [1.66-3.62]) born in the UK, and slightly higher for people of all ethnicities born in other low-incidence countries (1.74 [1.34-2.25]). However, the IRR was significantly higher for people born in countries with high TB incidence, with an IRR of 2.87 [2.00-4.11] for white people, 7.02 [5.87-8.40] for black Africans and 4.50 [3.48-5.80] for other ethnicities.

Age at HIV diagnosis was associated with incidence of TB; compared to people aged 35-44 years the IRR was elevated for people aged 25-34 (1.14 [1.02-1.128]) but not significantly different for other age groups.

TB incidence generally decreased as IMD decile increased; compared to the baseline of the lowest decile the IRR was 0.61 [0.48-0.78] for the 5th decile and 0.44 [0.29-0.66] for the 9th decile, although CIs overlapped.

Decreasing time-updated CD4 count was strongly and significantly associated with higher IRRs. Compared to a CD4 count >500 cells/µl, the IRR was 1.86 [1.57-2.19] for 350-499 cells/µl and this increased to 20.90 [17.04-25.64] for <50 cells/µl. Having ever initiated ART was significantly associated with substantially lower IRR of TB [0.25 [0.23-0.28].

Viral load at HIV diagnosis was also positively associated with risk of developing TB; compared to the baseline of ≤200 copies/ml the IRR for >200 copies/ml was 1.15 [0.97-1.36].

4.3.3.2 Multivariable results

All exposures were included in the multivariable Poisson regression model (Table 4.2), except viral load and IMD decile. IMD decile was excluded as there was a high degree of missing data and no association with the outcome in a multivariable model; it was included in a sensitivity analysis. CD4 count and age at HIV diagnosis were treated as categorical variables (tests for linearity P<0.001 and P=0.005, respectively); year of HIV diagnosis was included as linear variable (P=0.412). There was a statistically significant interaction between time-updated CD4 count and time-updated ART status (P<0.001).

Compared to MSM, PWID had increased rates of TB (incidence rate ratio [IRR] for men 5.47 [95% CI 4.07-7.35]; women 4.59 [2.75-7.67]). Rates were also higher in those infected through heterosexual sex (men 1.70 [1.38-2.10]; women 1.86 [1.51-2.29]). UK-born black Africans (1.97 [1.10-3.51]) and people of other ethnicities (1.92 [1.29-2.84]) were associated with increased incidence rates versus white UK-born individuals, as were those born in high TB incidence countries (black African 4.27 [3.42-5.33], white 2.19 [1.53-3.15], other ethnicities 3.36 [2.57-4.39]) and people born in low-incidence countries other than the UK (1.33 [1.02-1.73]).

Overall, and within each stratum of CD4 count, TB rates were greatly reduced in individuals who had received ART compared to those who had not (Table 4.3). Compared to never having initiated ART, the IRR for TB in people who had initiated ART was 0.07 [0.05-0.10] for people with a CD4 count ≥500 cells/µl, and 0.49 [0.35-0.69] for people with a CD4 count of <50 cells/µl.

Table 4.3: Multivariable Poisson regression of the association between time-updated ART status and TB disease, stratified by CD4 count, among PLHIV in England, Wales and Northern Ireland from 2000 to 2014.

	CD4 count (cells/µl)					
	≥500	350-499	200-349	100-199	50-99	0-49
Ever on ART	IRR (95% CI)	IRR (95% CI)	IRR (95% CI)	IRR (95% CI)	IRR (95% CI)	IRR (95% CI)
No	1.00	1.00	1.00	1.00	1.00	1.00
Yes	0.07 (0.05 - 0.10)	0.14 (0.11 - 0.18)	0.21 (0.17 - 0.25)	0.32 (0.26 - 0.40)	0.35 (0.25 - 0.49)	0.49 (0.35 - 0.69)

Incidence rate ratios derived from multivariable Poisson regression of the association between time-updated ART status and TB disease, stratified by CD4 count. Model adjusted for the variables in the multivariable model in Table 4.2. 62,684 PLHIV were included in this analysis; 32,319 were excluded from the model due to missing data on ethnicity and country of birth, route of HIV infection, CD4 count or age at HIV diagnosis. ART: anti-retroviral therapy, CI: confidence interval, IRR: incidence rate ratio, TB: tuberculosis.

Table 4.4: Multivariable Poisson regression of the association between time-updated CD4 count and TB disease, stratified by ART status, among PLHIV in England, Wales and Northern Ireland from 2000 to 2014.

	Ever on ART	
CD4 count	No	Yes
(cells/µl)	IRR (95% CI)	IRR (95% CI)
≥500	1.00	1.00
350-499	1.28 (1.06 - 1.55)	2.51 (1.77 - 3.56)
200-349	2.22 (1.84 - 2.66)	6.37 (4.66 - 8.72)
100-199	4.74 (3.79 - 5.93)	21.21 (15.59 - 28.85)
50-99	7.07 (5.26 - 9.51)	34.29 (24.10 - 48.77)
0-49	6.42 (4.87 - 8.46)	44.21 (30.90 - 63.24)

Incidence rate ratios derived from multivariable Poisson regression of the association between time-updated CD4 count and TB disease, stratified by ART status. 62,684 PLHIV were included in this analysis; 32,319 were excluded from the model due to missing data on ethnicity and country of birth, route of HIV infection, CD4 count or age at HIV diagnosis. Model adjusted for the variables in the multivariable model in Table 4.2. ART: anti-retroviral therapy, CI: confidence interval, IRR: incidence rate ratio, TB: tuberculosis.

When stratifying by ART initiation status, lower time-updated CD4 count was strongly associated with increased TB rates (Table 4.4). For individuals who had never initiated ART, the IRR for TB increased with decreasing CD4 count to 6.42 [4.87-8.46] for 0-49 cells/µl compared to ≥500 cells/µl. The increased risk at low CD4 count was higher in individuals who had initiated ART, with an IRR of 44.21 [30.90-63.24] for 0-49 cells/µl, compared to ≥500 cells/µl.

4.3.4 Anti-retroviral therapy initiation and discontinuation

In a post-hoc analysis of patients who had initiated ART, PLHIV who developed TB were more likely to have discontinued ART at their last clinic visit (27%, versus 6% of those without TB, P<0.001, Table 4.5). Rates of ART initiation, and the time period from the most recent clinic visit to the end of the study were similar for MSM, heterosexuals and PWID (Table 4.5).

4.3.5 Sensitivity analysis

Sensitivity analyses were conducted as follows: (1) excluding 241 individuals who acquired HIV infection through mother-to-child transmission, (2) excluding 595 individuals with TB whose probabilistic matching scores (linking to their HIV record) were in the lowest quartile, (3) excluding 137 individuals with TB who were matched to their HIV record using the three lowest-ranked deterministic criteria, (4) excluding 424 individuals diagnosed with TB 92-182 days after HIV diagnosis, (5) including IMD score and excluding data on 12,432 individuals missing IMD score. All analyses provided consistent results with the main model (Table 4.6, Table 4.7, Table 4.8).

Table 4.5: ART initiation rates, median time from date of last ART status to study end, and rates of ART discontinuation at the last clinic visit prior to the end of the study.

			Initiated ART by end of study	Median months from last ART status to study end	Patients who had initiated ART but were not on ART at their last clinic visit prior to study end*
Probable route of HIV infection	**Outcome**	**n**	**n (%)**	**n (95% CI)**	**n (%)**
All groups	No TB	61,093	52,207 (86%)	6.9 (2.8-16.3)	2,965 (6%)
	TB	1,591	1,095 (69%)	9.4 (3.6-20.3)	294 (27%)
Men who have sex with men	No TB	27,566	23,268 (84%)	6.3 (2.5-12.9)	1,058 (5%)
	TB	184	118 (64%)	8.0 (2.9-16.1)	35 (30%)
Heterosexual sex	No TB	31,773	27,476 (87%)	7.6 (3.1-26.8)	1,815 (7%)
	TB	1,311	913 (70%)	9.6 (3.6-21.5)	250 (27%)
People who inject drugs	No TB	1,297	1,058 (82%)	7.9 (2.8-34.9)	72 (7%)
	TB	77	47 (61%)	9.6 (3.8-21.5)	6 (13%)

*** Study end was TB diagnosis, death, or 31/12/2014, whichever was soonest. Patients with unknown ART status were classed as not being on ART. ART: anti-retroviral therapy, CI: confidence interval, TB: tuberculosis.**

Table 4.6: Sensitivity analyses of multivariable Poisson regression models of factors associated with incident TB disease among PLHIV in England, Wales and Northern Ireland from 2000 to 2014.

	Mother-to-child transmission excluded [1]	Weakest probabilistic matches (lowest 25%) excluded [2]	Weakest deterministic matches excluded [3]	182 day cut-off for 'simultaneous' diagnosis [4]	IMD decile included [5]
	IRR (95% CI)	IRR (95% CI)	IRR (95% CI)	IRR (95% CI)	IRR (95% CI)
Route of HIV infection					
MSM	1·00	1·00	1·00	1·00	1·00
Heterosexual men	1·79 (1·45 - 2·21)	1·63 (1·28 - 2·09)	1·84 (1·48 - 2·30)	1·83 (1·45 - 2·31)	1·83 (1·44 - 2·31)
Heterosexual women	1·95 (1·59 - 2·41)	1·94 (1·53 - 2·46)	2·04 (1·64 - 2·53)	1·96 (1·56 - 2·46)	1·97 (1·56 - 2·48)
Men who inject drugs	5·42 (4·03 - 7·29)	4·67 (3·26 - 6·68)	5·73 (4·22 - 7·79)	5·44 (3·91 - 7·56)	4·79 (3·35 - 6·85)
Women who inject drugs	4·89 (2·93 - 8·16)	4·01 (2·11 - 7·62)	5·35 (3·20 - 8·95)	5·76 (3·44 - 9·64)	6·18 (3·49 - 10·93)
Blood/Tissue transfer	3·08 (1·77 - 5·38)	2·82 (1·47 - 5·44)	2·92 (1·61 - 5·32)	3·28 (1·80 - 5·98)	2·79 (1·45 - 5·37)
Mother-to-child	-	2·74 (0·85 - 8·78)	4·03 (1·61 - 10·04)	3·78 (1·37 - 10·44)	5·58 (2·22 - 14·02)
Ethnicity/Country of birth					
White, UK-born	1·00	1·00	1·00	1·00	1·00
Black African, UK-born	1·94 (1·09 - 3·47)	1·66 (0·80 - 3·45)	2·00 (1·12 - 3·58)	1·54 (0·77 - 3·07)	1·67 (0·83 - 3·33)
Other ethnicity, UK-born	1·82 (1·23 - 2·70)	2·02 (1·29 - 3·15)	1·76 (1·17 - 2·65)	1·78 (1·16 - 2·71)	1·97 (1·28 - 3·05)
Ethnicity unknown, UK-born	-†	-†	-†	-†	-†
Born in low-TB incidence country	1·25 (0·96 - 1·63)	1·28 (0·94 - 1·75)	1·15 (0·87 - 1·52)	1·16 (0·87 - 1·55)	1·22 (0·91 - 1·64)
White, born in high-TB incidence country	2·02 (1·40 - 2·92)	2·37 (1·57 - 3·56)	1·93 (1·32 - 2·82)	1·90 (1·27 - 2·85)	1·85 (1·23 - 2·79)
Black African, born in high-TB incidence country	4·01 (3·21 - 5·02)	4·17 (3·21 - 5·42)	3·92 (3·12 - 4·93)	3·84 (3·02 - 4·89)	3·55 (2·75 - 4·57)
Other ethnicity, born in high-TB incidence country	3·27 (2·50 - 4·28)	3·29 (2·40 - 4·51)	3·01 (2·27 - 3·99)	3·13 (2·33 - 4·20)	3·12 (2·31 - 4·22)
Ethnicity unknown, born in high-TB incidence country	1·03 (0·14 - 7·42)	-†	1·07 (0·15 - 7·70)	1·35 (0·19 - 9·71)	2·30 (0·32 - 16·53)

	Mother-to-child transmission excluded [1]	Weakest probabilistic matches (lowest 25%) excluded [2]	Weakest deterministic matches excluded [3]	182 day cut-off for 'simultaneous' diagnosis [4]	IMD decile included [5]
	IRR (95% CI)	IRR (95% CI)	IRR (95% CI)	IRR (95% CI)	IRR (95% CI)
White, country of birth unknown	0·44 (0·25 - 0·80)	0·36 (0·17 - 0·78)	0·46 (0·26 - 0·84)	0·48 (0·26 - 0·89)	0·60 (0·33 - 1·09)
Other ethnicity, country of birth unknown	1·36 (0·99 - 1·86)	1·28 (0·88 - 1·87)	1·36 (0·99 - 1·88)	1·32 (0·94 - 1·87)	0·97 (0·65 - 1·43)
CD4 count					
≥500	*	*	*	*	*
350-499					
200-349					
100-199					
50-99					
0-49					
Ever on ART					
No	*	*	*	*	*
Yes					
Age at HIV diagnosis					
15-24	0·88 (0·73 - 1·05)	0·87 (0·71 - 1·07)	0·86 (0·71 - 1·04)	0·89 (0·73 - 1·08)	0·88 (0·70 - 1·09)
25-34	1·02 (0·91 - 1·15)	0·98 (0·86 - 1·12)	1·04 (0·92 - 1·17)	1·01 (0·89 - 1·15)	1·12 (0·97 - 1·29)
35-44	1·00	1·00	1	1·00	1·00
45-64	1·11 (0·95 - 1·31)	1·13 (0·94 - 1·36)	1·09 (0·92 - 1·29)	1·03 (0·86 - 1·24)	1·30 (1·09 - 1·56)
≥65	0·92 (0·50 - 1·67)	0·78 (0·37 - 1·66)	0·90 (0·48 - 1·68)	0·85 (0·42 - 1·71)	1·95 (1·03 - 3·67)
Year of HIV diagnosis					
(for each year increase from 2000)	1·02 (1·01 - 1·04)	1·02 (1·01 - 1·04)	1·03 (1·01 - 1·04)	1·00 (0·98 - 1·02)	1·01 (0·99 - 1·03)
IMD decile					

	Mother-to-child transmission excluded [1]	Weakest probabilistic matches (lowest 25%) excluded [2]	Weakest deterministic matches excluded [3]	182 day cut-off for 'simultaneous' diagnosis [4]	IMD decile included [5]
	IRR (95% CI)	IRR (95% CI)	IRR (95% CI)	IRR (95% CI)	IRR (95% CI)
1	-	-	-	-	0.92 (0.77 - 1.09)
2	-	-	-	-	0.89 (0.74 - 1.07)
3	-	-	-	-	0.85 (0.68 - 1.05)
4	-	-	-	-	0.81 (0.63 - 1.04)
5	-	-	-	-	0.94 (0.73 - 1.21)
6	-	-	-	-	0.92 (0.67 - 1.26)
7	-	-	-	-	0.84 (0.59 - 1.18)
8	-	-	-	-	0.65 (0.43 - 1.00)
9	-	-	-	-	1.03 (0.71 - 1.49)
10	-	-	-	-	0.92 (0.77 - 1.09)

[1] 241 individuals who acquired HIV infection through mother-to-child transmission were excluded. [2] 2,057 of 2,187 individuals with TB and HIV were linked via probabilistic matching. The lowest quartile of probabilistic match scores (595 individuals) were excluded. [3] 1951 of 2,187 individuals with TB and HIV were linked using 8 hierarchical deterministic matching criteria. 137 individuals matched using the three lowest-ranked criteria were excluded. [4] 424 individuals diagnosed with TB 92-182 days after HIV diagnosis were excluded. [5] IMD score was included in the analysis, 12,432 individuals with no data on their IMD score were excluded. *An interaction was present between time-updated CD4 count and time-updated ART status, see Table 4.7 and Table 4.8. †Not calculated as numerator was zero. ART: anti-retroviral therapy, CI: confidence interval, IMD: index of multiple deprivation (only available for patients in England and Wales), IRR: incidence rate ratio, MSM: men who have sex with men, PWID: people who inject drugs, PY: person-years, TB: tuberculosis.

Table 4.7: Sensitivity analyses of multivariable Poisson regression models of the association between time-updated CD4 count and TB disease, stratified by ART status, among PLHIV in England, Wales and Northern Ireland from 2000 to 2014.

	Mother-to-child transmission excluded [1]		Weakest probabilistic matches (lowest 25%) excluded [2]		Weakest deterministic matches excluded [3]		182 day cut-off for 'simultaneous' diagnosis [4]		IMD decile included [5]	
	Ever on ART		Ever on ART		Ever on ART		Ever on ART		Ever on ART	
	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
CD4 count	IRR (95% CI)	IRR (95% CI)	IRR (95% CI)	IRR (95% CI)	IRR (95% CI)	IRR (95% CI)	IRR (95% CI)	IRR (95% CI)	IRR (95% CI)	IRR (95% CI)
≥500	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00
350-499	2·22 (1·31 - 3·76)	2·98 (1·82 - 4·88)	2·05 (1·13 - 3·73)	2·88 (1·65 - 5·03)	2·09 (1·22 - 3·59)	2·91 (1·75 - 4·82)	2·27 (1·32 - 3·89)	2·99 (1·80 - 4·94)	2·29 (1·27 - 4·14)	2·71 (1·56 - 4·69)
200-349	2·2 (1·35 - 3·60)	5·83 (3·68 - 9·25)	2·01 (1·15 - 3·51)	5·62 (3·34 - 9·45)	2·23 (1·35 - 3·69)	5·91 (3·69 - 9·45)	2·21 (1·33 - 3·68)	5·29 (3·29 - 8·53)	1·91 (1·11 - 3·31)	5·31 (3·19 - 8·86)
100-199	2·35 (1·44 - 3·84)	18·46 (11·74 - 29·04)	2·25 (1·29 - 3·92)	17·89 (10·73 - 29·83)	2·5 (1·52 - 4·13)	18·99 (11·98 - 30·12)	2·66 (1·60 - 4·42)	16·98 (10·64 - 27·09)	1·85 (1·07 - 3·19)	16·95 (10·27 - 27·98)
50-99	2·35 (1·33 - 4·15)	29·15 (17·27 - 49·20)	2·05 (1·05 - 3·97)	25·23 (13·73 - 46·37)	2·35 (1·32 - 4·21)	29·42 (17·26 - 50·14)	2·6 (1·39 - 4·86)	22·58 (12·81 - 39·81)	1·92 (1·01 - 3·65)	27·35 (15·25 - 49·04)
0-49	3·84 (2·23 - 6·63)	46·89 (28·31 - 77·68)	3·47 (1·87 - 6·45)	45·33 (25·58 - 80·34)	3·77 (2·16 - 6·60)	45·99 (27·44 - 77·08)	4·38 (2·42 - 7·92)	36·91 (21·50 - 63·37)	2·56 (1·38 - 4·78)	39·07 (21·98 - 69·44)

Incidence rate ratios derived from multivariable Poisson regression models of the association between time-updated CD4 count and TB disease, stratified by ART status. Models adjusted for the variables in the multivariable model in Table 4.6. [1] 241 individuals who acquired HIV infection through mother-to-child transmission were excluded. [2] 2,057 of 2,187 individuals with TB and HIV were linked via probabilistic matching. The lowest quartile of probabilistic match scores (595 individuals) were excluded. [3] 1951 of 2,187 individuals with TB and HIV were linked using 8 hierarchical deterministic matching criteria. 137 individuals matched using the three lowest-ranked criteria were excluded. [4] 424 individuals diagnosed with TB 92-182 days after HIV diagnosis were excluded. [5] IMD score was included in the analysis, 12,432 individuals with no data on their IMD score were excluded. *An interaction was present between time-updated CD4 count and time-updated ART status, see Table 4.7 and Table 4.8. †Not calculated as numerator was zero. ART: anti-retroviral therapy, CI: confidence interval, IMD: index of multiple deprivation, IRR: incidence rate ratio, MSM: men who have sex with men, PWID: people who inject drugs, PY: person-years, TB: tuberculosis.

Table 4.8: Sensitivity analyses of multivariable Poisson regression models of the association between time-updated ART status and TB disease, stratified by CD4 count, among PLHIV in England, Wales and Northern Ireland from 2000 to 2014.

		CD4 count ≥500	350-499	200-349	100-199	50-99	0-49
Sensitivity analysis	**Ever on ART**	**IRR (95% CI)**	**IRR (95% CI)**	**IRR (95% CI)**	**IRR (95% CI)**	**IRR (95% CI)**	**IRR (95% CI)**
Mother-to-child transmission excluded [1]	No	1·00	1·00	1·00	1·00	1·00	1·00
	Yes	0·03 (0·02 - 0·05)	0·07 (0·05 - 0·10)	0·07 (0·06 - 0·09)	0·08 (0·06 - 0·10)	0·08 (0·05 - 0·11)	0·13 (0·09 - 0·18)
Weakest probabilistic matches (lowest 25%) excluded [2]	No	1·00	1·00	1·00	1·00	1·00	1·00
	Yes	0·04 (0·02 - 0·06)	0·07 (0·05 - 0·11)	0·07 (0·05 - 0·10)	0·08 (0·06 - 0·11)	0·07 (0·05 - 0·12)	0·12 (0·08 - 0·19)
Weakest deterministic matches excluded [3]	No	1·00	1·00	1·00	1·00	1·00	1·00
	Yes	0·03 (0·02 - 0·05)	0·07 (0·05 - 0·10)	0·08 (0·06 - 0·10)	0·09 (0·07 - 0·11)	0·08 (0·05 - 0·12)	0·13 (0·09 - 0·18)
182 day cut-off for 'simultaneous' diagnosis [4]	No	1·00	1·00	1·00	1·00	1·00	1·00
	Yes	0·03 (0·02 - 0·05)	0·08 (0·06 - 0·11)	0·08 (0·06 - 0·10)	0·09 (0·07 - 0·12)	0·09 (0·06 - 0·14)	0·15 (0·10 - 0·22)
IMD decile included [5]	No	1·00	1·00	1·00	1·00	1·00	1·00
	Yes	0·03 (0·02 - 0·05)	0·07 (0·05 - 0·11)	0·06 (0·05 - 0·08)	0·06 (0·04 - 0·08)	0·06 (0·04 - 0·10)	0·08 (0·05 - 0·12)

Incidence rate ratios derived from multivariable Poisson regression models of the association between time-updated CD4 count and TB disease, stratified by ART status. Models adjusted for the variables in the multivariable model in Table 4.6. [1] 241 individuals who acquired HIV infection through mother-to-child transmission were excluded. [2] 2,057 of 2,187 individuals with TB and HIV were linked via probabilistic matching. The lowest quartile of probabilistic match scores (595 individuals) were excluded. [3] 1951 of 2,187 individuals with TB and HIV were linked using 8 hierarchical deterministic matching criteria. 137 individuals matched using the three lowest-ranked criteria were excluded. [4] 424 individuals diagnosed with TB 92-182 days after HIV diagnosis were excluded. [5] IMD score was included in the analysis, 12,432 individuals with no data on their IMD score were excluded. *An interaction was present between time-updated CD4 count and time-updated ART status, see Table 4.7 and Table 4.8. †Not calculated as numerator was zero. ART: anti-retroviral therapy, CI: confidence interval, IMD: index of multiple deprivation, IRR: incidence rate ratio, MSM: men who have sex with men, PWID: people who inject drugs, PY: person-years, TB: tuberculosis.

4.4 Discussion

4.4.1 Summary of findings

Within our cohort of over 100,000 adults diagnosed with HIV from 2000 to 2014, six percent had a diagnosis of TB. Over half of these were simultaneously diagnosed with TB and HIV. Simultaneous diagnoses were most common among people who acquired their HIV infection via heterosexual sex, and/or who were born in countries with high TB incidence, and least common among MSM and PWID.

The incidence of TB in PLHIV after HIV diagnosis in this study was 344/100,000 PY, thirty times higher than the national TB incidence in the general adult (≥15 years) population of 11.5 per 100,000 population per year.[52] For patients not simultaneously diagnosed with HIV and TB, the risk of developing TB was highest in the first two years after HIV diagnosis. The differences in incidence between MSM, PWID, and people who acquired HIV infection through heterosexual sex, and between patients of different ethnicity, were also greatest in the first two years after HIV diagnosis. After two years TB incidence was relatively consistent across different risk groups.

The risk of developing TB following HIV diagnosis was highest for people who acquired HIV through injecting drug use (largely UK-born patients), and incidence rates in this population were comparable to those in black Africans born in high TB incidence countries; almost five-fold more than MSM after accounting for other factors including starting ART. TB incidence among MSM was ten times higher than in the general population (111/100,000 compared to 11.5/100,000)[52], whilst TB incidence in black African PLHIV was more than six times that in the general population (644/100,000 compared to 102/100,000 in black African people born outside the UK [irrespective of HIV status], [52] although these estimates are not directly comparable as the national figures for the general population were not stratified by TB incidence in country of birth).

Consistent with previous research, higher CD4 count [119, 164] and initiation of ART [200] were associated with lower incidence of TB.

4.4.2 Strengths and limitations of the study

This study benefitted from the very large national HIV-positive cohort, providing comprehensive results for England, Wales and Northern Ireland. The large sample size resulted in strong statistical evidence that the associations between the risk of TB and route of HIV infection, ethnicity, country of birth, CD4 count, and ART initiation (all $P<0.001$) did not occur by chance. The large sample also enabled composite variables to be created for sex and route of HIV infection, and for ethnicity and TB incidence in country of birth, to avoid these known interactions within the model, and to test for the interaction between CD4 count, ART and risk of TB. The algorithm linking patients with TB and HIV utilised ethnicity, year and country of birth; all variables with very high completeness: 97.3%, 99.9% and 90.5% respectively.

There were no substantial differences in the demographics or proportion of TB in patients missing data on each of these variables; however patients missing data on one variable were more likely to have other missing data. Patients missing data for multiple variables were less likely to be linked to a TB notification meaning that TB incidence may have been underestimated. It is likely that the low incidence of TB in patients with unknown ethnicity or route of HIV infection was a symptom of this, despite the fact that patients with extensive missing data may be more likely to be from populations at high risk for TB. Additionally, the record linkage algorithm is less sensitive to non-English names,[184] and therefore TB incidence in foreign populations may have been underestimated.

One limitation was missing CD4 count data for approximately a third of patients, who were therefore excluded from the risk factor analysis. This was partly due to difficulties linking data because of missing or incomplete PII in the CD4 database, and partly because some large hospitals do not supply CD4 count data to HARS. However,

there was no evidence that patients with missing CD4 count data were systematically different to the analysis cohort. As the sample size remained very large, it was decided to conduct a complete-case analysis rather than using multiple imputation. The decision was also influenced by the complexity of the dataset as a result of using time-updated data with intervals of varying lengths for CD4 count and ART initiation, and the fact that patients with missing data for one variable were also more likely to be missing data on other variables (leaving few complete variables from which to impute). There was also no evidence that patients missing data were systematically different to those with complete case data; although patients with missing data were less likely to be linked to a TB case, this was probably because they were also missing PII (used in the record linkage process) rather than because they were less likely to develop TB disease. As a result, it is likely that excluding records with incomplete data biased the results towards the null, and that the true effect sizes are larger than estimated. As this analysis used routinely collected surveillance data, all diagnoses were made according to standard definitions, which should have reduced measurement error and misclassification. TB diagnoses later found not to be *MTBC* were de-notified and not included in the analysis.

Composite variables were created for sex/route of HIV infection, and for ethnicity/country of birth, to avoid confounding as there are known associations between these variables; e.g. MSM cannot be women and white PLHIV are less likely to be born abroad. However, the results presented may have been confounded by other factors. The increased risk of TB among people who acquired HIV through injecting drug use may have been confounded by other social risk factors which increase susceptibility or exposure to TB, such as homelessness and imprisonment. Unfortunately no data were available on these risk factors, or on other sources of exposure to TB such as household contacts with TB, occupational exposure, or travel to areas where TB is endemic.

There are causal relationships between being on ART, CD4 count, and viral load (Figure 4.2). Whilst there was data on all CD4 counts for the majority of patients, data on viral load and whether or not a patient is on ART were only reported annually, from their most recent clinic visit; which is not necessarily representative of the whole year. As there was higher frequency data for CD4 count, viral load was excluded from the model, and a binary variable for whether individuals had ever started ART was used in order to utilise the time-updated CD4 count data without creating causal loops which could not be adequately reflected in the data. This enabled a more accurate assessment of the risk of TB at different CD4 counts; however it also meant that the effect of viral load on the risk of developing TB could not be assessed in the multivariable model. Data were available on ART discontinuation, but were of poor quality and could not be included in the model. Consequently, the association between starting ART and lower TB incidence may have been underestimated because it was assumed that all individuals remained on treatment for the duration of the study. Whether the patient was on ART at their last clinic visit during the study period was examined, however these figures are also likely to underestimate ART discontinuation as they cannot account for patients who have stopped and then restarted ART, or PLHIV who have dropped out of care completely.

Individuals entered the study cohort 92 days after HIV diagnosis or first presentation to UK health services; therefore TB incidence in people diagnosed abroad who were at risk prior to entering the UK may have been underestimated, as TB cases diagnosed during the initial period following HIV diagnosis when TB incidence is highest would have been missed. A recent study of PLHIV had 18% loss to follow-up over 4 years, and 14% of TB cases diagnosed more than 91 days after HIV diagnosis were in these patients.[196] As TB and HIV are sometimes treated (and usually reported) separately in the UK, and notification of TB cases is mandatory, dropping out of HIV care does not prevent notification of a TB diagnosis. This study therefore used passive censoring, continuing follow-up until 31/12/2014 rather than the date last seen

for HIV care. Consequently, migration out of the UK may have caused an underestimation of TB incidence.

A possible limitation of the Poisson regression model was censoring due to competing risks, specifically deaths from non-TB causes. However, very few patients died (3%) and the median time to death was 3.4 years, substantially longer than the median time to TB diagnosis (1.8 years). Therefore, any impact of such censoring is likely to be minimal.

There may have been some misclassification of the 'route of HIV diagnosis' variable. As this variable was assigned at HIV diagnosis; the PWID category does not necessarily reflect current injecting drug use at the time of TB diagnosis. Phylogenetic analyses have estimated 1-11% of male heterosexual PLHIV may be misclassified MSM, and this may be as high as 21% for black African men.[201] Consequently, the TB risk in black African MSM may have been underestimated due to some MSM not disclosing their behaviour. However, this misclassification is unlikely to affect which interventions are most effective; MSM who have been misclassified as heterosexual are likely to be also missed by HIV interventions which target MSM, and the low overall risk of TB for MSM means that targeting TB interventions to this population is unlikely to be cost-effective. Increased screening for both HIV and TB among black African communities may be more likely to be effective.

4.4.3 Implications of the research

PWID represented less than 2% of PLHIV in this cohort, but accounted for 3% of TB cases and more than 4% of cases diagnosed more than 91 days after HIV diagnosis. TB incidence in PWID in this study (876/100,000 PY in men and 605/100,000 in women) was higher than the 420/100,000 reported in a cohort of German PLHIV,[118] possibly because that cohort utilised active rather than passive follow-up and excluded patients who did not present to care for 6 months or more, who

may be more likely to develop TB disease than patients who remain engaged with care.

PWID are typically diagnosed with HIV late,[202] have slower rates of linkage to care and lower rates of viral suppression,[10] all of which may contribute to increased risk of TB. Rates of ART initiation and the time from the last clinic visit to the end of the study were comparable for MSM, heterosexuals and PWID; and PWID did not have higher rates of ART discontinuation at their last clinic visit prior to study end (Table 4.5). Consequently, it seems high rates of TB among PWID were caused by difficulties in linking to care and not lack of engagement with health services once linked. Increasing HIV testing for PWID and supporting them to link to HIV care could help to diagnose HIV infection sooner and reduce the loss of immune function, and thereby reduce the incidence of TB.

Many PWID have other co-morbidities which may cause immunosuppression, make HIV care more challenging, or be associated with increased risk of TB.[203] Additionally there are high rates of alcoholism and homelessness, and living in hostels is common.[204] These factors, in addition to injecting drugs in shared social settings, may drive close mixing of people with similar risk factors for TB disease resulting in TB transmission. High rates of smoking may also have impacts on both local lung immunity and TB transmission. Further studies are needed to explore the impact of these factors and to design effective interventions.

The majority of PWID were white (51%) and born in the UK or low TB incidence countries (72%). It is therefore likely that most TB in this group was acquired in the UK, meaning these cases may be preventable by diagnosing HIV sooner and ensuring prompt ART initiation. More could also be done to diagnose TB cases sooner; the impact of active case finding in PLHIV should be evaluated. In contrast, people who acquired HIV through heterosexual sex were typically of black African ethnicity (61%) and born in high TB incidence countries (69%), both populations which also have high

rates of TB among HIV-negative people. It is likely that their TB infections were acquired abroad, limiting the ability to prevent TB infection if they present with clinical TB at the time of HIV diagnosis.[205] As more than 60% of people who acquired HIV heterosexually were diagnosed with TB simultaneously or prior to HIV diagnosis, greater efforts to diagnose these HIV infections and initiate ART would reduce TB in this population. A greater focus on screening and treating latent TB infection (LTBI) could also prevent these cases.[206] There is little data available on the prevalence of LTBI and the use of preventive therapy among PLHIV in the UK. Rates of LTBI screening and uptake of preventive therapy vary substantially between HIV clinics,[137, 207] and a survey of UK HIV healthcare providers providing care to 90% of PLHIV in the UK found that only 54% offered LTBI screening and preventive therapy.[208] Health economics evaluations would be useful to determine the most effective screening measures for these populations.

Over half of all TB cases (55%) were diagnosed simultaneously with HIV infection, and of the 39% diagnosed later, the probability of a TB diagnosis was highest in the first year following HIV diagnosis (Figure 4.3). This suggests that TB disease is largely the result of TB infection acquired prior to HIV diagnosis. This could result from late diagnosis of existing active TB, particularly in migrants who have recently moved to the UK from high-burden countries and whose TB is largely attributable to reactivation of remotely acquired infection.[209] Additionally, the incidence of TB among migrants decreases with time since entry to the UK, as new TB infection is less likely in the UK than their country of origin. Other factors which could explain this trend are increased surveillance for opportunistic infections following HIV diagnosis, or ‘unmasking-type’ immune reconstitution inflammatory syndrome (IRIS) as a consequence of ART.[108]

The risk of TB in the first year after HIV diagnosis was particularly high for PLHIV of black African or Indian, Pakistani or Bangladeshi ethnicity, although these groups also saw the largest increases in cumulative risk over the remainder of the study period (Figure 4.5). Similarly, people who acquired HIV through heterosexual sex

or injecting drug use both had the greatest risk of developing TB in the first year following HIV diagnosis, and remained more likely to develop TB throughout the study period (Figure 4.4). This suggests that LTBI screening of these high-risk groups immediately following HIV diagnosis (and preventive therapy where appropriate) could have a substantial impact in reducing TB incidence after HIV diagnosis. Whilst TB incidence was lower after the first year following HIV diagnosis (Table 4.1), 25% of all TB cases occurred more than one year after HIV diagnosis. There is an even greater opportunity to prevent these cases with LTBI treatment, regardless of whether they were the result of reactivation of LTBI or more recent TB infection.

The British HIV Association (BHIVA) currently recommends IGRA testing and treatment of LTBI among PLHIV using criteria based on CD4 count, time on ART and TB incidence in country of birth.[110] In this study, the incidence of TB among PWID was comparable to that of black African patients born in countries with high TB incidence. Extending the BHIVA guidelines to include PWID, and providing additional screening and preventive therapy for PWID with LTBI should be considered to reduce TB incidence in this population.

Recent clinical trials [16, 17] indicated a protective effect against TB when HIV-positive individuals started ART as soon as possible rather than deferring until a lower CD4 threshold is reached. Our findings were consistent with this; patients who had initiated ART had greatly reduced rates of TB compared to those who had not (Table 4.3). BHIVA guidelines have recently been updated and now recommend all PLHIV initiate ART, regardless of CD4 count.[210] If implemented, these guidelines could lead to a significant reduction in TB cases in PLHIV, as many of the TB cases in ART-naïve patients may have been preventable with earlier initiation of ART. Whilst the majority of TB cases had not initiated ART at the time of their TB diagnosis (Table 4.1), some TB cases did still occur in patients on ART. These may have been existing disease that was previously undiagnosed, due to the limited sensitivity of some diagnostic tests for

LTBI in immunocompromised individuals. Other patients may have had subclinical disease when initiating ART, resulting in unmasking IRIS.

The risk of developing TB was greatest in the first year after HIV diagnosis for patients in all CD4 strata, but there was a substantial increase in the risk of TB among PLHIV who were diagnosed with HIV late (defined as a CD4 count below 350 cells/μl at diagnosis). This risk remained elevated throughout the study period (Figure 4.6) and was also associated with lower CD4 count; the cumulative risk was substantially higher for patients with CD4 count below 200 cells/μl at diagnosis than for patients with a CD4 count of 200-349. The cumulative risk appeared higher for patients with a CD4 count of 100-199 at diagnosis than for PLHIV diagnosed with a CD4 count below 100 cells/μl, but this is likely because there were a greater proportion of simultaneous HIV and TB diagnoses among PLHIV diagnosed with CD4 <100 cells/μl than for patients with CD4 of 100-199 at diagnosis (Table 4.1). Patients with very low CD4 counts at HIV diagnosis (<100 cells/μl) may also have initiated ART more rapidly, thereby reducing their ongoing risk of TB.

Time-updated CD4 count and ART initiation status interacted within the multivariable model. This could be a result of differing CD4 trajectories for patients who have or have not initiated ART; CD4 counts of patients who are on ART should be on an upwards trajectory, whereas CD4 count is likely to be declining for patients who are not on ART. Higher rate ratios for TB at low CD4 count in people on ART could also be attributable to late ART start (i.e. long periods of low CD4 count prior to initiating ART and then little time on ART prior to TB diagnosis), or due to ART discontinuation. The SMART trial demonstrated an association between stopping ART and increased risk of opportunistic disease and death.[211] The post-hoc analysis of patients who had started ART demonstrated that patients who went on to develop TB were more likely to have discontinued ART at their last study visit than individuals who remained TB-free (Table 4.5). This suggests ART discontinuation leaves patients at risk of new TB disease. Improving understanding of why PLHIV stop taking ART, and developing

interventions to support PLHIV to remain on ART long-term could also help reduce TB incidence among PLHIV.

4.4.4 Conclusions

In England, Wales and Northern Ireland, PLHIV who acquired HIV by injecting drugs had higher rates of TB after their HIV diagnosis than MSM, comparable to black Africans born in countries with high TB incidence. High rates of TB in PWID are likely to result from transmission within the UK. ART is highly protective against TB, but the majority of TB diagnoses were in people who have never started ART. ART discontinuation rates were much higher in people who subsequently developed TB than those who did not. Earlier diagnosis of HIV, quicker initiation of ART (as per the recently updated BHIVA guidelines), and improving retention in care and ART continuation should decrease incident TB in PLHIV.

5 Trends in factors associated with HIV co-infection among tuberculosis patients

5.1 Introduction

5.1.1 HIV co-infection of tuberculosis patients in the UK

The number of TB cases reported in England and Wales increased from 2000, peaking at 8,280 cases in 2011 (15.6/100,000 population), but then declined by a third to 5,758 in 2015.[52] HIV co-infection contributed substantially to the rise in TB from 1999 to 2003; 31% of the increase in TB was in people with HIV co-infection and by 2003 the prevalence of HIV co-infection was 8.3%.[123] Whilst the number of PLHIV in the UK increased by 10% to over 100,000 in 2015,[10] HIV co-infection among TB cases decreased to 3.2% (197/6,209) by 2014.[52] However, both the proportion and number of co-infected individuals remain among the highest in western Europe,[52, 116] and trends in TB-HIV epidemiology and risk factors for co-infection have not been recently described.

5.1.2 Known risk factors for HIV co-infection

A previous study of HIV among TB cases in England and Wales reported higher prevalence of HIV co-infection among women than men (6.1% vs. 5.2%), those aged 30-39 years at TB diagnosis (12.9%, vs. 5.6% for those aged 20-29, 8.1% for those aged 40-49, and an even lower prevalence outside of these age brackets).[123] Patients of black African ethnicity or who were born in Africa also had high rates of co-infection (20.5% and 19.8% respectively); overall the prevalence of HIV among TB patients born outside of the UK was 7.3%, compared to 1.5% in UK-born patients. The HIV prevalence was lower in TB patients who had been in the UK for more than five years (2.9%) than those who had arrived within the last five years (10.7%).[123] However, this study did not present any adjusted estimates of the risk or consider confounding or overlap in risk factors. Another study reported that HIV prevalence was

higher among healthcare workers with TB than other TB patients, particularly for nurses, women, and patients born in sub-Saharan Africa.[212]

A comparable study using surveillance data from the Netherlands reported that HIV co-infection was lower among women than men (2.6% vs. 5.1%, multivariable OR 0.62 [0.50-0.77]).[213] In this study HIV prevalence among TB patients from Africa (4.9%) was not significantly higher in a multivariable analysis than in patients born in the Netherlands (3.8%), possibly because African TB patients in the Netherlands tend to be from north Africa,[214] where TB is less prevalent than sub-Saharan Africa. However, HIV prevalence was higher among patients born outside the Netherlands in western Europe, north America, Israel or Australia (23.1% co-infected, multivariable OR 3.05 [2.06-4.50]) but lower among patients from Asia (1.2%, multivariable OR 0.22 [0.15-0.34]) or eastern Europe (1.7%, multivariable OR 0.43 [0.20-0.94]).[213] There was no significant association between HIV and site of TB disease.

A Europe-wide study of data reported to ECDC up until the end of 2014 reported a higher risk of HIV co-infection for men than women (4.2% vs. 4.3%, multivariable OR 1.25 [1.07-1.46]).[117] Compared to TB patients from within the EU/EEA (4.2% co-infected), the risk of HIV was higher for patients from other European countries (8.5%, multivariable OR 1.91 [1.15-3.17]) and Africa (17.7%, OR 3.28 [2.35-4.57]), but lower for patients from other regions (not including the Americas, 3.7%, multivariable OR 0.64 [0.42-0.98]). Patients with only extra-pulmonary disease were more likely to have HIV than patients with pulmonary disease (7.3% vs. 4.5%, multivariable OR 1.81 [1.50-2.19]), although extra-pulmonary disease varies in severity by site, and this study did not stratify by site.[117] In contrast, a prior UK study reported no association between extra-pulmonary TB and HIV co-infection compared with pulmonary TB alone, but found that HIV co-infection was a risk factor for miliary TB (OR 4.5 [3.4-5.8]),[215] and a 2013 meta-analysis reported a pooled OR of 1.3 (95% CI 1.05-1.6) associated with extra-pulmonary TB for PLHIV compared to HIV-negative people.[216]

Social risk factors such as drug misuse, homelessness and imprisonment may contribute to TB-HIV co-infection. HIV acquisition by injecting drug use is a known risk factor for developing TB;[151] however the interactions between different social risk factors and HIV co-infection have not been investigated in the UK. A study from the Netherlands also found substantially higher HIV prevalence among TB patients who were illegal residents (9.1%, multivariable OR 1.67 [1.08-2.40]), misused drugs (29.2%, OR 5.13 [3.81-6.89]) or were homeless (20.1%, OR 1.55 [1.01-2.40).[213] The proportion of TB patients in England with social risk factors rose from 9.8% in 2010 to 12% in 2015, although the numbers remained relatively constant.[52] Tackling TB in under-served populations is one of the key areas for action in the 2015-2020 Collaborative tuberculosis strategy for England.[217]

5.1.3 Summary and objectives

Understanding the factors associated with TB-HIV co-infection allows us to target screening for both HIV and LTBI to patients most at risk of developing TB disease. Testing for and treating LTBI has been recommended by both BHIVA and the National Institute for Health and Care Excellence (NICE) since 2011 for high-risk PLHIV, with criteria based on TB incidence in country of birth, time on ART and/or CD4 count.[110, 218] Although screening and treating LTBI is cost-effective,[219] a recent audit found that very few eligible patients were being screened.[207] The aims of this chapter were to describe temporal trends in HIV co-infection of TB patients, examine changes in the demographics of co-infected patients and the timing of TB and HIV diagnoses relative to each other, and to identify factors associated with co-infection, focussing on social risk factors (drug and alcohol misuse, homelessness and imprisonment). A secondary aim was to assess whether the current (2011) BHIVA guidelines on testing for HIV and LTBI remain appropriate.

5.2 Methods

5.2.1 Study population

This was a retrospective study of all adult (≥15 years) TB patients in England, Wales and Northern Ireland, notified to PHE's Enhanced TB Surveillance system (ETS) from 1st January 2000 until 31st December 2014.

5.2.2 Outcome: HIV status

Data from ETS was linked to HARS as described in Chapter 3, using a probabilistic matching algorithm (adapted from [184]) with supplementary deterministic matching to accept/reject borderline matches. Patients were considered to be co-infected with HIV if they were diagnosed with HIV or seen for HIV care between 1st January 2000 and 31st December 2014. TB cases notified to HARS as an AIDS-defining illness, which were not linked to a case notification in ETS, could not be included in this analysis. As in section 4.2.2, diagnoses were classed as 'simultaneous' if TB and HIV were diagnosed within 91 days of each other.

5.2.3 Exposure variables

Sociodemographic (sex, ethnicity, HIV prevalence in country of birth, IMD decile, and history of drug misuse, alcohol misuse, homelessness or imprisonment) and clinical (site[s] of disease, year of TB notification, age at TB notification) exposure variables were included. Data on social risk factors (IMD decile, current alcohol misuse, current or previous drug misuse, imprisonment or homelessness) were only available from 2010 onwards. As the prevalence of social risk factors was low, cases reported in 2010 or later with missing data for these variables were considered to not have these risk factors. Drug misuse was defined as problem drug use of illicit injecting drug use or long duration/regular use of illicit opiates, cocaine, and/or amphetamines, and/or daily/almost daily use of cannabis (or synthetic cannabinoids).

Composite variables were created, combining ethnicity and country of birth due to known interactions.[52] As a proxy for HIV exposure, countries of birth outside the

UK were grouped by HIV prevalence; 'high prevalence' was defined as >1% in the adult population living with HIV, as per WHO estimates.[74] Site of TB disease was categorised into three discrete groups; miliary or meningeal TB (with or without pulmonary disease), pulmonary disease with or without other extra-pulmonary disease (excluding miliary or meningeal TB), and other extra-pulmonary disease only. Miliary and meningeal TB were categorised separately as severe forms of disseminated TB may be more likely to develop for patients with HIV co-infection.[215]

5.2.4 Statistical Analysis

5.2.4.1 Descriptive analysis

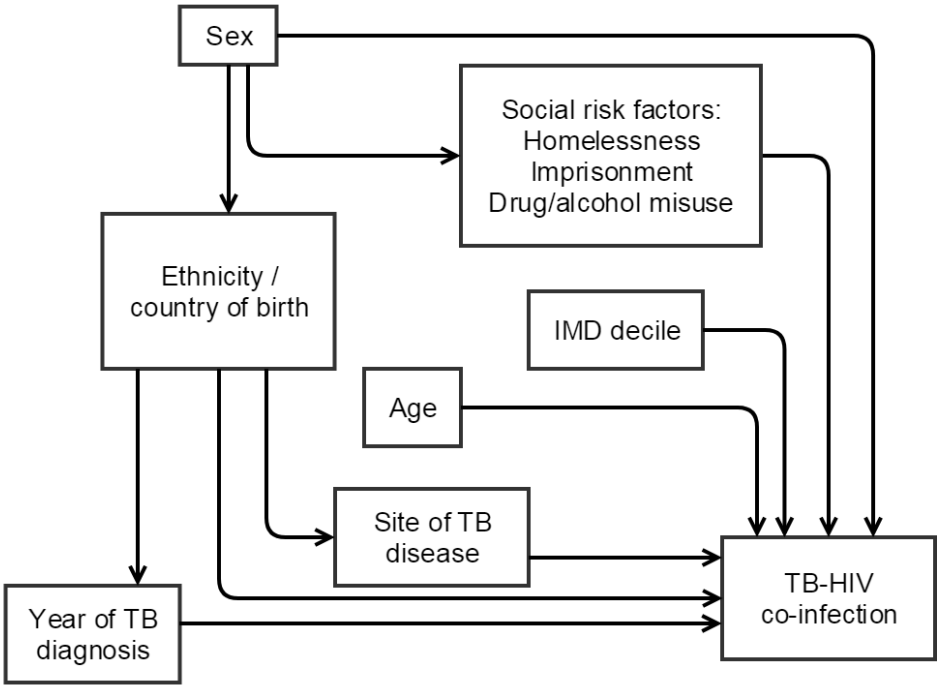
Data were analysed in Stata version 13.1. Descriptive analyses of the cohort were undertaken, examining the proportion of TB cases co-infected with HIV for the whole population and stratified by each exposure variable. Trends in the proportion of TB cases co-infected with HIV over the study period were examined for key variables (age, sex and ethnicity). The proportion of TB cases with different combinations of social risk factors, and the proportion of each of these combinations that were co-infected with HIV, were calculated. Among co-infected patients, trends in the order of TB and HIV diagnoses were examined.

5.2.4.2 Factors associated with HIV co-infection

To investigate factors associated with HIV co-infection, the proportion of cases co-infected with HIV stratified by exposure variable was calculated, and odds ratios were estimated using univariable logistic regression models. Potential confounders were prospectively identified,[199] and a causal framework was built (Figure 5.1). As there was no single 'main' exposure being investigated, there were no confounders in the traditional sense, and therefore the variables included in the multivariable models were informed by the causal framework defined *a priori*. Two multivariable models were built, a 'whole-cohort' model including all cases, and a model including only cases from 2010-2014 so that data on social risk factors could be included. Year of TB notification, age at TB notification, sex, ethnicity and country of birth and site of TB disease were

included in both models; social risk factors were only included in the '2010-2014' model. Linearity (of age group, year and IMD decile) and statistical interactions between social risk factors (drug misuse, alcohol misuse, homelessness and imprisonment) were assessed using likelihood-ratio tests. Variables were treated as categorical if P<0.05, and statistical interactions were considered significant at P<0.05. Patients missing data on one or more variables were excluded. To assess the impact of missing data, the distributions of all demographic factors for cases with missing vs. complete data were compared.

Figure 5.1: Conceptual framework of potential factors associated with HIV co-infection among TB cases.



IMD: index of multiple deprivation, TB: tuberculosis

5.2.4.3 Sensitivity analyses

To examine the impact of a different threshold for 'simultaneous' diagnoses, a sensitivity analysis was conducted using a 30 day threshold rather than 91 days for examining trends in timing of diagnoses. To evaluate the impact of the matching algorithm used on the results of this study, planned sensitivity analyses investigated the impact of excluding weaker matches between ETS and HARS from the two logistic regression models.

5.3 Results

5.3.1 Descriptive epidemiology and trends in HIV co-infection

106,829 cases of TB in adults (≥15 years) were reported to ETS in England, Wales and Northern Ireland for the period 2000-2014. Overall, 5,792 people (5.4%) were identified as being co-infected with HIV through record linkage. The proportion of TB patients co-infected with HIV peaked in 2004 (543/6,782, 8.0%, Table 5.1), although the absolute number was higher in 2005 (555/7,489). This decreased to 205/6,461 (3.2%) in 2014 (Figure 5.2).

Figure 5.2: The number and proportion of notified TB cases with and without HIV in England, Wales and Northern Ireland by year of TB diagnosis.

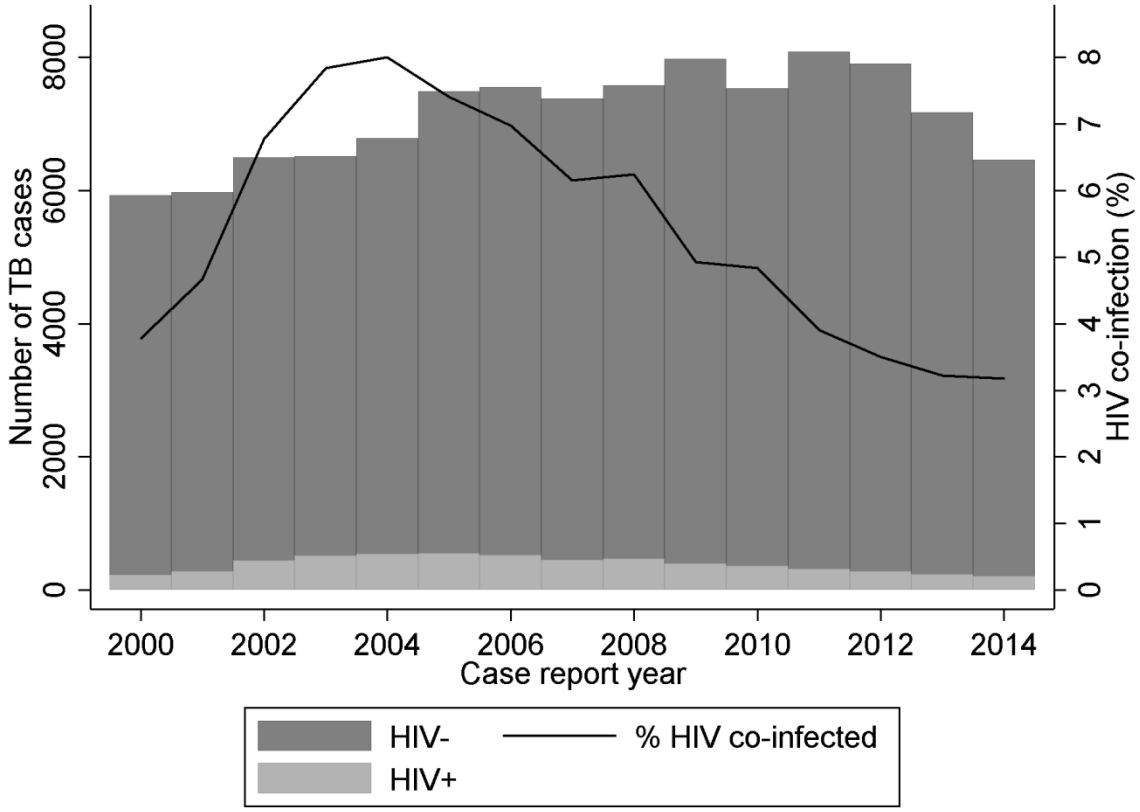


Table 5.1: The number and percentage of notified TB cases with and without HIV co-infection in England, Wales and Northern Ireland, 2000-2014.

	Number of TB cases (whole cohort)		
	HIV-negative	**HIV-positive (%)**	**Total**
Total	101,037	5,792 (5.4%)	106,829
Year			
2000	5,701	224 (3.8%)	5,925
2001	5,697	279 (4.7%)	5,976
2002	6,057	440 (6.8%)	6,497
2003	6,003	511 (7.8%)	6,514
2004	6,239	543 (8.0%)	6,782
2005	6,934	555 (7.4%)	7,489
2006	7,027	527 (7.0%)	7,554
2007	6,926	454 (6.2%)	7,380
2008	7,102	473 (6.2%)	7,575
2009	7,582	393 (4.9%)	7,975
2010	7,168	364 (4.8%)	7,532
2011	7,769	316 (3.9%)	8,085
2012	7,631	277 (3.5%)	7,908
2013	6,945	231 (3.2%)	7,176
2014	6,256	205 (3.2%)	6,461
Sex			
Female	43,761	2,935 (6.3%)	46,696
Male	57,085	2,846 (4.7%)	59,931
Missing	191	11 (5.4%)	202
Age group (years)			
15-24	16,628	320 (1.9%)	16,948
25-34	28,658	1,969 (6.4%)	30,627
35-44	17,521	2,281 (11.5%)	19,802
45-54	12,347	872 (6.6%)	13,219
55-64	9,283	255 (2.7%)	9,538
65+	16,576	94 (0.6%)	16,670
Missing	24	1 (4.0%)	25
Ethnicity/Country of birth			
White, low HIV prevalence	18,470	477 (2.5%)	18,947
Black African, low HIV prevalence	7,672	249 (3.1%)	7,921
Indian sub-continent, low HIV prevalence	38,745	179 (0.5%)	38,924
Other/unknown, low HIV prevalence	8,692	197 (2.2%)	8,889
White, high HIV prevalence	232	20 (7.9%)	252
Black African, high HIV prevalence	7,419	3,537 (32.3%)	10,956
Indian sub-continent, high HIV prevalence	1,325	29 (2.1%)	1,354
Other/unknown, high HIV prevalence	1,673	254 (13.2%)	1,927
Country of birth unknown	16,809	850 (4.8%)	17,659

	Number of TB cases (whole cohort)		
	HIV-negative	HIV-positive (%)	Total
Site of TB disease			
Pulmonary, +/- extra-pulmonary*	52,770	3,227 (5.8%)	55,997
Miliary/meningeal TB	3,747	780 (17.2%)	4,527
Extra-pulmonary only	44,218	1,761 (3.8%)	45,979
Missing	302	24 (7.4%)	326
Homelessness†			
No	32,129	1,136 (3.4%)	33,265
Yes	947	92 (8.9%)	1,039
Missing	2,693	165 (5.8%)	2,858
Imprisonment†			
No	31,092	1,102 (3.4%)	32,194
Yes	958	61 (6.0%)	1,019
Missing	3,719	230 (5.8%)	3,949
Drug misuse†			
No	31,823	1,125 (3.4%)	32,948
Yes	980	86 (8.1%)	1,066
Missing	2,966	182 (5.8%)	3,148
Alcohol misuse†			
No	31,297	1,098 (3.4%)	32,395
Yes	1,134	55 (4.6%)	1,189
Missing	3,338	240 (6.7%)	3,578
IMD decile†			
1	7,547	350 (4.4%)	7,897
2	6,836	287 (4.0%)	7,123
3	5,531	218 (3.8%)	5,749
4	4,207	134 (3.1%)	4,341
5	3,050	112 (3.5%)	3,162
6	2,294	89 (3.7%)	2,383
7	1,669	61 (3.5%)	1,730
8	1,408	48 (3.3%)	1,456
9	1,228	45 (3.5%)	1,273
10	994	19 (1.9%)	1,013
Missing	1,005	30 (2.9%)	1,035

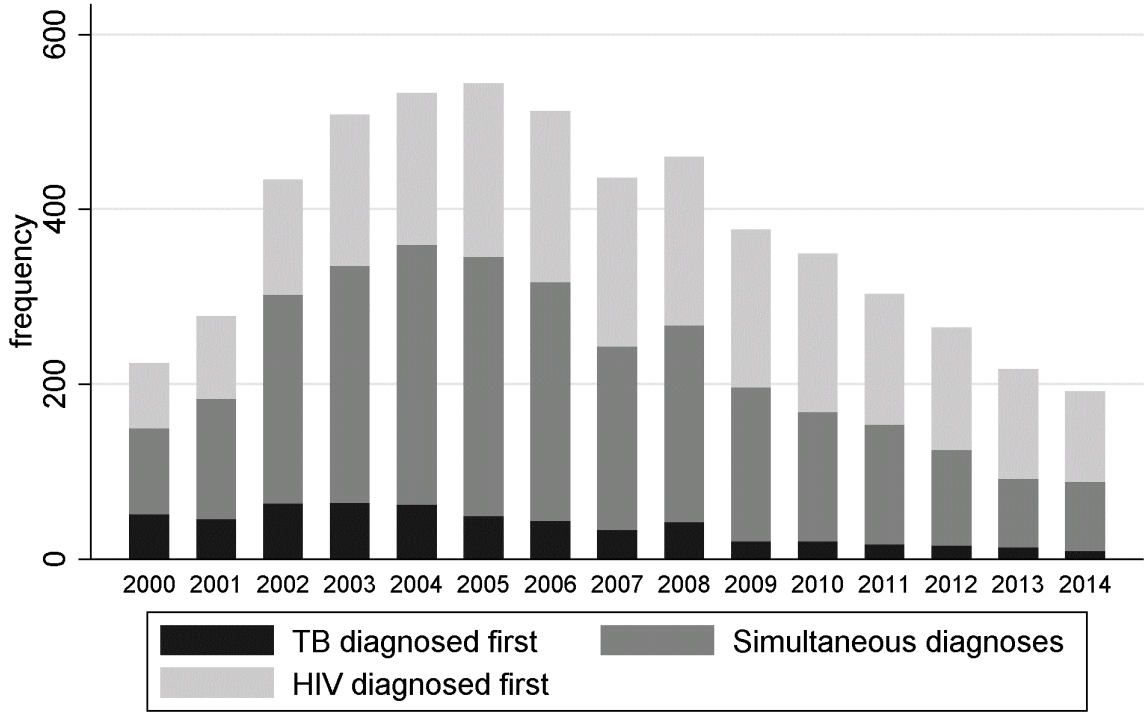
*** Excluding miliary and meningeal tuberculosis. †2010-2014 only. IMD: index of multiple deprivation, TB: tuberculosis.**

5.3.1.1 Timing of tuberculosis and HIV diagnoses

Overall, 2,787/5,792 (49%) co-infected people were diagnosed with TB and HIV simultaneously, whilst 549 (9%) were diagnosed with TB before and 2,456 (42%) after HIV diagnosis. The relationship between TB and HIV diagnosis changed over time (Figure 5.3). The absolute number of TB cases in people with known HIV rose from 75/224 (33.5%) in 2000 to 210/454 (46.2%) in 2007, then fell to 117/205 (57.1%) by

2014, but increased as a proportion of all TB-HIV cases from 2000 to 2014 (Table 5.2), whilst the number of cases diagnosed with TB first fell from 22.8% to 4.4%. The number of simultaneous diagnoses peaked at 54.9% in 2004, but has since decreased to 38.5%. There were corresponding declines in the proportion diagnosed simultaneously with TB and HIV from 298/543 (55%) in 2004 to 79/205 (39%) in 2014 and those first diagnosed with TB (51/224 (23%) in 2000 to 9/205 (4%) in 2014). The trend in decreasing proportions of simultaneous diagnoses and increasing proportions of TB being diagnosed after HIV was consistent in a sensitivity analysis using a 30 day threshold rather than 91 days.

Figure 5.3: The first diagnosis for TB-HIV co-infected patients diagnosed with TB from 2000 to 2014.



TB: tuberculosis. Simultaneous diagnoses were defined as TB and HIV diagnosed within 91 days of each other.

Table 5.2: The relationship between HIV and TB diagnoses in patients diagnosed with HIV and TB between 2000 and 2014.

Case report year	Simultaneous diagnoses		HIV diagnosed first		TB diagnosed first		Total
	n	%	n	%	n	%	
2000	98	43.8	75	33.5	51	22.8	224
2001	138	49.5	96	34.4	45	16.1	279
2002	241	54.8	135	30.7	64	14.5	440
2003	271	53.0	175	34.2	65	12.7	511
2004	298	54.9	183	33.7	62	11.4	543
2005	296	53.3	209	37.7	50	9.0	555
2006	274	52.0	209	39.7	44	8.3	527
2007	211	46.5	210	46.3	33	7.3	454
2008	227	48.0	204	43.1	42	8.9	473
2009	179	45.5	194	49.4	20	5.1	393
2010	148	40.7	196	53.8	20	5.5	364
2011	137	43.4	163	51.6	16	5.1	316
2012	111	40.1	151	54.5	15	5.4	277
2013	79	34.2	139	60.2	13	5.6	231
2014	79	38.5	117	57.1	9	4.4	205
Total	2,787		2,456		549		5,792

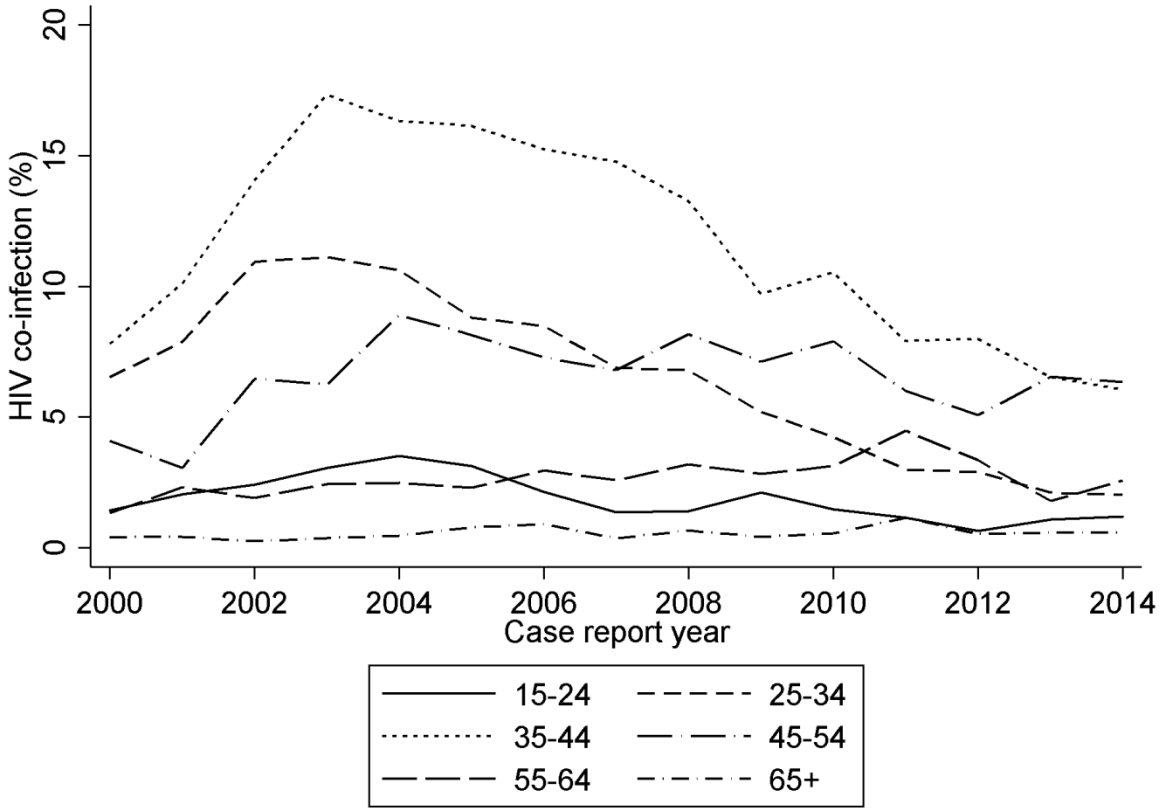
n: number, TB: tuberculosis. Simultaneous diagnosis was defined as TB and HIV diagnosed within 91 days of each other.

5.3.1.2 Trends in the demographics of co-infected patients

During the study period, HIV co-infection was higher in women (2,935/46,696, 6.3%) than men (2,846/59,931, 4.7%); however the decline was larger in women (Figure 5.4). There were also declines in HIV co-infection in TB patients aged 15-44, but not in patients aged 45 or older (Figure 5.5).

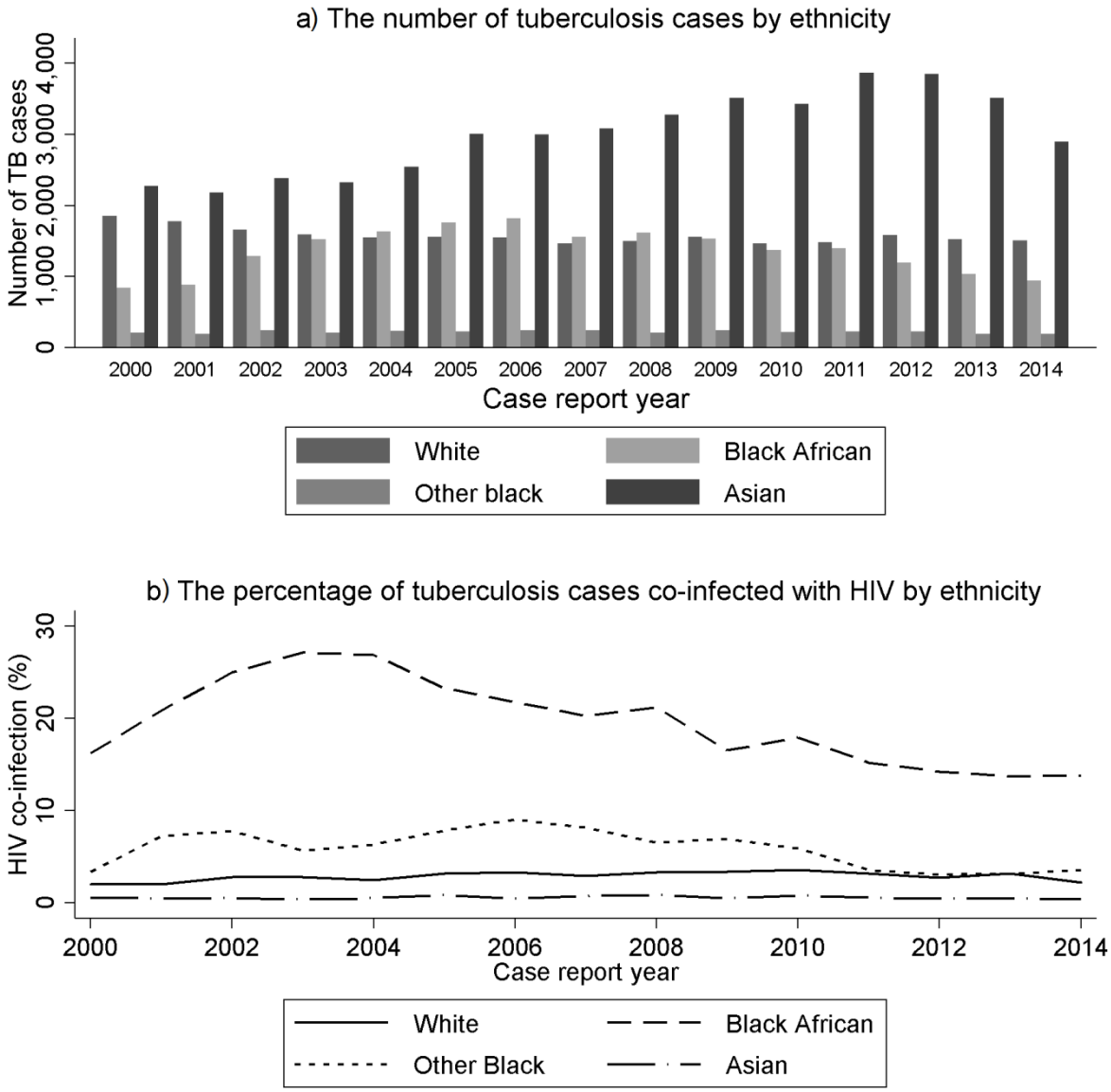
Figure 5.4: The percentage of notified TB cases co-infected with HIV by sex in England, Wales and Northern Ireland, 2000 to 2014.

Figure 5.5: The percentage of notified TB cases co-infected with HIV by age in England, Wales and Northern Ireland, 2000-2014.



The percentage of people with HIV co-infection was highest in TB patients of black African ethnicity born in countries with a high HIV prevalence (3,537/10,956, 32.3%), compared to 477/18,947 (2.5%) in the white UK-born population. Temporal changes in the ethnicity of TB patients (Figure 5.6) correlate with overall trends in the number and percentage of people with TB and HIV; TB in people of black African ethnicity peaked in 2006 and has since been decreasing, whereas the number of Asian TB patients was increasing until 2011.

Figure 5.6: The number (a) and percentage (b) of notified TB cases co-infected with HIV by ethnicity in England, Wales and Northern Ireland, 2000 to 2014.



5.3.2 Factors associated with HIV co-infection

5.3.2.1 Univariable results

In univariable analyses, the odds ratio (OR) for HIV co-infection was higher for the years 2001-2010 compared to 2000 (peak OR 2.22 [1.89-2.60]) in 2004) (Table 5.3). The ORs for 2011-2014 were not significantly different to 2000. Men were less likely to be co-infected with HIV than women (OR 0.74 [0.70-0.78]). Compared to TB patients aged 25-34 years, HIV co-infection was more likely for patients aged 35-44 (OR 1.89 [1.78-2.02]) but less likely for patients aged 15-24 (OR 0.28 [0.25-0.32]), 55-64 (OR 0.40 [0.35-0.46]) or over 65 years (OR 0.08 [0.07-0.10]).

Compared to TB patients of white ethnicity born in countries with low HIV prevalence, the odds of HIV were highest for patients of black African ethnicity born in countries with high HIV prevalence (OR 18.46 [16.71-20.39]). The odds of HIV were also higher among black African patients born in countries with low HIV prevalence (OR 1.26 [1.08-1.47]) and patients born in countries with high HIV prevalence who were of white (OR 3.34 [2.10-5.32]) or 'other'/unknown ethnicity (OR 5.88 [5.01-6.90]); but substantially lower among patients from countries with low HIV prevalence who were of Indian/Pakistani/Bangladeshi origin (OR 0.18 [0.15-0.21]).

HIV co-infection was greater among TB patients with miliary or meningeal TB compared to patients with pulmonary disease (with or without other extra-pulmonary disease) (OR 3.40 [3.13-3.71]) but lower for patients with other extra-pulmonary disease only (OR 0.65 [0.61-0.69]).

Table 5.3: Results from univariable and two multivariable logistic regression models of factors associated with HIV co-infection in notified TB cases in England, Wales and Northern Ireland, for the periods 2000-2014 and 2010-2014.

		Univariable results (whole cohort)		Multivariable results (whole cohort)		Multivariable results (2010-2014)	
		OR (95% CI)	P value	OR (95% CI)	P value	OR (95% CI)	P value
Year							
	2000	1.00		1.00		-	
	2001	1.25 (1.04-1.49)		1.22 (1.00-1.48)		-	
	2002	1.85 (1.57-2.18)		1.45 (1.20-1.74)		-	
	2003	2.17 (1.84-2.55)		1.51 (1.26-1.80)		-	
	2004	2.22 (1.89-2.60)		1.56 (1.31-1.87)		-	
	2005	2.04 (1.74-2.39)		1.40 (1.17-1.67)		-	
	2006	1.91 (1.63-2.24)		1.36 (1.14-1.62)		-	
	2007	1.67 (1.42-1.97)	<0.001	1.22 (1.01-1.46)	<0.001	-	
	2008	1.70 (1.44-1.99)		1.25 (1.04-1.50)		-	
	2009	1.32 (1.12-1.56)		1.04 (0.86-1.25)		-	
	2010	1.29 (1.09-1.53)		1.06 (0.87-1.28)		1.00	
	2011	1.04 (0.87-1.23)		0.83 (0.68-1.01)		0.76 (0.61-0.93)	
	2012	0.92 (0.77-1.11)		0.82 (0.67-1.00)		0.81 (0.65-0.99)	<0.001
	2013	0.85 (0.70-1.02)		0.76 (0.62-0.94)		0.66 (0.53-0.83)	
	2014	0.83 (0.69-1.01)		0.72 (0.58-0.89)		0.68 (0.54-0.85)	
Sex							
	Female	1.00	<0.001	1.00	<0.001	1.00	0.03
	Male	0.74 (0.70-0.78)		0.84 (0.79-0.90)		0.85 (0.74-0.98)	
Age group (years)							
	15-24	0.28 (0.25-0.32)	<0.001	0.28 (0.24-0.32)	<0.001	0.31 (0.22-0.44)	<0.001
	25-34	1.00		1.00		1.00	

	Univariable results (whole cohort)		Multivariable results (whole cohort)		Multivariable results (2010-2014)	
	OR (95% CI)	P value	OR (95% CI)	P value	OR (95% CI)	P value
35-44	1.89 (1.78-2.02)		1.97 (1.83-2.13)		2.16 (1.79-2.59)	
45-54	1.03 (0.95-1.12)		1.33 (1.21-1.47)		1.97 (1.60-2.44)	
55-64	0.40 (0.35-0.46)		0.62 (0.54-0.72)		1.35 (1.02-1.78)	
65+	0.08 (0.07-0.10)		0.16 (0.13-0.19)		0.31 (0.20-0.49)	
Ethnicity/Country of birth						
White, low HIV prevalence	1.00		1.00		1.00	
Black African, low HIV prevalence	1.26 (1.08-1.47)		1.07 (0.91-1.26)		1.07 (0.74-1.55)	
Indian sub-continent, low HIV prevalence	0.18 (0.15-0.21)		0.16 (0.14-0.20)		0.20 (0.14-0.28)	
Other/unknown, low HIV prevalence	0.88 (0.74-1.04)	<0.001	0.73 (0.62-0.87)	<0.001	0.67 (0.46-0.96)	<0.001
White, high HIV prevalence	3.34 (2.10-5.32)		2.43 (1.51-3.91)		3.85 (1.45-10.23)	
Black African, high HIV prevalence	18.46 (16.71-20.39)		13.02 (11.69-14.50)		10.39 (8.13-13.29)	
Indian sub-continent, high HIV prevalence	0.85 (0.58-1.24)		0.69 (0.47-1.01)		0.51 (0.20-1.26)	
Other/unknown, high HIV prevalence	5.88 (5.01-6.90)		4.56 (3.85-5.39)		4.07 (2.79-5.94)	
Country of birth unknown	1.96 (1.75-2.19)		1.62 (1.44-1.82)		1.51 (1.14-2.00)	
Site of TB disease						
Pulmonary, +/- extra-pulmonary*	1.00		1.00		1.00	
Miliary/meningeal TB	3.40 (3.13-3.71)	<0.001	3.30 (2.96-3.68)	<0.001	3.81 (3.02-4.80)	<0.001
Extra-pulmonary only	0.65 (0.61-0.69)		0.70 (0.65-0.75)		0.73 (0.63-0.86)	
Homelessness						
No	1.00	<0.001	-		1.00	0.26
Yes	2.75 (2.20-3.43)		-		1.22 (0.87-1.72)	

		Univariable results (whole cohort)		Multivariable results (whole cohort)		Multivariable results (2010-2014)	
		OR (95% CI)	P value	OR (95% CI)	P value	OR (95% CI)	P value
Imprisonment							
	No	1.00	<0.001	-		1.00	0.19
	Yes	1.80 (1.38-2.34)		-		0.77 (0.52-1.15)	
Drug misuse							
	No	0.00	<0.001	-		1.00	<0.001
	Yes	2.48 (1.98-3.12)		-		2.70 (1.90-3.84)	
Alcohol misuse							
	No	1.00	0.03	-		1.00	0.73
	Yes	1.38 (1.05-1.82)		-		0.93 (0.64-1.37)	
IMD decile							
	(for each unit increase)	0.95 (0.93-0.97)	<0.001	-		0.99 (0.96-1.02)	0.47

*** Excluding miliary/meningeal tuberculosis. CI: confidence interval, IMD: index of multiple deprivation, OR: odds ratio, TB: tuberculosis. The whole-cohort model excluded 543 (0.5%) TB cases missing data on sex (n=202), age (n=25) and/or site of TB disease (n=326), and was adjusted for year of TB notification, age, sex, ethnicity and HIV prevalence in country of birth and site of TB disease. The 2010-2014 model excluded 7,187 (19%) TB cases missing data on sex (n=60), site of TB disease (n=137), homelessness (n=2,693), imprisonment (n=3,719), drug misuse (n=2,966), alcohol misuse (n=3,338) and/or IMD decile (n=1,035) and adjusted for all variables in the table.**

5.3.2.2 Multivariable results

Factors associated with HIV co-infection were consistent between the univariable and multivariable results (Table 5.3). Year and age were included as categorical variables (tests for linearity both P<0.001).

The odds of HIV co-infection were higher from 2001 to 2008 (peak OR 1.56 [95% CI 1.31-1.87] in 2004) and lower in 2013 and 2014 (OR 0.72 [0.58-0.89] in 2014) than in 2000. The odds of HIV co-infection were lower for men than women (OR 0.84 [0.79-0.90]). Compared to white patients born in countries with low HIV prevalence, the odds of HIV co-infection were higher for people of black African ethnicity who were born in countries with high HIV prevalence (OR 13.02 [11.69-14.50]).

People with miliary/meningeal TB had increased odds of HIV co-infection (OR 3.30 [2.96-3.68]) versus people with pulmonary TB, whilst people with other extra-pulmonary disease only were less likely to have HIV (OR 0.70 [0.65-0.75]).

Sensitivity analyses investigating the impact of excluding weaker matches between ETS and HARS showed results consistent with the main models (Table 5.4).

Table 5.4: Sensitivity analyses for two multivariable logistic regression models of factors associated with HIV co-infection in notified TB cases in England, Wales and Northern Ireland, in 2000-2014 and 2010-2014, which excluded co-infected patients who were matched with low probabilistic scores or the weakest deterministic criteria.

	Whole cohort*		2010-2014≠	
	OR (95% CI)	P value	OR (95% CI)	P value
Year				
2000	1.00		-	
2001	1.21 (0.98-1.49)		-	
2002	1.49 (1.23-1.81)		-	
2003	1.54 (1.27-1.86)		-	
2004	1.61 (1.33-1.94)		-	
2005	1.43 (1.19-1.72)		-	
2006	1.42 (1.18-1.71)		-	
2007	1.27 (1.05-1.54)	<0.001	-	
2008	1.29 (1.07-1.56)		-	
2009	1.10 (0.90-1.33)		-	
2010	1.12 (0.92-1.37)		1.00	0.00
2011	0.88 (0.72-1.08)		0.75 (0.60-0.94)	
2012	0.86 (0.69-1.05)		0.80 (0.64-0.99)	
2013	0.78 (0.63-0.97)		0.62 (0.49-0.79)	
2014	0.76 (0.61-0.96)		0.68 (0.53-0.86)	
Sex				
Female	1.00	<0.001	1.00	0.03
Male	0.82 (0.77-0.88)		0.85 (0.73-0.98)	
Age group (years)				
15-24	0.27 (0.24-0.31)			
25-34	1.00		0.31 (0.22-0.45)	<0.001
35-44	1.99 (1.84-2.14)	<0.001	1.00	
45-54	1.33 (1.20-1.47)		2.23 (1.84-2.71)	
55-64	0.62 (0.53-0.72)		2.06 (1.65-2.56)	
65+	0.15 (0.12-0.19)		1.34 (1.00-1.80)	
Ethnicity/Country of birth			0.26 (0.16-0.43)	
White, UK-born	1.00			
Black African, low HIV prevalence	0.99 (0.83-1.17)			
Indian sub-continent, low HIV prevalence	0.13 (0.11-0.16)		1.00	<0.001
Other/unknown, low HIV prevalence	0.64 (0.54-0.78)		1.03 (0.69-1.52)	
White, high HIV prevalence	2.43 (1.49-3.94)	<0.001	0.17 (0.12-0.25)	
Black African, high HIV prevalence	12.54 (11.23-14.00)		0.57 (0.39-0.85)	
Indian sub-continent, high HIV prevalence	0.70 (0.47-1.04)		4.18 (1.57-11.14)	
Other/unknown, high HIV prevalence	4.42 (3.72-5.26)		10.23 (7.92-13.21)	
Country of birth unknown	1.55 (1.37-1.75)		0.55 (0.22-1.37)	

	Whole cohort*		2010-2014≠	
	OR (95% CI)	P value	OR (95% CI)	P value
Site of TB disease			4.23 (2.87-6.24)	
Pulmonary, +/- extra-pulmonary*	1.00		1.46 (1.09-1.97)	
Miliary/meningeal TB	3.34 (2.99-3.74)	<0.001		
Extra-pulmonary only	0.70 (0.65-0.75)		1.00	<0.001
Homelessness			3.84 (3.02-4.89)	
No	-		0.72 (0.61-0.85)	
Yes	-			
Imprisonment				
No	-		1.00	0.60
Yes	-		1.11 (0.77-1.60)	
Drug misuse				
No	-		1.00	<0.001
Yes	-		2.61 (1.80-3.78)	
Alcohol misuse				
No	-		1.00	0.52
Yes	-		0.88 (0.58-1.32)	
IMD decile				
(for each unit increase)	-		0.98 (0.95-1.02)	0.33

IMD: index of multiple deprivation, OR: odds ratio, TB: tuberculosis. Sensitivity analyses were conducted for both the whole-cohort model and the 2010-2014 model; excluding HIV-positive individuals with the lowest 5% of probabilistic matching scores (linking the TB and HIV records) and individuals whose records were matched using the three weakest deterministic criteria.

*** The whole-cohort sensitivity analysis excluded 506 TB cases co-infected with HIV.**

≠ The 2010-2014 sensitivity analysis excluded 45 TB cases co-infected with HIV.

5.3.3 Social risk factors associated with HIV co-infection

From 2010-2014, 37,162 TB cases were notified, 1,393 (3.8%) were co-infected with HIV. Complete data on social risk factors (drug and alcohol misuse, homelessness, imprisonment and IMD decile) were available for 30,105 patients (81%). The proportion of patients with HIV (Table 5.1) was higher in people with current or previous drug misuse (8.1%), homelessness (8.9%), imprisonment (6.0%) or alcohol misuse (4.6%) than for patients with no social risk factors (3.2%, Table 5.5).

Table 5.5 shows the percentage of TB patients with combinations of different risk factors, and the proportion with HIV. There was substantial correlation between different social risk factors among TB patients. HIV co-infection was highest among TB patients with a history of drug misuse and homelessness (11.5% co-infected). HIV prevalence among TB patients with no history of drug misuse, homelessness or imprisonment was 3.2%.

In univariable analyses HIV co-infection was positively associated with homelessness (OR 2.75 [2.20-3.43]), imprisonment (OR 1.80 [1.38-2.34]), drug misuse (OR 2.48 [1.98-3.12]) and alcohol misuse (OR 1.38 [1.05-1.82]). The associations for other explanatory variables were consistent with the whole-cohort model. IMD decile was retained as a linear variable (test for linearity P=0.14). In a multivariable model adjusted for all variables shown in Table 5.1, the only social risk factor with strong statistical evidence for an association with HIV co-infection was drug misuse (OR 2.70 [1.90-3.84], P<0.001). There was no evidence for statistical interactions between the social risk factors in the regression model (P>0.25 for all combinations).

There was no substantial difference in the year of notification, age, or site of disease for patients with completed social risk factor data and patients missing data on one or more social risk factors. However, patients with missing social risk factor data were more likely to be women, and/or from the Indian subcontinent.

Table 5.5: The number of notified TB cases with and without a history of drug misuse, homelessness or imprisonment, and the HIV prevalence in each of these groups in England, Wales and Northern Ireland, 2010-14.

		Drug misuse		Total
		No	Yes	
Total (homeless and not homeless)	**No Prison**	30,711 (3.2%[3.1-3.4])	506 (8.3%[5.9-10.7])	31,217 (3.3%[3.1-3.5])
n (% with HIV [95% CI])	**Prison**	480 (4.4%[2.5-6.2])	365 (6.8%[4.3-9.4])	845 (5.4%[3.9-7.0])
	Total	31,191 (3.3%[3.1-3.5])	871 (7.7%[5.9-9.5])	32,062 (3.4%[3.2-3.6])
Homeless	**No Prison**	465 (7.1%[4.8-9.4])	86 (14.0%[6.6-21.3])	551 (8.2%[5.9-10.5])
n (% with HIV [95% CI])	**Prison**	99 (3.0%[-0.4-6.4])	149 (10.1%[5.2-14.9])	248 (7.3%[4.0-10.5])
	Total	564 (6.4%[4.4-8.4])	235 (11.5%[7.4-15.6])	799 (7.9%[6.0-9.8])
Not homeless	**No Prison**	30,246 (3.2%[3.0-3.4])	420 (7.1%[4.7-9.6])	30,666 (3.2%[3.0-3.4])
n (% with HIV [95% CI])	**Prison**	381 (4.7%[2.6-6.9])	216 (4.6%[1.8-7.4])	597 (4.7%[3.0-6.4])
	Total	30,627 (3.2%[3.0-3.4])	636 (6.3%[4.4-8.2])	31,263 (3.3%[3.1-3.5])

CI: confidence interval, n: number, TB: tuberculosis. Excluding 5,899 (16%) cases missing data on homelessness (n=2,693), imprisonment (n=3,719) and/or drug misuse (n=2,966).

5.4 Discussion

5.4.1 Summary of findings

In England, Wales and Northern Ireland, there was a substantial decline in both the number and proportion of notified TB cases with HIV since 2005 and 2004 respectively, particularly in women. Within the decline in the number of co-infected TB patients, the proportion who were aware of their HIV infection at the time of their TB diagnosis rose during the study period, whilst the proportion who were diagnosed with TB before their HIV diagnosis decreased. HIV co-infection was greatest for people of black African ethnicity born in countries with a high HIV prevalence. Drug misuse was the only social risk factor independently associated with HIV co-infection.

5.4.2 Strengths and limitations of the study

This study benefitted from fifteen years of case notifications, representing comprehensive coverage of TB cases in England, Wales and Northern Ireland. As with the results presented in Chapter 4, the large sample size meant the study had plenty of power to detect associations. The narrow confidence intervals and low P values (predominantly $P<0.001$, Table 5.3) for the variables associated with lower risk of HIV co-infection indicate that these associations are unlikely to be due to chance.

There were some limitations. The outcome of this study was HIV co-infection; however HIV is a causal factor in the development of TB. The associations reported here are therefore not necessarily causal relationships (particularly the associations between site of disease and the likelihood of HIV co-infection) but should be interpreted as indicators of which TB patients may be more likely to also have HIV. The epidemiological implications of the limitations of the matching algorithm used to identify co-infected patients were described in section 4.4.2, but additionally there were 695 patients reported to HARS as having TB as an AIDS-defining illness which were not matched to notified TB cases (Chapter 3). As notification of TB cases to PHE is mandatory, it is likely that these cases were not matched due to a combination of

missing data and common names or other PII, rather than the TB cases not having been reported to ETS; although it is possible that extra-pulmonary TB cases managed outside of respiratory departments, and TB cases treated in the private sector, are under-reported to ETS. Treating these patients as HIV-negative in this analysis may have caused underestimation of the prevalence of HIV co-infection and may have biased the odds ratios. Sensitivity analyses were conducted, excluding weaker matches between TB and HIV records to assess whether associations were different for cases whose HIV status was less certain. These provided consistent results, suggesting the matching algorithm did not affect the conclusions drawn.

Most variables had little missing data, however 19% of patients from 2010-2014 were missing some social risk factor data. These patients were more likely to be women or from the Indian sub-continent; both groups have low levels of social risk factors.[52] Consequently, any bias in the results is likely to be towards the null, underestimating the association between TB and social risk factors. The use of routinely collected surveillance data meant that all diagnoses of TB and HIV were made according to standard definitions for both infections, thereby minimising measurement error and misclassification. Any TB diagnoses later found not to be *MTBC* were de-notified and excluded from the analysis.

Similarly to Chapter 4, a composite variable was created for ethnicity and country of birth due to known interactions between some strata of these variables; however, time since migration to the UK (for foreign-born patients) was not adjusted for. Recent migration has previously been associated with HIV co-infection in foreign-born TB cases,[123] although the proportion of TB cases in recent migrants has decreased over the last decade since this association was last reported.[52] Instead, HIV prevalence in country of birth was adjusted for, as a proxy for exposure to HIV which is a causal factor for HIV co-infection. Similarly, it was not possible to adjust for other factors influencing exposure to HIV such as having a sexual partner with HIV, or time spent in countries where HIV prevalence is much higher.

Data on social risk factors (drug and alcohol misuse, homelessness and imprisonment) were included in the multivariable model, and the prevalence of HIV co-infection in TB patients with different combinations of these risk factors was examined. Whilst there was substantial overlap between risk factors (a third of cases with social risk factors had more than one social risk factor [52]), there was no evidence for statistical interactions between these variables in the multivariable model. However, the estimated odds ratios may have been confounded by other social risk factors for which data was not available; in particular, it was not possible to differentiate between injecting drug use (an established risk factor for HIV infection) and other drug misuse, and it is possible that there are differential associations between these and other social risk factors. Higher deciles of IMD scores were associated with HIV co-infection in univariable analyses but not in a multivariable model, and consequently as there was a high degree of missing data for this variable due to missing postcode data (from which it is derived) it was excluded from the final multivariable model. As IMD score is a measure of area-level deprivation, it is not necessarily an accurate measure of socioeconomic status for any individual patient, both factors which may influence the risk of TB.

It was only possible to identify co-infected individuals with diagnosed HIV, and an estimated 13% of HIV infections were undiagnosed in 2015.[10] However, 94% of people diagnosed with TB in 2015 received HIV testing or were already aware of their HIV status,[52] and therefore a very low prevalence of undiagnosed HIV would be expected in this population. There may have been more undiagnosed HIV among TB patients prior to 2008, when BHIVA guidelines were released recommending HIV testing for all TB patients;[220] however, it is likely that any of these TB patients with undiagnosed HIV would have since presented to care and been diagnosed with HIV. As the matching algorithm linked all TB cases from 2000 to 2014 to all PLHIV between 2000 and 2014, the probability that co-infected patients were not identified is low. As this was a retrospective, observational study of TB cases, causality between HIV co-

infection and TB disease could not be established,[221] but it is likely that HIV infection precedes TB disease.

The analysis of trends in the timings of TB and HIV diagnoses of co-infected patients was limited by the need to define a threshold for defining 'simultaneous' diagnosis. A 91 day threshold for defining simultaneous diagnoses was a pragmatic choice to account for possible delays in diagnosis and reporting, and to exclude ART-induced unmasking immune reconstitution inflammatory syndrome. However, the sensitivity analysis using a 30 day threshold for a stricter definition of 'simultaneous diagnosis' showed a consistent trend, suggesting the results were not biased by the choice of threshold.

5.4.3 Implications of the research

BHIVA and NICE have recommended HIV testing for all patients diagnosed with TB since 2011.[218, 220] Between 2011 and 2015 over 90% of TB patients with previously unknown HIV status were tested for HIV.[52] The proportion of TB patients diagnosed simultaneously with TB and HIV started declining before the introduction of these guidelines, and has since decreased further, however a substantial proportion of patients (38.5%) were still diagnosed simultaneously in 2014. The routine testing of new TB patients for HIV provides no opportunity to prevent TB disease, and TB in patients who were simultaneously diagnosed cases may have been preventable if HIV had been diagnosed earlier and ART had been initiated sooner. In 2008, BHIVA introduced specific guidelines recommending more HIV testing in populations at high risk of infection, and increased the CD4 count threshold at which they recommend PLHIV start ART. This is reflected in the rise in median CD4 count representing a healthier population of PLHIV,[222] contributing to the observed decline in co-infections and the decreasing number and proportion of simultaneous diagnoses. However, more community HIV testing in populations at risk for TB is necessary to diagnose HIV sooner, to enable previously undiagnosed PLHIV to begin ART and prevent progression of LTBI to active disease. The recent change in BHIVA guidelines to

recommend ART for all PLHIV regardless of their CD4 count should also further decrease TB incidence in PLHIV. As a substantial proportion of co-infected patients are still diagnosed simultaneously, this study demonstrates the continued need for universal HIV testing for all patients diagnosed with TB, and highlights the groups that should be prioritised for interventions that improve the uptake of HIV testing.

Changes in the pattern of co-infections mean that a greater proportion of TB-HIV is now occurring in people with known HIV infection. This may be explained by increased HIV testing,[223] resulting in earlier HIV diagnosis. However, as the number of people with HIV in the UK increases, sustained success requires better management of latent TB infection to prevent the occurrence of TB disease in people diagnosed with HIV. The increasing proportion of TB cases in people with diagnosed HIV represents a missed opportunity to prevent these TB cases by diagnosing and treating latent and incipient TB. This study demonstrates the continued risk of TB in people diagnosed with HIV, highlighting the need for more testing and treatment of LTBI in high-risk PLHIV.

Part of the decline in HIV co-infection can be attributed to changing migration patterns, resulting in less TB in black African individuals from countries with a high HIV prevalence and more among individuals from Asian countries with low HIV prevalence.[82] The prevalence of TB in most African countries also declined during the study period,[224] as did the number of new HIV infections in Africa between 2010 and 2015,[225] which may have contributed to the decreasing number of TB and HIV infections among migrants. However, whilst the largest decreases in HIV co-infection were in black African TB patients, they remain the most at-risk, and the BHIVA guidelines recommending LTBI screening in this population of PLHIV remain appropriate.

HIV co-infection was greater among women with TB than men. This reflects the gender ratio of heterosexual PLHIV in the UK; where there are more women diagnosed with HIV than men, particularly among black Africans.[10] This is partly because of

higher diagnostic rates in women [10] due to antenatal HIV testing, and partly due to differences in sexual mixing; black African men tend to have more new partners than women, leading to a disproportionate number of new HIV infections in women. The decline of co-infection in women therefore reflects both the changes in ethnicity of TB patients and the healthier cohort of women with HIV. Co-infection declined in TB patients aged 15-44, but not in older people, which may be due to the ageing cohort of PLHIV and higher rates of late HIV diagnosis in older patients, increasing the risk of TB.[10]

HIV co-infection was higher in TB patients with any social risk factor (drug and alcohol misuse, imprisonment and homelessness) than in those without, and overall even higher among TB cases with multiple social risk factors. However, after adjusting for other social risk factors in the multivariable model, only drug misuse remained associated with greater odds of HIV co-infection; probably the result of HIV acquisition from injecting drug use. As TB incidence among PLHIV who inject drugs is comparable to that among black Africans from countries with high TB burden (Chapter 4), LTBI screening and treatment for PLHIV with social risk factors, particularly drug misuse, could decrease TB in this population.

There was a much greater proportion of HIV co-infection among patients with miliary and meningeal TB than pulmonary TB, and a lower proportion among patients with other extra-pulmonary disease. The results were consistent with other studies reporting higher prevalence of HIV co-infection among cases of miliary [226] and meningeal [227] TB, and lower HIV prevalence among other forms of extra-pulmonary TB. This was consistent over the study period and with previous work [215], probably because HIV contributes to a greater risk of severe disseminated TB disease, and because other forms of extra-pulmonary TB are common in TB patients from the Indian sub-continent [215] where HIV prevalence is low. Recent work in the UK showed patients with severe extra-pulmonary TB have worse outcomes than pulmonary TB, whilst patients with other extra-pulmonary TB generally have better outcomes, although

HIV was not adjusted for.[52] It is unclear whether this is solely because of severe disease presentation or whether they could be influenced by higher rates of HIV co-infection (and corresponding clinical complexity) in these patients. Regardless, as diagnosis of extra-pulmonary TB is often difficult, increasing awareness of extra-pulmonary TB symptoms among PLHIV and populations with high rates of undiagnosed HIV might allow earlier diagnosis of extra-pulmonary TB.

5.4.4 Conclusions

HIV co-infection in TB patients has substantially decreased over the past decade in England, Wales and Northern Ireland, as has the proportion of patients diagnosed simultaneously with both infections. However, sub-populations of patients with high rates of co-infection remain. Whilst the current policy of testing all patients diagnosed with TB for HIV infection is important in ensuring appropriate management of TB patients, many of these TB cases would be preventable if HIV could be diagnosed before TB develops. Increasing HIV testing and ensuring early treatment of HIV infection in black African populations (particularly people born in countries with high HIV prevalence) and people with a history of drug misuse could help prevent these TB cases. The BHIVA guidelines on LTBI testing for PLHIV from sub-Saharan Africa remain relevant, and LTBI screening for PLHIV with a history of drug misuse, homelessness or imprisonment should also be considered.

6 The role of HIV in tuberculosis transmission

6.1 Introduction

6.1.1 HIV and tuberculosis transmission

HIV infection increases susceptibility to TB disease by increasing the rate of progression from LTBI to active disease.[85, 88] However, there is also some evidence that patients with HIV co-infection may be less infectious; contact studies have shown lower prevalence of TST positivity and lower rates of TST conversion among contacts of HIV-positive index patients than HIV-negative index patients,[93-95, 228] particularly when the index patients with HIV were immunosuppressed.[96] This may be mediated through a shorter duration of infectiousness due to accelerated TB disease progression and earlier diagnosis,[88, 89] more rapid commencement of TB treatment,[96] lower rates of cavitary[94, 96] or sputum smear-positive[94, 95] TB, or a shorter duration of cough[94] among HIV-positive index patients.

Several studies in low-incidence settings which examined HIV as a risk factor for being part of a strain type cluster found no association,[71, 229, 230] including one meta-analysis.[231] One study in the USA reported a positive association between HIV and being part of a cluster for US-born patients in univariable analyses, but no association for foreign-born patients; and their stepwise approach to building a multivariable model resulted in HIV status being excluded from the multivariable model.[232] Two other studies reported that patients with HIV were less likely to be part of a strain type cluster; clusters were defined using 24-locus MIRU-VNTR in combination with WGS or spoligotyping.[233, 234] These studies all suffer from one major limitation, which is that the risk factors for being part of a strain type cluster could either be factors which affect susceptibility to infection or factors which influence infectiousness (or both). Consequently, if HIV co-infection lowers onwards transmission but increases susceptibility to infection, these effects could cancel each other out and no association would be observed. There is also weak evidence that HIV positivity

among the first cases of a cluster may be associated with increased numbers of secondary cases in clusters, suggesting that TB patients with HIV may be more infectious or mix with more susceptible contacts.[72, 235] Similarly, there is some evidence that cases arising from recent infection are more likely to be HIV-positive, which could be the result of greater susceptibility to infection or more rapid progression from LTBI to active TB disease.[236]

6.1.2 Clustering of tuberculosis cases in the UK

Molecular strain typing data can help identify cases which may be part of the same chain of transmission.[237] Since 2010, all culture-positive *MTBC* isolates in England, Wales and Northern Ireland have been prospectively strain typed using 24-locus mycobacterial interspersed repetitive units-variable number tandem repeats (MIRU-VNTR) strain typing. Between 2010 and 2015, 58.4% of TB cases in England were part of a strain type cluster with at least one other case.[52] Being part of a strain type cluster was more common for UK-born TB patients (69.9% clustered) than for patients born abroad (54.7%).

There have been a limited number of studies examining risk factors associated with clustering in the UK, most of which have not included data on HIV status. The last national study of clustering of TB cases in the UK used RFLP data from 1998 and reported that pulmonary disease, previous treatment for TB and homelessness were positively associated with being part of a strain type cluster, whilst TB patients from Asia and the Indian sub-continent were less likely to be clustered.[238] That study did not include data on HIV status, but a similar study of TB cases in London reported no association between HIV and being part of a strain type cluster.[70] In London, higher rates of clustering have been reported among UK-born patients, those with a history of drug or alcohol misuse or imprisonment, pulmonary TB (compared to extra-pulmonary TB) and those who have had TB previously.[239] Clusters where one of the first two cases had a history of imprisonment were more likely to become a large cluster (five or more cases) than where neither case had ever been imprisoned.[239]

6.1.3 Summary and objectives

Most risk factors for TB transmission have the same direction of effect on both susceptibility to infection and likelihood of onward transmission. In contrast, HIV is believed to increase susceptibility to infection and progression to active disease, but its impact on onward transmission is unclear. The aims of this chapter were to utilise molecular strain typing data, in combination with the national TB and HIV surveillance datasets, to examine for the first time whether there are associations between HIV co-infection and onward transmission of TB, and whether TB is more often due to reactivation of LTBI or recent infection in patients with and without HIV.

6.2 Methods

6.2.1 Study population

This was a retrospective study of culture-confirmed patients with *MTBC* in adults (aged ≥15 years) in England, Wales and Northern Ireland, notified to Public Health England (PHE)'s Enhanced TB Surveillance System (ETS) between 2010 and 2014. All patients with disease strain typed with at least 23 loci, using 24-loci MIRU-VNTR genotyping at PHE's mycobacteriology reference laboratories, were included.[237] Recurrence of TB was identified by record linkage within ETS. Recurrent cases of TB were excluded if the strain type of the second notification was the same as that of the first (i.e. plausible instances of relapse).

6.2.2 Outcome: strain type clusters

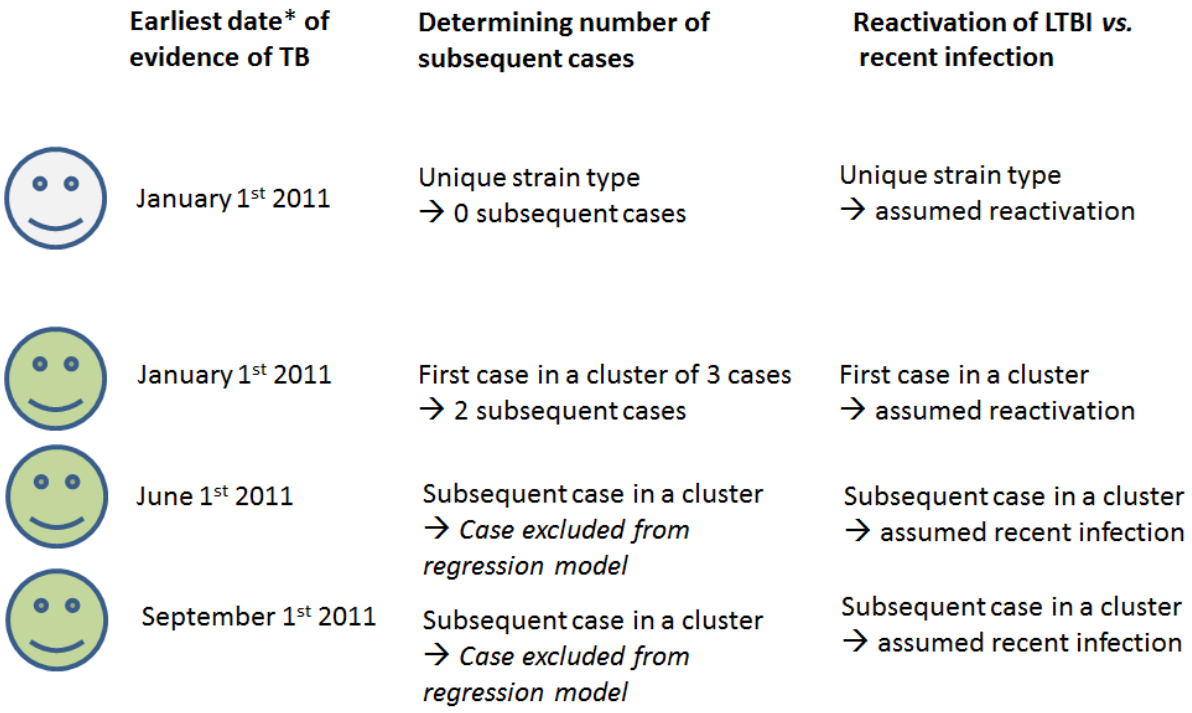
The outcome of interest was whether a TB patient was part of a strain type cluster. PHE defines a strain type cluster as two or more persons with TB caused by indistinguishable 24-loci MIRU-VNTR strain types.[237, 239] At least one case in the cluster must be typed at all 24 loci; additional cases can have one unmapped locus (which can differ between cases in the cluster). Some loci are prone to typing failures, despite repeated attempts, due to very high numbers of tandem repeats; in such instances all cases in a cluster are permitted to have an untypeable locus at the same

position.[237] TB cases whose strain type differed from all others at 1 or more loci were considered to have a unique strain type and therefore not part of a cluster. Where cases could belong to multiple clusters (due to untyped loci and thus missing information, or mixed infections), the cluster with the earliest sampling date was used, unless epidemiological data was available which supported the case belonging to a different cluster.

The earliest date of evidence of TB disease for each patient (including symptom onset date, date of presentation to healthcare, earliest specimen date, diagnosis date, treatment start date and case notification date) was used to define the order of cases within clusters (Figure 6.1). The first case in a cluster was defined as the patient with the earliest date and all later cases were considered subsequent cases. Patients who were not part of strain type clusters were classed as having zero subsequent cases. I examined whether cases were the first case in a cluster (as a surrogate for personal infectiousness and likelihood of onwards transmission) and if so, how many subsequent cases there were in a cluster. Whether or not a case was a subsequent case was also analysed, as a surrogate for recent infection compared to reactivation of previous LTBI.

Cases of TB in children (aged <15 years) were included in the dataset for defining strain type clusters and the size of clusters, but excluded from the subsequent analysis as record linkage to HIV surveillance to determine HIV status was only possible for adults. The adult TB cases which formed clusters with children were included in the analysis. Clusters were not limited by geographical area within England, Wales and Northern Ireland.

Figure 6.1: A diagram of how the order of cases within clusters, the number of subsequent cases, and whether a case was the result of reactivation or recent infection was determined.



Faces the same colour indicate a common strain type.
*** Dates made up as examples**

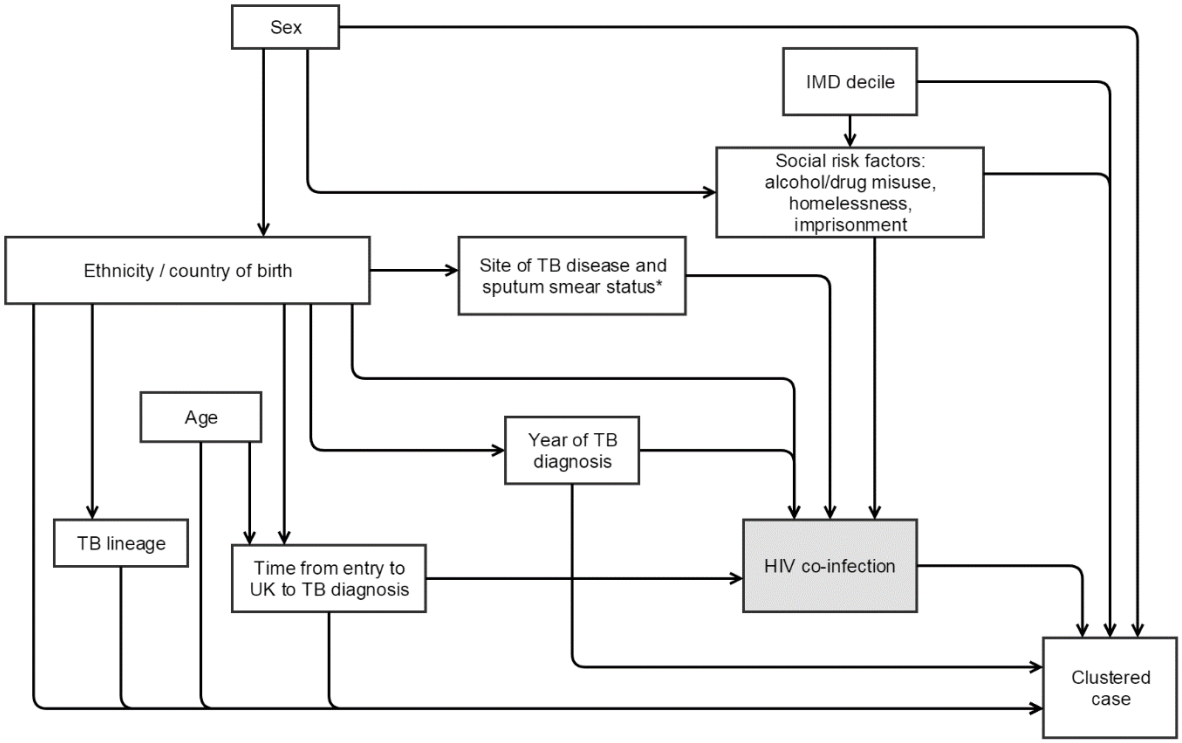
6.2.3 Exposure variables

The primary exposure variable was HIV status, which was determined through linkage to the national HIV and AIDS Reporting System (HARS),[98, 240] described in Chapter 3.

Potential confounders for the relationship between HIV status and the outcomes were identified prospectively in Chapter 4, and a causal framework was built (Figure 6.2).[199] Exposure variables for both outcomes included year of TB diagnosis, age at TB diagnosis, sex, ethnicity, place of birth and time from entry to the UK (if born abroad) to TB diagnosis, social risk factors (current or prior imprisonment, homelessness, drug misuse, or alcohol misuse hindering self-administration of TB treatment) and index of multiple deprivation [IMD] decile. When considering risk factors for being the first case in a cluster, the site of TB disease (smear positive/negative

pulmonary disease with or without extra-pulmonary disease, or extra-pulmonary disease only) was also considered.

Figure 6.2: Conceptual framework of potential factors associated with HIV co-infection and strain type clustering among TB cases.



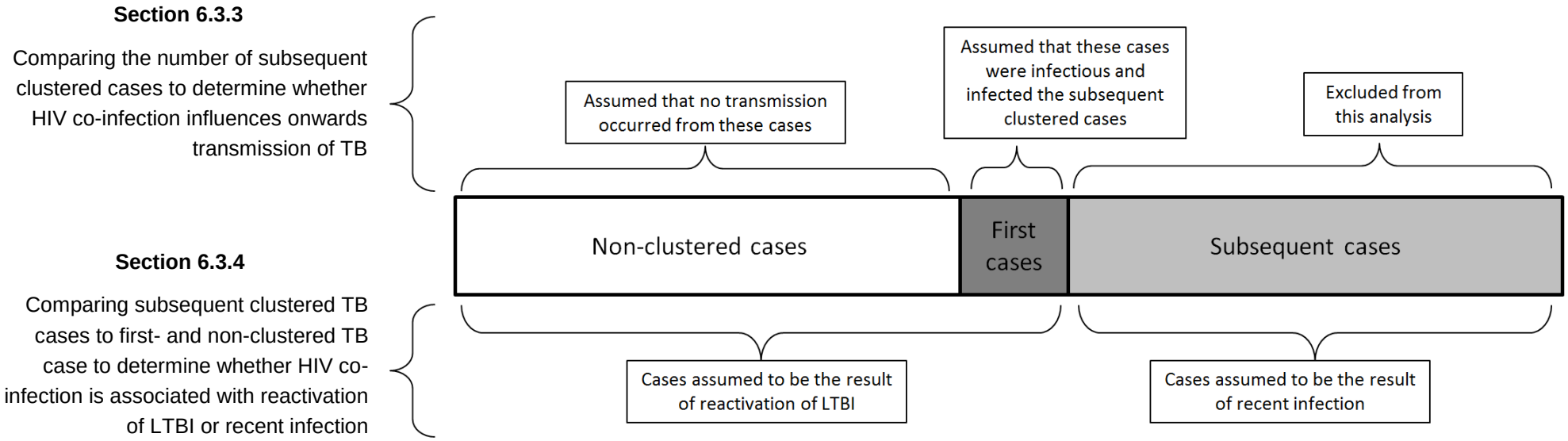
***Site of TB disease and sputum smear status were only considered causal factors onwards transmission of TB, not for comparing reactivation and recent infection.**
IMD: index of multiple deprivation, TB: tuberculosis.

6.2.4 Statistical Analysis

Data were analysed in Stata version 13.1. Descriptive analyses of the cohort were undertaken examining the proportion of cases belonging to a strain type cluster and how many of these cases were first cases compared to subsequent cases, stratified by HIV status. The number of subsequent cases following the first case of pulmonary TB in a cluster, stratified by HIV status of the first case in the cluster, was also examined.

To investigate whether HIV was a risk factor for potential transmission of TB, two analyses were conducted. Figure 6.3 shows which TB cases were included in each analysis.

Figure 6.3: Schematic diagram of which TB cases are included in which clustering analyses.



Firstly, to investigate the impact of HIV on the onward transmission of TB, zero-inflated Poisson regression [241] was used to examine factors associated with the number of subsequent cases after the first pulmonary case of each strain type cluster, as it is generally accepted that patients with only extra-pulmonary disease are not infectious. Poisson regression was considered, but the Poisson distribution was not a good fit for the data as there were a high number of zero counts (non-clustered TB patients, who had zero subsequent cases). Zero-inflated Poisson regression assumes there are two latent groups within the data; a group which can only have zero as its count (in this study, patients who are not infectious) and a group which can have a zero or non-zero count, with a Poisson distribution (the number of subsequent cases generated by a TB patient who is infectious). Zero-inflated Poisson models give an output in two parts; an incidence rate ratio for the number of subsequent cases (given that there has been transmission of infection), and an odds ratio for the likelihood of having zero subsequent cases by default. The Vuong test was used to determine whether a Poisson model or the zero-inflated Poisson model were a better fit for the data.

After calculating the number of subsequent cases for each TB case (zero for non-clustered cases, one or more for the first case of each cluster), subsequent cases in clusters were excluded from the regression model of factors associated with clustering, as it was not possible to determine the exact chain of transmission within clusters from MIRU-VNTR strain typing. For each case, the model was offset by the time since the earliest date of evidence of TB to the end of the study period (31st December 2014). As the model did not explain all the variation in the number of subsequent cases seen in the descriptive analyses, a second analysis was conducted examining the number of subsequent cases if the first case in a cluster had only extra-pulmonary disease.

Secondly, to investigate the impact of HIV on susceptibility to recent TB infection compared to reactivation of latent infection, logistic regression was used to

examine risk factors for being a subsequent case in a strain type cluster compared to being a non-clustered case or the first case of a cluster. All cases were included in this analysis.

For both analyses, univariable results were estimated and multivariable models were built, excluding patients missing data on one or more variables. Within-patient clustering (in patients with recurrent TB with different strain types) was not adjusted for as there were very few cases for which it was relevant. Linearity (of age group, year and IMD decile) was assessed using likelihood-ratio tests and variables were treated as categorical if P>0.05. To examine the representativeness of the cohort, the distributions of variables were compared for cases included and excluded from the study.

6.3 Results

6.3.1 Descriptive analysis

There were 37,162 cases of TB in adults aged ≥15 years notified to PHE in England, Wales and Northern Ireland between 2010 and 2014. 261/37,162 (0.7%) were recurrent cases within this time period. 23,146/37,162 (62.3%) were culture confirmed, and 18,913 (81.7%) were strain typed at ≥23 loci. 68 recurrent cases of TB had strain typing data available for both the first and subsequent isolates; 49 cases of recurrent TB were excluded where the isolates from the recurrent episode had the same strain type as the original isolate, however 19 recurrent cases where the strain type differed between the first and subsequent isolate(s) were included. Ultimately 18,864 TB cases were included in the analysis, representing 50.8% of TB cases notified in England, Wales and Northern Ireland from 2010-2014. The earliest date of evidence of TB was symptom onset date for 76% of cases, the remaining 24% used date of presentation to health services, specimen date, diagnosis date or case report date. The dates used to define earliest date of evidence of TB did not differ for HIV-positive and HIV-negative patients.

Of the cases included in the analysis, 10,709 (56.8%) were part of 2,284 strain type clusters. 2,238 (20.9%) were the first cases in a cluster (in 46 clusters the first case was aged <15 years and therefore excluded from the analysis) and 8,471 (79.1%) were subsequent cases. The median time between the first case in a cluster and the second case was 10 months (IQR 4-19).

6.3.2 HIV and clustering of tuberculosis cases

759/18,864 TB cases were co-infected with HIV (4.0%); 410 (54.0%) were clustered, of which 99 (24.2%) were the first case in a cluster and 311 (75.9%) were subsequent clustered cases.

Of the 8,471 subsequent cases in clusters, 572 (6.8%) had an HIV-positive first case, 7,775 (91.8%) had an HIV-negative first case, and the HIV status of the first case was unknown for 124 (1.5%) patients where the first case in the cluster was a child. Other demographic, socioeconomic and clinical factors are shown in Table 6.1.

The HIV status of the first case of a cluster was positively associated with the HIV status of subsequent cases (P<0.001). The prevalence of HIV among subsequent cases was higher in clusters with an HIV-positive first case (10.7%) than in clusters with an HIV-negative first case (3.2%). 6.4% of HIV-negative subsequent cases had an HIV-positive first case, compared to 19.9% of HIV-positive subsequent cases. 1,998/2,284 (87.5%) of clusters consisted of only HIV-negative TB patients, 11 clusters (0.5%) consisted of only HIV-positive TB patients, and 275 (12.0%) clusters were mixed.

The mean cluster size in the cohort was 4.8 (median 3, inter-quartile range 2-4, range 2-198); 4.7 for clusters where the first patient was HIV-negative, and 7.0 for clusters with an HIV-positive first case.

Table 6.1: The clustering status of TB cases by risk factor in England, Wales and Northern Ireland, 2010-2014.

	Total cases	Clustered cases (%)	Subsequent cases (% of clustered cases)	First cases (% of clustered cases)
Total	18,864	10,709 (56.8)	8,471 (79.1)	2,238 (20.9)
HIV status				
Negative	18,105	10,299 (56.9)	8,160 (79.2)	2,139 (20.8)
Positive	759	410 (54.0)	311 (75.9)	99 (24.1)
Year of TB notification				
2010	3,174	1,795 (56.6)	874 (48.7)	921 (51.3)
2011	4,296	2,443 (56.9)	1,786 (73.1)	657 (26.9)
2012	4,327	2,525 (58.4)	2,150 (85.1)	375 (14.9)
2013	3,696	2,130 (57.6)	1,940 (91.1)	190 (8.9)
2014	3,371	1,816 (53.9)	1,721 (94.8)	95 (5.2)
Sex				
Female	7,521	4,153 (55.2)	3,272 (78.8)	881 (21.2)
Male	11,323	6,547 (57.8)	5,196 (79.4)	1,351 (20.6)
Missing	20	9 (45.0)	3 (33.3)	6 (66.7)
Age (years)				
15-24	3,238	2,059 (63.6)	1,652 (80.2)	407 (19.8)
25-34	5,632	3,139 (55.7)	2,453 (78.1)	686 (21.9)
35-44	3,578	2,041 (57.0)	1,601 (78.4)	440 (21.6)
45-54	2,388	1,423 (59.6)	1,149 (80.7)	274 (19.3)
55-64	1,488	890 (59.8)	717 (80.6)	173 (19.4)
≥65	2,540	1,157 (45.6)	899 (77.7)	258 (22.3)
Ethnicity				
White	3,991	2,442 (61.2)	1,959 (80.2)	483 (19.8)
Black African	3,211	2,031 (63.3)	1,603 (78.9)	428 (21.1)
Black Other	588	458 (77.9)	391 (85.4)	67 (14.6)
Indian sub-continent	8,079	4,198 (52.0)	3,300 (78.6)	898 (21.4)
Mixed/other	2,525	1,330 (52.7)	1,029 (77.4)	301 (22.6)
Missing	470	250 (53.2)	189 (75.6)	61 (24.4)
Time since entry to the UK				
UK born	4,431	3,000 (67.7)	2,495 (83.2)	505 (16.8)
Within 2 years	2,535	1,313 (51.8)	979 (74.6)	334 (25.4)
2-5 years	2,999	1,509 (50.3)	1,154 (76.5)	355 (23.5)
5-10 years	2,743	1,485 (54.1)	1,149 (77.4)	336 (22.6)
More than 10 years	4,115	2,329 (56.6)	1,870 (80.3)	459 (19.7)
Missing	2,041	1,073 (52.6)	824 (76.8)	249 (23.2)
TB lineage				
Beijing	1,041	770 (74.0)	667 (86.6)	103 (13.4)
Euro-American	7,313	4,300 (58.8)	3,352 (78.0)	948 (22.0)
Central Asian Strain	5,280	3,285 (62.2)	2,674 (81.4)	611 (18.6)
East Asian Indian	2,674	1,046 (39.1)	769 (73.5)	277 (26.5)
Other	2,554	1,306 (51.1)	1,008 (77.2)	298 (22.8)
Missing	2			

	Total cases	Clustered cases (%)	Subsequent cases (% of clustered cases)	First cases (% of clustered cases)
IMD decile				
1	3,933	2,360 (60.0)	1,868 (79.2)	492 (20.8)
2	3,645	2,130 (58.4)	1,678 (78.8)	452 (21.2)
3	3,008	1,704 (56.6)	1,334 (78.3)	370 (21.7)
4	2,301	1,314 (57.1)	1,066 (81.1)	248 (18.9)
5	1,655	906 (54.7)	695 (76.7)	211 (23.3)
6	1,183	652 (55.1)	516 (79.1)	136 (20.9)
7	838	453 (54.1)	375 (82.8)	78 (17.2)
8	728	398 (54.7)	302 (75.9)	96 (24.1)
9	610	307 (50.3)	241 (78.5)	66 (21.5)
10	474	243 (51.3)	194 (79.8)	49 (20.2)
Missing	489	242 (49.5)	202 (83.5)	40 (16.5)
Drug misuse				
No	16,536	9,241 (55.9)	7,291 (78.9)	1,950 (21.1)
Yes	702	551 (78.5)	473 (85.8)	78 (14.2)
Missing	1,626	917 (56.4)	707 (77.1)	210 (22.9)
Alcohol misuse				
No	16,260	9,160 (56.3)	7,251 (79.2)	1,909 (20.8)
Yes	776	528 (68.0)	441 (83.5)	87 (16.5)
Missing	1,828	1,021 (55.9)	779 (76.3)	242 (23.7)
Homelessness				
No	16,771	9,480 (56.5)	7,500 (79.1)	1,980 (20.9)
Yes	666	449 (67.4)	372 (82.9)	77 (17.1)
Missing	1,427	780 (54.7)	599 (76.8)	181 (23.2)
Imprisonment				
No	16,210	9,097 (56.1)	7,200 (79.1)	1,897 (20.9)
Yes	649	484 (74.6)	410 (84.7)	74 (15.3)
Missing	2,005	1,128 (56.3)	861 (76.3)	267 (23.7)
Site of TB disease/Smear status†				
Pulmonary, smear positive	4,959	3,137 (63.3)	2,448 (78.0)	689 (22.0)
Pulmonary, smear negative/unknown	6,952	4,084 (58.7)	3,279 (80.3)	805 (19.7)
Extra-pulmonary	6,947	3,486 (50.2)	2,742 (78.7)	744 (21.3)
Missing	6	2 (33.3)	2 (100.0)	0 (0.0)

IMD: index of multiple deprivation score, TB: tuberculosis. † Patients with both pulmonary and extra-pulmonary disease were classed as having pulmonary disease.

6.3.3 Risk factors for being the first case in a cluster (a surrogate for onwards transmission of tuberculosis)

The number of subsequent cases for each TB case differed substantially by both HIV status, site of disease and smear status (Table 6.2); pulmonary smear-positive patients had a higher mean number of subsequent cases (1.1) than smear-negative or extra-pulmonary patients (0.8 and 0.7 respectively). However, extra-pulmonary TB patients with HIV had a higher mean number of subsequent cases compared to HIV-negative extra-pulmonary TB patients (2.5 versus 0.6, Table 6.2).

Table 6.2: The mean number of subsequent clustered cases, stratified by the HIV status, site of disease and smear status of the first TB case.

Site of disease† and smear status	HIV status of first case		
	HIV-negative Mean (SE)	HIV-positive Mean (SE)	Total Mean (SE)
Pulmonary smear positive	1.1 (0.02)	0.6 (0.07)	1.1 (0.02)
Pulmonary smear negative/unknown	0.8 (0.01)	0.9 (0.07)	0.8 (0.01)
Extra-pulmonary disease	0.6 (0.01)	2.5 (0.14)	0.7 (0.01)
Total	0.8 (0.01)	1.3 (0.05)	0.8 (0.01)

SE: standard error of the mean (Poisson distribution), TB: tuberculosis. †Patients with both pulmonary and extra-pulmonary disease were classed as having pulmonary disease.

Among pulmonary TB cases, there was no association between HIV co-infection and being the first case of a strain type cluster (compared to not being part of a strain type cluster) in univariable or multivariable zero-inflated Poisson regression models (multivariable OR 1.10 [0.79-1.53], Table 6.3). However, among pulmonary TB cases who were the first case of a strain type cluster, HIV co-infection was associated with a decreased number of subsequent clustered cases in the univariable model (IRR 0.76 [0.68-0.87], Table 6.3). This association remained in a multivariable zero-inflated Poisson regression model adjusted for year of TB notification, age, sex, ethnicity and country of birth, TB lineage, IMD decile, drug misuse, alcohol misuse, homelessness, imprisonment and smear status (IRR 0.76 [0.66-0.87]).

Extra-pulmonary TB cases with HIV co-infection were less likely to be the first case of a cluster (multivariable OR 1.93 [1.12-3.33], Table 6.4). However, where an

extra-pulmonary TB case was the first case in a cluster, HIV co-infection was associated with an increased number of subsequent cases (multivariable IRR 3.62 [3.12-4.19]).

The Vuong test was P<0.001 for both models, indicating that a zero-inflated Poisson model was a better fit for the data than the Poisson model.

Table 6.3: Univariable and multivariable zero-inflated Poisson regression of factors associated with the number of subsequent clustered cases for pulmonary TB cases in England, Wales and Northern Ireland, 2010-2014.

	Total pulmonary cases	Clustered cases (%)	First pulmonary cases (% of clustered cases)	Univariable (Number of subsequent cases)	Univariable (Non-clustered case)	Multivariable≠ (Number of subsequent cases)	Multivariable≠ (Non-clustered case)
				IRR (95% CI)	OR (95% CI)	IRR (95% CI)	OR (95% CI)
Total	11,911	7,221 (60.6)	2,056 (28.5)				
HIV status							
Negative	11,366	6,910 (60.8)	1,950 (28.2)	1.00	1.00	1.00	1.00
Positive	545	311 (57.1)	106 (34.1)	0.76 (0.68-0.87)	0.94 (0.72-1.23)	0.75 (0.65-0.86)	1.10 (0.79-1.53)
Year of TB diagnosis							
2010	2,028	1,205 (59.4)	716 (59.4)	1.00	1.00	1.00	1.00
2011	2,696	1,638 (60.8)	546 (33.3)	0.63 (0.59-0.66)	1.69 (1.46-1.96)	0.64 (0.60-0.68)	1.52 (1.29-1.80)
2012	2,650	1,670 (63.0)	379 (22.7)	0.39 (0.35-0.43)	1.87 (1.56-2.24)	0.38 (0.34-0.43)	1.53 (1.25-1.88)
2013	2,354	1,456 (61.9)	230 (15.8)	0.42 (0.35-0.49)	2.79 (2.20-3.53)	0.40 (0.34-0.48)	2.38 (1.83-3.11)
2014	2,183	1,252 (57.4)	185 (14.8)	0.64 (0.52-0.79)	4.53 (3.34-6.14)	0.59 (0.47-0.74)	4.04 (2.87-5.69)
Sex							
Female	4,562	2,661 (58.3)	765 (28.7)	1.00	1.00	1.00	1.00
Male	7,333	4,552 (62.1)	1,285 (28.2)	1.04 (0.99-1.09)	0.86 (0.76-0.96)	1.01 (0.95-1.06)	0.81 (0.70-0.93)
Missing	16	8 (50.0)	6 (75.0)				
Age (years)							
15-24	2,254	1,504 (66.7)	405 (26.9)	0.93 (0.87-1.00)	0.78 (0.65-0.92)	0.86 (0.79-0.93)	0.78 (0.63-0.96)
25-34	3,250	1,947 (59.9)	575 (29.5)	1.00	1.00	1.00	1.00
35-44	2,089	1,314 (62.9)	395 (30.1)	1.30 (1.22-1.39)	0.89 (0.75-1.05)	1.33 (1.24-1.43)	0.99 (0.81-1.22)
45-54	1,566	1,000 (63.9)	252 (25.2)	0.90 (0.83-0.99)	0.89 (0.73-1.08)	0.96 (0.87-1.07)	1.08 (0.84-1.38)
55-64	999	633 (63.4)	167 (26.4)	1.19 (1.09-1.30)	1.10 (0.87-1.39)	1.01 (0.91-1.13)	1.36 (1.01-1.82)

	Total pulmonary cases	Clustered cases (%)	First pulmonary cases (% of clustered cases)	Univariable (Number of subsequent cases)	Univariable (Non-clustered case)	Multivariable≠ (Number of subsequent cases)	Multivariable≠ (Non-clustered case)
				IRR (95% CI)	OR (95% CI)	IRR (95% CI)	OR (95% CI)
≥65	1,753	823 (46.9)	262 (31.8)	1.04 (0.96-1.13)	1.61 (1.34-1.94)	1.03 (0.93-1.14)	1.97 (1.53-2.53)
Ethnicity							
White	3,481	2,205 (63.3)	522 (23.7)	1.00	1.00	1.00	1.00
Black African	1,926	1,270 (65.9)	370 (29.1)	0.96 (0.89-1.02)	0.76 (0.64-0.91)	1.23 (1.12-1.36)	0.91 (0.70-1.19)
Black Other	406	322 (79.3)	68 (21.1)	0.89 (0.77-1.02)	0.51 (0.35-0.74)	0.86 (0.74-1.01)	0.58 (0.37-0.93)
Indian sub-continent	4,174	2,354 (56.4)	758 (32.2)	0.94 (0.88-0.99)	1.15 (1.00-1.33)	0.93 (0.85-1.01)	1.19 (0.94-1.51)
Mixed/other	1,621	894 (55.2)	273 (30.5)	0.60 (0.55-0.66)	1.03 (0.85-1.26)	0.68 (0.60-0.77)	1.10 (0.83-1.46)
Missing	303	176 (58.1)	65 (36.9)				
Time since entry to the UK							
UK born	3,631	2,526 (69.6)	540 (21.4)	1.00	1.00	1.00	1.00
Within 2 years	1,536	833 (54.2)	311 (37.3)	0.70 (0.64-0.75)	1.10 (0.91-1.32)	0.65 (0.59-0.71)	1.27 (0.99-1.63)
2-5 years	1,549	815 (52.6)	267 (32.8)	0.75 (0.69-0.81)	1.24 (1.02-1.50)	0.76 (0.69-0.84)	1.35 (1.05-1.74)
5-10 years	1,543	897 (58.1)	283 (31.5)	0.82 (0.76-0.89)	1.10 (0.91-1.33)	0.76 (0.69-0.83)	1.25 (0.97-1.60)
More than 10 years	2,423	1,460 (60.3)	423 (29.0)	0.89 (0.83-0.95)	1.23 (1.04-1.46)	0.84 (0.77-0.91)	1.10 (0.87-1.40)
Missing	1,229	690 (56.1)	232 (33.6)				
TB lineage							
Beijing	706	525 (74.4)	93 (17.7)	1.00	1.00	1.00	1.00
Euro-American	5,306	3,233 (60.9)	898 (27.8)	0.51 (0.47-0.56)	1.05 (0.80-1.39)	0.46 (0.41-0.50)	1.11 (0.80-1.54)
Central Asian Strain	2,955	1,948 (65.9)	547 (28.1)	0.72 (0.66-0.79)	1.05 (0.79-1.40)	0.78 (0.70-0.86)	1.01 (0.72-1.43)
East Asian Indian	1,271	551 (43.4)	235 (42.6)	0.42 (0.37-0.48)	1.59 (1.16-2.17)	0.52 (0.45-0.59)	1.54 (1.06-2.23)
Other	1,673	964 (57.6)	283 (29.4)	0.48 (0.43-0.53)	1.17 (0.86-1.59)	0.43 (0.38-0.48)	1.18 (0.82-1.69)

	Total pulmonary cases	Clustered cases (%)	First pulmonary cases (% of clustered cases)	Univariable (Number of subsequent cases)	Univariable (Non-clustered case)	Multivariable≠ (Number of subsequent cases)	Multivariable≠ (Non-clustered case)
				IRR (95% CI)	OR (95% CI)	IRR (95% CI)	OR (95% CI)
Missing	2						
IMD decile							
1	2,581	1,654 (64.1)	440 (26.6)	-	-	-	-
2	2,238	1,383 (61.8)	396 (28.6)	-	-	-	-
3	1,851	1,117 (60.3)	335 (30.0)	-	-	-	-
4	1,425	873 (61.3)	247 (28.3)	-	-	-	-
5	1,039	609 (58.6)	191 (31.4)	-	-	-	-
6	737	437 (59.3)	125 (28.6)	-	-	-	-
7	525	306 (58.3)	80 (26.1)	-	-	-	-
8	486	276 (56.8)	82 (29.7)	-	-	-	-
9	390	224 (57.4)	63 (28.1)	-	-	-	-
10	305	171 (56.1)	54 (31.6)	-	-	-	-
Missing	334	171 (51.2)	43 (25.1)	-	-	-	-
For each decile increase	-	-	-	0.96 (0.95-0.97)	1.02 (0.99-1.04)	0.96 (0.95-0.97)	1.00 (0.97-1.03)
Drug misuse							
No	10,165	6,061 (59.6)	1,768 (29.2)	1.00	1.00	1.00	1.00
Yes	639	507 (79.3)	82 (16.2)	1.14 (1.02-1.27)	0.61 (0.45-0.83)	0.88 (0.76-1.01)	0.84 (0.56-1.28)
Missing	1,107	653 (59.0)	206 (31.5)				
Alcohol misuse							
No	10,039	6,043 (60.2)	1,747 (28.9)	1.00	1.00	1.00	1.00
Yes	670	470 (70.1)	87 (18.5)	1.85 (1.71-2.01)	0.97 (0.74-1.26)	1.69 (1.54-1.86)	1.18 (0.84-1.66)
Missing	1,202	708 (58.9)	222 (31.4)				
Homelessness							

	Total pulmonary cases	Clustered cases (%)	First pulmonary cases (% of clustered cases)	Univariable (Number of subsequent cases)	Univariable (Non-clustered case)	Multivariable≠ (Number of subsequent cases)	Multivariable≠ (Non-clustered case)
				IRR (95% CI)	OR (95% CI)	IRR (95% CI)	OR (95% CI)
No	10,398	6,277 (60.4)	1,799 (28.7)	1.00	1.00	1.00	1.00
Yes	567	393 (69.3)	85 (21.6)	0.90 (0.80-1.02)	0.74 (0.55-0.99)	0.63 (0.54-0.72)	0.88 (0.59-1.30)
Missing	946	551 (58.2)	172 (31.2)				
Imprisonment							
No	9,990	5,978 (59.8)	1,725 (28.9)	1.00	1.00	1.00	1.00
Yes	553	423 (76.5)	82 (19.4)	1.07 (0.96-1.20)	0.86 (0.76-0.96)	1.10 (0.97-1.26)	0.85 (0.57-1.26)
Missing	1,368	820 (59.9)	249 (30.4)				
Smear status							
Smear positive	4,959	3,137 (63.3)	901 (28.7)	1.00	1.29	1.00	1.00
Smear negative or unknown	6,952	4,084 (58.7)	1,155 (28.3)	0.87 (0.83-0.92)	1.94 (1.78-2.12)	0.83 (0.79-0.88)	1.17 (1.02-1.34)

CI: confidence interval, IMD: index of multiple deprivation score, IRR: incidence rate ratio (Poisson part) for an increased number of subsequent clustered cases, OR: odds ratio (zero-inflated part) for the odds of being a non-clustered case, compared to being the first pulmonary case of a cluster, TB: tuberculosis. ≠ Adjusted for all variables shown in the table. The multivariable model included 5,694 TB cases after 1,052 were excluded due to missing data on one or more of sex (n=14), ethnicity (n=192), time since entry to the UK (n=771) or IMD score (n=206). †Cases missing data were considered not to have these social risk factors.

Table 6.4: Univariable and multivariable zero-inflated Poisson regression of factors associated with the number of subsequent clustered cases for extra-pulmonary TB cases in England, Wales and Northern Ireland, 2010-2014.

	Total extra-pulmonary cases	Clustered cases (%)	First extra-pulmonary cases (% of clustered cases)	Univariable (Number of subsequent cases) IRR (95% CI)	Univariable (Non-clustered case) OR (95% CI)	Multivariable≠ (Number of subsequent cases) IRR (95% CI)	Multivariable≠ (Non-clustered case) OR (95% CI)
Total	6,953	3,488 (50.2)	744 (21.3)				
HIV status							
Negative	6,739	3,389 (50.3)	722 (21.3)	1.00	1.00	1.00	1.00
Positive	214	99 (46.3)	22 (22.2)	4.16 (3.71-4.67)	1.38 (0.86-2.19)	3.62 (3.12-4.19)	1.93 (1.12-3.33)
Year of TB diagnosis							
2010	1,146	590 (51.5)	293 (49.7)	1.00	1.00	1.00	1.00
2011	1,600	805 (50.3)	242 (30.1)	0.77 (0.71-0.84)	1.65 (1.34-2.02)	0.72 (0.66-0.80)	1.45 (1.15-1.84)
2012	1,677	855 (51.0)	122 (14.3)	0.56 (0.48-0.66)	2.84 (2.21-3.64)	0.60 (0.51-0.71)	2.57 (1.93-3.41)
2013	1,342	674 (50.2)	62 (9.2)	0.39 (0.29-0.51)	2.83 (1.94-4.13)	0.45 (0.34-0.61)	2.82 (1.88-4.22)
2014	1,188	564 (47.5)	25 (4.4)	0.90 (0.74-1.10)	6.86 (4.36-10.80)	1.11 (0.82-1.51)	7.82 (4.67-13.11)
Sex							
Female	2,959	1,492 (50.4)	323 (21.6)	1.00	1.00	1.00	1.00
Male	3,990	1,995 (50.0)	421 (21.1)	1.25 (1.15-1.35)	1.11 (0.94-1.31)	1.22 (1.12-1.34)	0.99 (0.81-1.21)
Missing	4	1 (25.0)	(0.0)				
Age (years)							
15-24	984	555 (56.4)	111 (20.0)	2.26 (2.01-2.54)	1.10 (0.85-1.42)	1.66 (1.46-1.89)	1.07 (0.80-1.45)
25-34	2,382	1,192 (50.0)	266 (22.3)	1.00	1.00	1.00	1.00
35-44	1,489	727 (48.8)	156 (21.5)	1.67 (1.49-1.87)	1.27 (1.01-1.60)	1.43 (1.26-1.61)	1.32 (1.00-1.75)
45-54	822	423 (51.5)	83 (19.6)	1.37 (1.19-1.59)	1.19 (0.89-1.58)	1.39 (1.18-1.63)	1.46 (1.02-2.10)
55-64	489	257 (52.6)	52 (20.2)	1.73 (1.48-2.02)	1.18 (0.84-1.66)	1.92 (1.60-2.31)	1.45 (0.94-2.24)
≥65	787	334 (42.4)	76 (22.8)	1.08 (0.92-1.26)	1.34 (1.00-1.81)	0.94 (0.78-1.14)	1.40 (0.92-2.12)

	Total extra-pulmonary cases	Clustered cases (%)	First extra-pulmonary cases (% of clustered cases)	Univariable (Number of subsequent cases) IRR (95% CI)	Univariable (Non-clustered case) OR (95% CI)	Multivariable≠ (Number of subsequent cases) IRR (95% CI)	Multivariable≠ (Non-clustered case) OR (95% CI)
Ethnicity							
White	510	237 (46.5)	49 (20.7)	1.00	1.00	1.00	1.00
Black African	1,285	761 (59.2)	150 (19.7)	1.76 (1.45-2.14)	0.74 (0.51-1.08)	0.85 (0.65-1.10)	0.49 (0.27-0.89)
Black Other	182	136 (74.7)	17 (12.5)	3.69 (2.92-4.66)	0.62 (0.32-1.19)	2.84 (2.18-3.70)	0.57 (0.26-1.25)
Indian sub-continent	3,905	1,844 (47.2)	414 (22.5)	1.21 (1.00-1.46)	0.96 (0.68-1.35)	0.64 (0.49-0.83)	0.64 (0.36-1.12)
Mixed/other	904	436 (48.2)	101 (23.2)	0.97 (0.78-1.21)	0.80 (0.54-1.20)	0.58 (0.44-0.78)	0.50 (0.27-0.93)
Missing	167	74 (44.3)	13 (17.6)				
Time since entry to the UK							
UK born	800	474 (59.3)	86 (18.1)	1.00	1.00	1.00	1.00
Within 2 years	999	480 (48.0)	107 (22.3)	1.75 (1.52-2.01)	1.51 (1.09-2.10)	2.06 (1.70-2.50)	2.56 (1.62-4.05)
2-5 years	1,450	694 (47.9)	156 (22.5)	0.72 (0.61-0.84)	1.16 (0.85-1.59)	0.99 (0.81-1.22)	1.72 (1.10-2.70)
5-10 years	1,200	588 (49.0)	134 (22.8)	0.87 (0.74-1.01)	1.20 (0.87-1.65)	1.16 (0.94-1.42)	1.83 (1.15-2.89)
More than 10 years	1,692	869 (51.4)	185 (21.3)	0.95 (0.83-1.09)	1.15 (0.85-1.56)	1.28 (1.04-1.57)	1.42 (0.91-2.23)
Missing	812	383 (47.2)	76 (19.8)				
TB lineage							
Beijing	335	245 (73.1)	34 (13.9)	1.00	1.00	1.00	1.00
Euro-American	2,007	1,067 (53.2)	236 (22.1)	0.46 (0.39-0.54)	1.21 (0.79-1.87)	0.41 (0.34-0.49)	1.16 (0.70-1.93)
Central Asian Strain	2,325	1,337 (57.5)	255 (19.1)	0.75 (0.65-0.87)	1.36 (0.89-2.09)	0.76 (0.64-0.90)	1.37 (0.82-2.26)
East Asian Indian	1,403	495 (35.3)	133 (26.9)	0.48 (0.41-0.58)	2.18 (1.40-3.42)	0.55 (0.45-0.67)	2.07 (1.23-3.48)
Other	881	342 (38.8)	85 (24.9)	0.66 (0.56-0.79)	2.17 (1.36-3.47)	0.62 (0.51-0.75)	1.92 (1.12-3.30)
Missing	2						

	Total extra-pulmonary cases	Clustered cases (%)	First extra-pulmonary cases (% of clustered cases)	Univariable (Number of subsequent cases) IRR (95% CI)	Univariable (Non-clustered case) OR (95% CI)	Multivariable≠ (Number of subsequent cases) IRR (95% CI)	Multivariable≠ (Non-clustered case) OR (95% CI)
IMD decile							
1	1,352	706 (52.2)	160 (22.7)	-	-	-	-
2	1,407	747 (53.1)	156 (20.9)	-	-	-	-
3	1,157	587 (50.7)	136 (23.2)	-	-	-	-
4	876	441 (50.3)	72 (16.3)	-	-	-	-
5	616	297 (48.2)	75 (25.3)	-	-	-	-
6	446	215 (48.2)	45 (20.9)	-	-	-	-
7	313	147 (47.0)	26 (17.7)	-	-	-	-
8	242	122 (50.4)	30 (24.6)	-	-	-	-
9	220	83 (37.7)	17 (20.5)	-	-	-	-
10	169	72 (42.6)	12 (16.7)	-	-	-	-
Missing	155	71 (45.8)	15 (21.1)	-	-	-	-
For each decile increase	-	-	-	0.93 (0.92-0.95)	1.03 (1.00-1.07)	0.97 (0.95-0.99)	1.03 (0.99-1.08)
Drug misuse							
No	6,371	3,180 (49.9)	675 (21.2)	1.00	1.00	1.00	1.00
Yes	63	44 (69.8)	7 (15.9)	0.41 (0.21-0.82)	0.34 (0.10-1.18)	0.49 (0.27-0.90)	0.31 (0.06-1.66)
Missing	519	264 (50.9)	62 (23.5)				
Alcohol misuse							
No	6,221	3,117 (50.1)	654 (21.0)	1.00	1.00	1.00	1.00
Yes	106	58 (54.7)	13 (22.4)	1.44 (1.13-1.83)	0.89 (0.47-1.66)	1.79 (1.34-2.38)	1.09 (0.47-2.51)
Missing	626	313 (50.0)	77 (24.6)				
Homelessness							
No	6,373	3,203 (50.3)	679 (21.2)	1.00	1.00	1.00	1.00
Yes	99	56 (56.6)	7 (12.5)	0.29 (0.12-0.72)	0.71 (0.21-2.33)	0.23 (0.09-0.58)	0.62 (0.10-3.94)
Missing	481	229 (47.6)	58 (25.3)				

	Total extra-pulmonary cases	Clustered cases (%)	First extra-pulmonary cases (% of clustered cases)	Univariable (Number of subsequent cases) IRR (95% CI)	Univariable (Non-clustered case) OR (95% CI)	Multivariable≠ (Number of subsequent cases) IRR (95% CI)	Multivariable≠ (Non-clustered case) OR (95% CI)
Imprisonment							
No	6,220	3,119 (50.1)	657 (21.1)	1.00	1.00	1.00	1.00
Yes	96	61 (63.5)	8 (13.1)	0.06 (0.03-0.13)	1.11 (0.94-1.31)	0.17 (0.04-0.82)	0.36 (0.01-8.94)
Missing	637	308 (48.4)	79 (25.6)				

CI: confidence interval, IMD: index of multiple deprivation score IRR: incidence rate ratio (Poisson part) for an increased number of subsequent clustered cases, OR: odds ratio (zero-inflated part) for the odds of being a non-clustered case, compared to being the first extra-pulmonary case of a cluster, TB: tuberculosis. ≠ Adjusted for all variables shown in the table. The multivariable model included 3,576 extra-pulmonary TB cases after 633 were excluded due to missing data on one or more of sex (n=3), ethnicity (n=106), time since entry to the UK (n=505), IMD score (n=99) or TB lineage (n=1). †Cases missing data were considered not to have these social risk factors.

6.3.4 Risk factors for being a subsequent case in a cluster (a surrogate for recent infection with tuberculosis)

HIV co-infection was negatively associated with being a subsequent case in a cluster in univariable analysis (OR 0.85 [0.73-0.98], Table 6.5) and multivariable analysis (OR 0.82 [0.69-0.98]) adjusted for year of TB diagnosis, age, sex, ethnicity and country of birth, TB lineage, IMD decile, drug misuse, alcohol misuse, homelessness, and imprisonment.

6.3.5 Assessment of potential bias and misclassification

The distribution of variables for cases included in the study population was generally similar to those that were excluded for being culture-negative or strain typed at <23 loci (Table 6.6). The cases included in the analysis did not substantially differ in terms of age, sex, ethnicity, place of birth (UK or abroad), time since entry to the UK, year of TB diagnosis or presence of social risk factors, but TB patients with only extra-pulmonary disease were less likely to be culture-confirmed and typed at ≥23 loci (and therefore less likely to be included).

Of the 18,864 cases included in the analysis, 5,462 (29.0%) were typed at 23 loci and 1,186 (11.0% of clustered cases) could have belonged to more than one cluster. The prevalence of HIV was similar for cases typed with 23 loci (4.1%) and cases typed with 24 loci (4.0%). 122 cases (0.7%) had more than one strain type due to mixed infection, and 4 cases (0.02%) belonged to more than one cluster.

Table 6.5: Univariable and multivariable logistic regression of factors associated with being a subsequent TB case in a cluster (a surrogate for recent infection) compared to being the first case or a non-clustered case, in England, Wales and Northern Ireland from 2010-2014.

	Univariable OR (95% CI)	Multivariable≠ OR (95% CI)
HIV status		
Negative	1.00	1.00
Positive	0.85 (0.73-0.98)	0.82 (0.69-0.98)
Year of TB notification		
2010	1.00	1.00
2011	1.87 (1.70-2.07)	2.06 (1.84-2.31)
2012	2.60 (2.36-2.87)	3.06 (2.74-3.43)
2013	2.91 (2.63-3.22)	3.38 (3.02-3.80)
2014	2.74 (2.48-3.04)	3.17 (2.82-3.56)
Sex		
Female	1.00	
Male	1.10 (1.04-1.17)	1.09 (1.02-1.17)
Age (years)		
15-24	1.35 (1.24-1.47)	1.19 (1.08-1.32)
25-34	1.00	1.00
35-44	1.05 (0.96-1.14)	0.92 (0.83-1.02)
45-54	1.20 (1.09-1.32)	0.90 (0.80-1.01)
55-64	1.21 (1.07-1.35)	0.96 (0.83-1.10)
≥65	0.71 (0.64-0.78)	0.51 (0.45-0.57)
Ethnicity		
White	1.00	1.00
Black African	1.03 (0.94-1.13)	1.51 (1.31-1.73)
Black Other	2.06 (1.72-2.47)	2.25 (1.82-2.78)
Indian sub-continent	0.72 (0.66-0.77)	0.92 (0.81-1.04)
Mixed/other	0.71 (0.65-0.79)	0.98 (0.85-1.13)
Time since entry to the UK		
UK born	1.00	1.00
Within 2 years	0.49 (0.44-0.54)	0.41 (0.36-0.47)
2-5 years	0.49 (0.44-0.53)	0.39 (0.35-0.44)
5-10 years	0.56 (0.51-0.62)	0.49 (0.43-0.55)
More than 10 years	0.65 (0.59-0.70)	0.61 (0.54-0.69)
TB lineage		
Beijing	1.00	1.00
Euro-American	0.47 (0.41-0.54)	0.38 (0.33-0.45)
Central Asian Strain	0.58 (0.50-0.66)	0.63 (0.54-0.74)
East Asian Indian	0.23 (0.19-0.26)	0.23 (0.19-0.28)
Other	0.37 (0.31-0.42)	0.32 (0.27-0.38)
IMD decile		
For each decile increase	0.97 (0.96-0.98)	0.98 (0.96-0.99)
Drug misuse		
No	1.00	1.00
Yes	2.62 (2.24-3.08)	1.53 (1.25-1.87)

		Univariable OR (95% CI)	Multivariable≠ OR (95% CI)
Alcohol misuse			
	No	1.00	1.00
	Yes	1.65 (1.43-1.91)	1.21 (1.01-1.45)
Homelessness			
	No	1.00	1.00
	Yes	1.58 (1.35-1.84)	1.03 (0.85-1.24)
Imprisonment			
	No	1.00	1.00
	Yes	2.16 (1.84-2.54)	1.26 (1.03-1.54)

CI: confidence interval, IMD: index of multiple deprivation score, OR: odds ratio, TB: tuberculosis. ≠ Adjusted for all variables shown in the table. The multivariable model included 16,171 TB cases after 2,693 were excluded due to missing data on one or more of sex (n=20), ethnicity (n=470), time since entry to the UK (n=2,041), IMD score (n=489) and/or TB lineage (n=2). †Cases missing data were considered not to have these social risk factors.

Table 6.6: Distribution of the variables included in the regression models, for cases included and excluded from the study population.

Variable		Excluded cases (not strain typed) n (column %)	Included cases (strain typed n (column %)
HIV			
	Negative	17,664 (96.5)	18,105 (96.0)
	Positive	634 (3.5)	759 (4.0)
Year			
	2010	4,358 (23.8)	3,174 (16.8)
	2011	3,789 (20.7)	4,296 (22.8)
	2012	3,581 (19.6)	4,327 (22.9)
	2013	3,480 (19.0)	3,696 (19.6)
	2014	3,090 (16.9)	3,371 (17.9)
Sex			
	Female	8,063 (44.1)	7,521 (39.9)
	Male	10,195 (55.7)	11,323 (60.0)
Age group			
	15-24	2,351 (12.9)	3,238 (17.2)
	25-34	4,958 (27.1)	5,632 (29.9)
	35-44	3,622 (19.8)	3,578 (19.0)
	45-54	2,552 (14.0)	2,388 (12.7)
	55-64	1,964 (10.7)	1,488 (7.9)
	≥65	2,851 (15.6)	2,540 (13.5)
Ethnicity			
	White	3,574 (19.5)	3,991 (21.2)
	Black African	2,739 (15.0)	3,211 (17.0)
	Black Other	472 (2.6)	588 (3.1)
	Indian sub-continent	9,020 (49.3)	8,079 (42.8)
	Mixed/other	2,069 (11.3)	2,525 (13.4)
Time since entry to the UK			
	UK born	4,286 (23.4)	4,431 (23.5)
	<2 years	1,964 (10.7)	2,535 (13.4)
	2-5 years	2,448 (13.4)	2,999 (15.9)
	5-10 years	2,647 (14.5)	2,743 (14.5)
	≥10 years	4,907 (26.8)	4,115 (21.8)
Drug misuse			
	No	16,412 (89.7)	16,536 (87.7)
	Yes	364 (2.0)	702 (3.7)
IMD decile			
	1	3,964 (21.7)	3,933 (20.9)
	2	3,478 (19.0)	3,645 (19.3)
	3	2,741 (15.0)	3,008 (16.0)
	4	2,040 (11.2)	2,301 (12.2)
	5	1,507 (8.2)	1,655 (8.8)
	6	1,200 (6.6)	1,183 (6.3)
	7	892 (4.9)	838 (4.4)
	8	728 (4.0)	728 (3.9)

Variable		Excluded cases (not strain typed) n (column %)	Included cases (strain typed n (column %)
	9	663 (3.6)	610 (3.2)
	10	539 (3.0)	474 (2.5)
Alcohol misuse			
	No	16,135 (88.2)	16,260 (86.2)
	Yes	413 (2.3)	776 (4.1)
Homelessness			
	No	16,494 (90.1)	16,771 (88.9)
	Yes	373 (2.0)	666 (3.5)
Imprisonment			
	No	15,984 (87.4)	16,210 (85.9)
	Yes	370 (2.0)	649 (3.4)
Site of TB disease			
	Pulmonary smear positive	1,280 (7.0)	4,959 (26.3)
	Pulmonary smear negative	6,044 (33.0)	6,952 (36.9)
	Extra-pulmonary disease only	10,843 (59.3)	6,947 (36.8)

IMD: index of multiple deprivation, n: number, TB: tuberculosis.

6.4 Discussion

6.4.1 Summary of findings

The aims of this chapter were to investigate the impact of HIV on the infectiousness of TB patients, by examining the association between HIV co-infection of TB patients and the number of subsequent clustered cases, and to examine whether TB patients with HIV were more often due to recent infection or reactivation of LTBI. Pulmonary TB patients with HIV had fewer subsequent clustered cases than HIV-negative pulmonary TB patients, and there was no association between HIV co-infection and not being clustered. Extra-pulmonary TB cases with HIV were less likely to be the first case of a cluster; but if they were the first case, had a greater number of subsequent clustered cases than extra-pulmonary TB patients without HIV. There was a negative association between HIV co-infection and being a subsequent case in a cluster, compared to being the first case or a non-clustered case.

6.4.2 Strengths and limitations of the study

This study benefitted from a large sample of all culture-positive TB cases strain typed at ≥23 loci in England, Wales and Northern Ireland over a five-year period, and represents over 50% of all TB cases in the country during this time. The coverage was comparable to national studies of a similar size in the Netherlands,[72, 242] and considerably higher than the 31% coverage in a previous study in England which did not include data on HIV co-infection.[238] Studies in Norway and Denmark have achieved higher rates of coverage nationally (67-69% of all TB cases), however these studies had limited or no information on HIV status and much smaller overall sample sizes.[243, 244]

24-loci MIRU-VNTR is a highly discriminative, high-throughput method of genotyping *MTBC*,[245, 246] and has been widely used in TB cluster investigations. Strain typing was done at national reference laboratories according to national guidelines, thereby reducing measurement error and misclassification. However, analyses using WGS have demonstrated that identical 24-loci MIRU-VNTR profiles do not always have sufficiently high resolution to distinguish between closely related, but distinct, lineages,[69, 234] particularly in clusters of foreign-born patients.[234] As HIV co-infection was negatively associated with both the number of subsequent clustered cases (section 6.3.3) and with being recently infected (section 6.3.4), over-estimation of clustering among foreign-born TB patients (who are more likely to be co-infected with HIV, Chapter 5) may have biased our estimates of the effect of HIV on TB transmission towards the null. It is therefore likely that the true magnitude of the association between HIV and reactivation of LTBI is even greater than reported here.

Strain typing cannot be used to confirm transmission between patients with identical strain types, only to refute it where strain types are different. Strain typing also cannot identify the direction of transmission between two individuals, which can sometimes be inferred (although not proven) by using WGS data to examine the accumulation of small genomic mutations. Furthermore, shared strain types may not

represent recent transmission, particularly in patients born abroad who may have been infected with common endemic strain types before entering the UK.[52] It is likely that the 54% of TB cases who were clustered in this analysis was an overestimation of the proportion of TB attributable to recent transmission.

Whilst the sample size was large, only approximately 50% of TB cases nationally were included in this analysis as strain typing relies on culture of mycobacterial samples. Low sampling fractions result in underestimation of the extent of clustering,[60, 62] as cases can be misclassified as not-clustered if the case they cluster with has not been strain typed. However, it has been shown that a low sampling fraction does not bias estimations of risk factors associated with clustering when the study is sufficiently powered, because risk factors for being clustered are generally also risk factors for increased cluster size.[60, 62] Some of the non-strain typed cases may have been the first cases of clusters, meaning some subsequent cases in clusters may have been misclassified as the first cases. Some cases may have been misclassified as not being clustered due to the time limiting of the sample, however, the effect of this is minimal compared to that of low sampling fractions.[62] It is also possible that non-clustered TB cases, which were assumed to be reactivation of LTBI, could be the result of recent infection acquired outside of England, Wales and Northern Ireland.

Clustered TB cases were classed as being the first case or a subsequent case in clusters according to their earliest date of evidence of TB. Consequently, the order of patients within clusters may have been misclassified as patients may not present to care in the order in which they were infected. In particular, TB patients with HIV may reactivate more quickly than those who are HIV-negative, meaning they are more likely to develop active TB sooner, whilst people who have been diagnosed with HIV may be more engaged with health services and/or diagnosed with TB sooner. If this is the case, differential misclassification of TB patients with HIV as the first case in a cluster would be expected, when in fact they may just be the first patient in that cluster who presented to care. However, HIV-positive cases typically had fewer subsequent cases

and were less likely to be subsequent cases in clusters, and so any misclassification to this effect would have biased the results towards the null and caused underestimation of the impact of HIV. Furthermore, less than 50% of TB patients were aware of their HIV infection when diagnosed with TB (Chapter 4) and therefore this would not have influenced the time it took them to present to care. Where possible, symptom onset date was used to determine the order of patients in clusters, as much onward transmission will occur before a TB patient is diagnosed.

TB cases included in the analysis were more likely to have smear-positive pulmonary TB and less likely to have extra-pulmonary TB only compared to cases which were excluded as they were not strain typed at ≥23 loci. This is unlikely to have biased the association between HIV and whether a case is a subsequent case in a cluster (section 6.3.4), as site of disease is not on the causal pathway between HIV and infection with or reactivation of TB. However, it may have led to overestimation of the odds ratio for the association between HIV and being the first case in a cluster; cases with only extra-pulmonary TB were less likely to be the first case in a cluster but also less likely to be HIV-positive (Chapter 5), and this over-representation of HIV-negative extra-pulmonary TB cases in the analyses may have resulted in odds ratios closer to the null.

Cases which were only typed at 23 loci could have been assigned to more than one cluster. Misclassifying which cluster a case belonged to could have led to misclassification of whether the case was the first or a subsequent case in a cluster. However, HIV status was not associated with whether a case was typed at 23 or 24 loci, and therefore any misclassification is likely to be non-differential, and therefore should not have biased the results. Some cases had a mixed infection, and therefore could have belonged to more than one cluster; however only four cases with mixed infection belonged to more than one cluster and therefore any bias would be very minimal.

Similarly to Chapters 4 and 5, a composite variable was created for country of birth and time since entry to the UK. Due to the smaller sample size available for this analysis, this variable was not stratified by the incidence of TB or the prevalence of HIV in the patients' country of birth, which may have confounded the associations between HIV and the clustering outcomes. To minimise confounding, factors associated with HIV co-infection of TB cases in Chapter 5 were included in the multivariable models, as were other factors which were considered *a priori* to potentially be associated with onwards transmission or reactivation of LTBI. However, the results may have been confounded by other factors not included in the model. Non-completion of treatment and drug resistance may differentially bias the association between HIV co-infection and onwards transmission of TB. Missing data meant this study lacked power to investigate these associations, which would benefit from further research. The relatively small numbers of patients with HIV and/or social risk factors in this analysis meant that the study did not have sufficient power to investigate interactions between the different social risk factors, HIV, and clustering. This could be a useful area for future research.

As this was a study of routinely collected TB surveillance data, no data was available on the HIV status of the contacts of TB cases, which may have confounded the association between HIV status and the number of subsequent clustered cases. If the contacts of HIV-positive TB patients were more immunosuppressed, then a higher number of subsequent clustered cases would be expected; however pulmonary TB patients with HIV had fewer subsequent cases than HIV-negative pulmonary TB patients. Therefore, the reduction in onwards transmission of TB as a result of HIV co-infection may have been underestimated in this study.

Data on HIV status was not available for children (aged <15 years), and therefore children could not be included in this analysis. To limit bias, children were included when determining whether TB cases were clustered, the size of clusters and whether a case was the first or a subsequent case in a cluster; but then excluded from

the risk factor analyses. TB in children living with HIV is relatively rare in the UK (estimated at 196/100,000 PY from 1996-2014),[172] and therefore the impact of HIV on TB transmission from children is likely to be minimal.

6.4.3 Implications of the research

Among pulmonary TB cases, HIV co-infection was negatively associated with the number of subsequent clustered cases (Table 6.3), particularly for smear-positive patients (Table 6.2). This finding is consistent with the results of contact studies across high- and low-burden settings, which have found lower risks of LTBI and TB disease among the contacts of HIV-positive patients than HIV-negative TB patients, adding weight to the suggestion that patients with pulmonary TB and HIV may be less infectious. Among extra-pulmonary TB cases, there was a strong association between HIV co-infection and not being the first case of a cluster, again suggesting that patients with HIV are less infectious, in this case as a result of their extra-pulmonary disease. However, where HIV-positive extra-pulmonary TB patients were the first case of a cluster, they had a substantially higher number of subsequent clustered cases than HIV-negative extra-pulmonary TB patients. As it is generally accepted that patients with only extra-pulmonary TB disease are not infectious, it is unlikely that these patients are driving transmission within these larger clusters. It is possible that transmission within these clusters was driven by patients with undiagnosed TB, or patients whose strain type was unknown. Increased cluster size may also be the result of transmission chains occurring within clusters. The prevalence of HIV was higher among subsequent cases in clusters with an HIV-positive first case than clusters with HIV-negative first cases, and so it is likely that the increased size of these clusters is because the contacts of the infectious case are more susceptible to infection and progression to active disease. Regardless of whether these HIV-positive cases are the ‘true’ first case in a cluster or merely the first observed case in a cluster, the first observed patient is still a point at which interventions can be considered and targeted for cluster investigations.

Whilst extra-pulmonary TB cases may not drive transmission, the knowledge that extra-pulmonary TB cases with HIV can be the first detected case of a substantially larger cluster is important for directing cluster investigations. Furthermore, as around 50% of co-infected patients are only diagnosed with HIV at the time of their TB diagnosis (Chapter 4), targeting HIV screening to the contacts of TB patients with HIV could result in earlier diagnosis of HIV infections, providing the opportunity to initiate ART and prevent TB disease from occurring. Targeting LTBI treatment to these high-risk contacts could also further decrease the incidence of TB.

A negative association was found between HIV co-infection and being a subsequent case in a cluster, compared to being the first case or a non-clustered case. This suggests that TB in patients with HIV is more often the result of reactivation of remotely-acquired LTBI than recent infection. These cases of TB may be preventable if PLHIV, particularly those born abroad, could be tested and treated for LTBI. This finding contrasts with that of a meta-analysis of the association between HIV and clustering of TB cases in HIV-endemic populations, which concluded that HIV-associated TB was more often the result of recent infection than reactivation of LTBI.[92] This difference is likely the result of the different settings; the higher incidence of TB in the general population in countries where HIV is endemic will lead to a greater force of infection which may differentially affect PLHIV, who are more likely to be immunosuppressed. In contrast, in the UK (and other low-burden settings) the majority of TB cases are in foreign-born patients and transmission is generally considered to be minimal.[52] As there is generally less exposure to TB, HIV contributes more to reactivation of LTBI than to new TB infections.

6.4.4 Conclusions

In conclusion, pulmonary TB patients with HIV had fewer subsequent clustered cases than patients without HIV. However, when patients with HIV and extra-pulmonary TB were the first case of a cluster, they had a higher number of subsequent cases. HIV prevalence was higher among the subsequent cases of HIV-positive first

cases than the subsequent cases of HIV-negative first cases, suggesting that the higher number of subsequent cases for extra-pulmonary TB patients with HIV is because their contacts are more susceptible to infection and progression of disease. These findings suggest that overall, pulmonary TB patients with HIV are less infectious. Furthermore, TB cases with HIV were less likely to be a subsequent case within a cluster, which suggests that HIV-associated TB is more often due to reactivation of LTBI rather than recent infection. The contacts of HIV-positive TB patients could be a useful target population for HIV testing, as the prevalence of HIV was higher among contacts of HIV-positive TB cases. More widespread testing for LTBI and preventive therapy among people living with HIV could decrease the incidence of TB occurring due to reactivation of LTBI.

7 Discussion

7.1 Key findings

This thesis aimed to improve our understanding of the epidemiology of TB-HIV co-infections and the role of HIV in TB transmission in England, Wales and Northern Ireland. This was addressed with the following objectives:

- To conduct a review of the literature for risk factors for TB among people with HIV in countries with low TB incidence (Chapter 2).
- To link the national TB and HIV surveillance datasets held by Public Health England to identify co-infected patients (Chapter 3).
- To describe the epidemiology of TB-HIV co-infections in England, Wales, and Northern Ireland, and investigate risk factors for developing TB for people diagnosed with HIV (Chapter 4).
- To examine trends in HIV co-infection among TB cases in England, Wales and Northern Ireland, and identify factors associated with HIV co-infection of TB cases (Chapter 5).
- To investigate the influence of HIV on TB transmission by examining associations between HIV co-infection of TB cases and the number of subsequent clustered cases, and determining whether HIV co-infection is associated with reactivation of LTBI or recent infection with TB (Chapter 6).

These objectives were successfully achieved, with some limitations – as described in the relevant chapters. The results and implications from these studies are described below.

The results chapters of this thesis commenced with a review of the literature on risk factors for latent and active TB among PLHIV in countries with low TB incidence in Chapter 2. This review found that sociodemographic risk factors for LTBI and active TB among PLHIV were similar, but that there were differences in clinical risk factors. There were higher rates of both LTBI and active TB among PLHIV of black African, Asian and

Hispanic ethnicities than people of white ethnicity, and higher rates among PLHIV born outside of low-incidence countries, particularly those born in countries with high TB incidence. People who had acquired HIV through injecting drug use or heterosexual sex had higher rates of both LTBI and active TB than MSM. There was some evidence that social risk factors such as drug misuse, homelessness, and imprisonment were associated with TB. Exposure to a contact with TB and having had TB previously were both associated with higher risk of LTBI, but the studies of active TB did not report on these risk factors. There was strong evidence that low CD4 count, high viral load and being ART-naïve increased the risk of active TB for PLHIV; however, there was generally no association between these factors and the risk of LTBI. This may be because diagnostic tests for LTBI are less sensitive in immunocompromised people; or because immunosuppression does not increase the risk of infection, just the risk of progression from LTBI to active TB disease.

In Chapter 3, a new method was developed to link the TB and HIV surveillance datasets for England, Wales and Northern Ireland, which was relatively quick, repeatable, and had no subjective bias. The linked datasets were then used to analyse risk factors associated with TB among PLHIV and risk factors for HIV co-infection among TB patients in England, Wales and Northern Ireland in Chapters 4 and 5, respectively.

In England, Wales and Northern Ireland, six percent of PLHIV diagnosed with HIV between 2000 and 2014 were found to have had a TB diagnosis, and over half of these were not aware of their HIV infection prior to their TB diagnosis (Chapter 4). Patients diagnosed simultaneously with TB and HIV were often born in countries with high TB incidence and acquired HIV through heterosexual sex, both factors which were also associated with development of TB after an HIV diagnosis. Other factors associated with developing TB for PLHIV included HIV acquisition by injecting drug use, low CD4 count and not having initiated ART.

In chapter 5, trend analyses showed that the number and proportion of persons diagnosed with TB who were co-infected with HIV declined from 2005 to 2014; likely a result of an increase in the number of TB patients from south-east Asian countries, where the prevalence of HIV is low, and a decrease in both the number and proportion of TB patients with HIV from sub-Saharan Africa, due to decreases in migration from this area and declining rates of HIV co-infection among TB patients in these countries.

Within the overall decline in the number of TB patients with HIV over the study period, there was a shift in the pattern of the timing of TB and HIV diagnoses with an increasing proportion of co-infected patients being aware of their HIV infection at the time of their TB diagnosis, reflecting improved efforts to diagnose HIV early and a reduction in PLHIV who were only diagnosed with HIV as a result of their TB diagnosis. The risk of HIV co-infection was higher for women and for persons born in countries with high HIV prevalence (particularly people of black African ethnicity). HIV co-infection was more common among patients with miliary or meningeal TB than pulmonary TB, but less common among patients with other forms of extra-pulmonary TB only. In a sub-analysis of five years of data (2010-2014) for which data on social risk factors was available, drug misuse was the only social risk factor independently associated with HIV infection, although co-infection was more common in patients with any social risk factor, and there was substantial overlap between risk factors.

Chapter 6 examined the role of HIV in the transmission of TB at the population-level in England, Wales and Northern Ireland. Patients with pulmonary TB and HIV had fewer subsequent clustered cases than pulmonary TB patients without HIV. However, when patients with HIV and extra-pulmonary TB were the first case of a cluster, they had a higher number of subsequent clustered cases. As it is unlikely that these patients are actually the first in the transmission chain, it is possible that the increased cluster size is a result of more extensive TB screening among their contacts in order to find the source case. Furthermore, HIV prevalence was higher among the subsequent cases of HIV-positive first cases than the subsequent cases of HIV-negative first cases. The

higher number of subsequent clustered cases for some HIV-positive TB cases may therefore be because their contacts are more susceptible, rather than because they are more infectious. TB cases with HIV were less likely to be a subsequent case in a cluster, suggesting they are more likely the result of reactivation of LTBI than recent infection.

7.2 Implications and recommendations

A number of factors have contributed to the observed decrease in the number of patients co-infected with TB and HIV. The demographics of TB patients in the UK have changed; a decrease in migration from sub-Saharan Africa has led to fewer TB patients from this area,[52, 247] and the decreasing proportion who are co-infected with HIV reflects decreasing rates of HIV co-infection among TB patients in sub-Saharan African countries.[1] There has also been an increase in the number of TB patients from south-east Asian countries where the prevalence of HIV is low,[52] which also contributed to the decline in the proportion of TB cases with HIV co-infection in the UK. The decline can also partly be attributed to the improved health of PLHIV in the UK. There have been increases in HIV testing,[223] which led to earlier diagnosis of HIV (evidenced by an increase in the average CD4 count at HIV diagnosis) and earlier initiation of ART;[222] with the overall result being less immunosuppression among PLHIV.

Within the decline of the number of patients co-infected with TB and HIV from 2004 to 2014, there was a shift in the pattern of diagnoses with an increasing proportion of patients being aware of their HIV infection prior to their TB diagnosis. The decline in the number of co-infected patients who were diagnosed with TB more than 91 days before their HIV diagnosis is a result of widespread HIV testing in TB clinics. HIV testing of all new TB patients has been recommended by BHIVA since 2008,[220] but has increased in recent years. In 2012, national data on HIV testing was available for 64% of TB cases with previously unknown HIV status, of whom 87% were tested for HIV; by 2015 data on HIV testing was reported for 92% of TB cases with previously

unknown HIV status and 93% had received HIV testing.[52, 124] Despite the decline in the number of co-infected patients, the substantial proportion of patients who are diagnosed with TB and HIV simultaneously demonstrate the continuing importance of testing new TB patients for HIV. Over the duration of the study period, BHIVA increased the CD4 threshold at which initiation of ART is recommended, which led to a healthier population of PLHIV; it is likely that this contributed to the decrease in the number of patients who developed TB after their HIV diagnosis. The recent (2015) recommendation that ART be initiated regardless of CD4 count should continue this trend and further decrease TB in PLHIV.

- **Recommendation:** ART should continue to be offered to PLHIV regardless of their CD4 count to reduce the incidence of TB.

Whilst the number of people with TB and HIV in the UK is relatively small and is declining, TB-HIV co-infection remains an important clinical problem for those who are affected by it. HIV co-infection complicates diagnosis of TB, due to decreased sensitivity of diagnostic tests for LTBI among PLHIV,[103, 104] lower rates of sputum smear-positivity among PLHIV with TB,[105] and atypical presentation of TB through higher rates of miliary and other extra-pulmonary TB which are difficult to diagnose.[110, 248] HIV co-infection also complicates the management of TB patients due to drug interactions, IRIS, and higher rates of relapse and acquired drug resistance on some TB treatment regimens.[110] TB also substantially increases the risk of death for PLHIV; a recent study found an almost five-fold increased hazard of death among PLHIV with TB than those without.[98]

The increasing proportion of patients who are diagnosed with HIV first, and TB subsequently, suggests that there are missed opportunities to prevent TB from developing in people with diagnosed HIV. This is supported by the data presented in Chapter 6 which showed that TB cases with HIV were less likely to be later cases in strain type clusters and more likely to be the first case or a non-clustered case, suggesting that HIV-associated TB is more often the result of reactivation of remotely-

acquired LTBI than recent infection. BHIVA and NICE both recommend screening for and treating LTBI among PLHIV; the guidelines on eligibility for LTBI testing are based on combinations of TB incidence in the person's country of birth, CD4 count and whether the patient has initiated ART.[110, 249] However, a recent survey of healthcare providers found that LTBI screening of PLHIV was only offered in 57% of geographical areas across the UK; and of those, few complied fully with the BHIVA and NICE guidelines (33.5% and 6.5% respectively).[208] Greater promotion of testing for and treating LTBI among PLHIV could prevent reactivation of LTBI from occurring and reduce the incidence of HIV-associated TB. Further research is required to understand whether the barriers to testing and treating LTBI are occurring at the clinic level (i.e., are testing and treatment for LTBI being offered) or at the patient level, and how these barriers can be overcome to improve the uptake of LTBI testing and treatment.

- **Research question:** Is low uptake of LTBI testing and treatment a result of it not being offered, not being taken up, or treatment not being completed?
- **Research question:** What interventions could be used to improve the uptake of LTBI testing and treatment, and are they cost effective for PLHIV?

Targeting testing and treatment of LTBI to high-risk populations of PLHIV is essential to make best use of resources. The risk factor analyses presented in Chapters 4 and 5 highlighted that people of black African ethnicity or who were born in countries with high levels of TB and HIV have substantially higher rates of co-infection, as do PLHIV who were diagnosed late, those with low CD4 counts, and those who have not initiated ART. Increasing HIV testing among people born in high-incidence countries, people of black African ethnicity, and PWID is essential for earlier diagnosis of HIV. The investigation of risk factors for co-infection was limited to persons who were diagnosed and reported to the national TB and HIV surveillance systems (HARS and ETS). Therefore, as well as the populations that were identified as high-risk for TB-HIV co-infection, further work could examine whether there are other sociodemographic factors associated with late HIV diagnosis and development of TB

which could be used to develop more targeted screening programs. In particular, the literature review showed some evidence that imprisonment and homelessness were risk factors for developing TB for PLHIV. In Chapter 5, both of these factors were associated with high proportions of HIV co-infection in univariable analyses, although not independently. The review of the literature in Chapter 2 found no studies which prospectively examined homelessness or imprisonment as risk factors for active TB; and the studies which examined associations with LTBI had mixed results and were often small and potentially lacking in power. This could be a useful area for further research and targeted screening.

- **Research question:** What is the incidence of TB among PLHIV who are homeless or in prison?
- **Research question:** Is screening people with social risk factors for HIV and LTBI an effective and cost-effective method of active case-finding?
- **Research question:** Are other sociodemographic factors, such as smoking status, occupation, income or housing density, associated with higher rates of TB-HIV co-infection?

The current BHIVA and NICE guidelines for testing and treating LTBI are based on TB incidence in country of birth, CD4 count and ethnicity; the results presented in this thesis support these recommendations. However, the incidence of TB among PLHIV who acquired their HIV infection through injecting drug use was comparable to that of black African PLHIV born in countries with high TB incidence, and this was reflected in the high proportion of HIV co-infection among TB patients with a history of drug misuse. Consequently, testing for and treating PWID for LTBI is likely to be of benefit. As drug use was associated with recent infection with *MTBC*, whereas being born abroad was associated with reactivation of LTBI (section 6.3.4), repeated testing of PWID for LTBI may be warranted due to the ongoing risk of new infection. The optimal frequency at which testing should be offered to maximise efficacy and cost-effectiveness is unknown, and could be the subject of future research. There were

also higher rates of HIV co-infection among TB patients with other social risk factors, such as homelessness and imprisonment. Whilst these factors were not independently associated with HIV co-infection, the substantial proportion of overlap in these populations mean that they could be useful target populations for screening for and treating LTBI. The efficacy and cost-effectiveness of LTBI screening among these groups, and the frequency at which it should be conducted, should be evaluated.

- **Recommendation:** BHIVA guidelines for offering LTBI testing and treatment should be extended to include PWID.
- **Research question:** At what frequency should PWID be offered LTBI testing and treatment to maximise efficacy and cost-effectiveness of the screening programme?

Pulmonary TB cases with HIV co-infection had fewer subsequent clustered cases than pulmonary TB cases without HIV, contributing to the body of evidence that TB patients with HIV are less infectious. This, in combination with the fact that TB patients with HIV are more likely to be the result of reactivation of LTBI than recent infection, suggests that HIV is not a substantial contributor to TB transmission in England, Wales and Northern Ireland. PHE guidelines on when to conduct cluster investigations suggest that a case within a cluster with HIV may be an indicator of recent transmission.[237] The results from this study, which used national surveillance data over a five year period, suggested that HIV co-infection does not increase infectiousness and was more frequently associated with reactivation of LTBI than recent infection. However, as PLHIV have higher rates of progression from LTBI to active TB, the absolute risk of TB as a result of recent infection is likely higher for PLHIV than for HIV-negative people, and consequently clusters of TB involving PLHIV may still merit investigation.

In contrast with pulmonary TB cases, extra-pulmonary TB cases with HIV had a higher number of subsequent cases than HIV-negative extra-pulmonary TB cases. TB cases with only extra-pulmonary disease are very unlikely to be the source case of a

cluster, but investigating these cases when they first present with symptoms could lead to earlier diagnosis of other patients within the cluster, thereby improving treatment outcomes and reducing further transmission from other cases within the cluster. Furthermore, HIV was more prevalent among the subsequent cases of HIV-positive first cases than the subsequent cases of HIV-negative first cases. As a substantial number of co-infected patients were unaware of their HIV infection until their TB diagnosis, targeting HIV testing towards the contacts of co-infected patients could be a useful method for identifying undiagnosed HIV infections in people at risk of developing TB. Diagnosing these people with HIV sooner, initiating ART and testing for and treating LTBI could further reduce TB incidence. The potential yield and cost-effectiveness of such a screening program could be investigated.

- **Research question:** Is offering HIV testing to the contacts of TB patients with HIV a cost-effective method of identifying undiagnosed HIV infections in people at high risk of TB?

7.3 Conclusions

The linkage of two comprehensive national surveillance datasets for TB and HIV allowed epidemiological investigations into TB-HIV co-infection on a national scale. There was a substantial decline in TB-HIV co-infection in England, Wales and Northern Ireland from 2005 to 2014. The number and proportion of co-infected patients who were not aware of their HIV infection at the time of their TB diagnosis has decreased, however a substantial proportion of co-infected patients are still diagnosed simultaneously with TB and HIV, and this could be further reduced. More screening for HIV in populations at risk of both TB and HIV – particularly people of black African ethnicity, people born in high-incidence countries, and people who inject drugs, but potentially also people with a history of homelessness or imprisonment – would lead to earlier diagnosis of HIV and enable TB to be prevented by initiating ART. Among people with diagnosed HIV, ART was highly protective against TB, however many of the people who developed TB in this study had not initiated ART when diagnosed with

TB. Linkage to HIV care and ART initiation have improved over the course of this study and continue to improve, which is likely to further decrease TB among PLHIV. TB patients with HIV appear to be less infectious than HIV-negative TB patients, and HIV co-infection of TB cases was associated with reactivation of LTBI rather than recent infection. Consequently, there is an opportunity to prevent these TB cases by screening for and treating LTBI. Improving adherence to the BHIVA and NICE guidelines on LTBI screening and treatment – and expanding these guidelines to include people who inject drugs – could prevent reactivation of LTBI in PLHIV, thereby reducing the incidence of TB.

8 References

1. WHO. **Global tuberculosis report.** Geneva, Switzerland: World Health Organization, 2016.

2. WHO. The top 10 causes of death 2017. [Accessed 21/09/2017]. Available from: http://www.who.int/mediacentre/factsheets/fs310/en/.

3. WHO. HIV/AIDS 2017. [Accessed 21/08/2016]. Available from: http://www.who.int/mediacentre/factsheets/fs360/en/.

4. McCune JM. **The dynamics of CD4+ T-cell depletion in HIV disease**. Nature. 2001;**410**(6831):974-9.

5. CDC. About HIV/AIDS 2017. [Accessed 21/08/2017]. Available from: https://www.cdc.gov/hiv/basics/whatishiv.html.

6. Doitsh G, Galloway NL, Geng X, Yang Z, Monroe KM, Zepeda O, et al. **Cell death by pyroptosis drives CD4 T-cell depletion in HIV-1 infection**. Nature. 2014;**505**(7484):509-14.

7. Barre-Sinoussi F, Ross AL, Delfraissy JF. **Past, present and future: 30 years of HIV research**. Nature reviews Microbiology. 2013;**11**(12):877-83.

8. AIDSmap. Seroconversion 2017. [Accessed 21/08/2017]. Available from: http://www.aidsmap.com/Seroconversion/page/1322973/s.

9. Fanales-Belasio E, Raimondo M, Suligoi B, Butto S. **HIV virology and pathogenetic mechanisms of infection: a brief overview**. Annali dell'Istituto superiore di sanita. 2010;**46**(1):5-14.

10. Kirwan PD, Chau C, Brown AE, Gill ON, Delpech VC, Contributors. **HIV in the UK - 2016 report**. Public Health England, London. 2016.

11. Terrence Higgins Trust. Testing for HIV 2015. [Accessed 21/08/2017]. Available from: http://www.tht.org.uk/sexual-health/about-hiv/testing-for-hiv.

12. CDC. Act against AIDS: Get tested 2017. [Accessed 23/09/2017]. Available from: https://www.cdc.gov/actagainstaids/campaigns/doingit/gettested/index.html.

13. CDC. Terms, Definitions, and Calculations Used in CDC HIV Surveillance Publications 2017. [Accessed 21/08/2017]. Available from: https://www.cdc.gov/hiv/statistics/surveillance/terms.html.

14. WHO. **WHO case definitions of HIV for surveillance and revised clinical staging and immunological classification of HIV-related disease in adults and children.** Geneva, Switzerland.: World Health Organization, 2007.

15. Strategies for Management of Antiretroviral Therapy Study G, El-Sadr WM, Lundgren J, Neaton JD, Gordin F, Abrams D, et al. **CD4+ count-guided interruption of antiretroviral treatment**. The New England journal of medicine. 2006;**355**(22):2283-96.

16. Insight Start Study Group, Lundgren JD, Babiker AG, Gordin F, Emery S, Grund B, et al. **Initiation of Antiretroviral Therapy in Early Asymptomatic HIV Infection**. The New England journal of medicine. 2015;**373**(9):795-807.

17. Grinsztejn B, Hosseinipour MC, Ribaudo HJ, Swindells S, Eron J, Chen YQ, et al. **Effects of early versus delayed initiation of antiretroviral treatment on clinical outcomes of HIV-1 infection: results from the phase 3 HPTN 052 randomised controlled trial**. The Lancet Infectious diseases. 2014;**14**(4):281-90.

18. PARTNER study group, Rodger AJ, Cambiano V, Bruun T, Vernazza P, Collins S, et al. **Sexual activity without condoms and risk of HIV transmission in serodifferent couples when the HIV-positive partner is using suppressive antiretroviral therapy**. Jama-J Am Med Assoc. 2016;**316**(19):2048-.

19. Fisher M, Benn P, Evans B, Pozniak A, Jones M, Maclean S, et al. **UK Guideline for the use of post-exposure prophylaxis for HIV following sexual exposure**. Int J STD AIDS. 2006;**17**(2):81-92.

20. Young TN, Arens FJ, Kennedy GE, Laurie JW, Rutherford G. **Antiretroviral post-exposure prophylaxis (PEP) for occupational HIV exposure**. The Cochrane database of systematic reviews. 2007(1):CD002835.

21. National Institute for Health and Care Excellence. Pre-exposure prophylaxis of HIV in adults at high risk: Truvada (emtricitabine/tenofovir disoproxil) 2016. [Accessed 21/08/2017]. Available from: https://www.nice.org.uk/advice/esnm78/chapter/key-points-from-the-evidence.

22. Patel P, Borkowf CB, Brooks JT, Lasry A, Lansky A, Mermin J. **Estimating per-act HIV transmission risk: a systematic review**. AIDS. 2014;**28**(10):1509-19.

23. Boily MC, Baggaley RF, Wang L, Masse B, White RG, Hayes RJ, et al. **Heterosexual risk of HIV-1 infection per sexual act: systematic review and meta-analysis of observational studies**. The Lancet Infectious diseases. 2009;**9**(2):118-29.

24. Dosekun O, Fox J. **An overview of the relative risks of different sexual behaviours on HIV transmission**. Current opinion in HIV and AIDS. 2010;**5**(4):291-7.

25. CDC. HIV transmission 2017. [Accessed 21/08/2017]. Available from: https://www.cdc.gov/hiv/basics/transmission.html.

26. UNAIDS. Global HIV statistics 2017. [Accessed 21/08/2017]. Available from: http://www.unaids.org/sites/default/files/media_asset/UNAIDS_FactSheet_en.pdf.

27. Autran B, Carcelain G, Li TS, Blanc C, Mathez D, Tubiana R, et al. **Positive effects of combined antiretroviral therapy on CD4+ T cell homeostasis and function in advanced HIV disease**. Science. 1997;**277**(5322):112-6.

28. Hogg RS, O'Shaughnessy MV, Gataric N, Yip B, Craib K, Schechter MT, et al. **Decline in deaths from AIDS due to new antiretrovirals**. Lancet (London, England). 1997;**349**(9061):1294.

29. AVERT. HIV prevention programmes overview 2017. [Accessed 21/09/2017]. Available from: https://www.avert.org/professionals/hiv-programming/prevention/overview.

30. Temah CT. **What drives HIV/AIDS epidemic in sub-Saharan Africa?** Revue d'economie du development 2009/5. 2009;**17**(5):41-70.

31. Nakagawa F, Miners A, Smith CJ, Simmons R, Lodwick RK, Cambiano V, et al. **Projected Lifetime Healthcare Costs Associated with HIV Infection**. PloS one. 2015;**10**(4):e0125018.

32. Palella FJ, Jr., Baker RK, Moorman AC, Chmiel JS, Wood KC, Brooks JT, et al. **Mortality in the highly active antiretroviral therapy era: changing causes of death and disease in the HIV outpatient study**. Journal of acquired immune deficiency syndromes (1999). 2006;**43**(1):27-34.

33. Flynn JL, Chan J. **Tuberculosis: latency and reactivation**. Infection and immunity. 2001;**69**(7):4195-201.

34. BMJ Best Practice. Extrapulmonary tuberculosis 2017. [Accessed 19/08/2017]. Available from: http://bestpractice.bmj.com/best-practice/monograph/166/diagnosis/step-by-step.html.

35. Underwood J, Cresswell F, Salam AP, Keeley AJ, Cleland C, John L, et al. **Complications of miliary tuberculosis: low mortality and predictive biomarkers from a UK cohort**. BMC infectious diseases. 2017;**17**(1):295.

36. Thwaites GE, Tran TH. **Tuberculous meningitis: many questions, too few answers**. The Lancet Neurology. 2005;**4**(3):160-70.

37. CDC. Latent tuberculosis infection: A guide for primary healthcare providers. Diagnosis of latent TB infection 2014. [Accessed 19/08/2017]. Available from: https://www.cdc.gov/tb/publications/ltbi/diagnosis.htm.

38. CDC. Tuberculin skin testing 2016. [Accessed 19/08/2017]. Available from: https://www.cdc.gov/tb/publications/factsheets/testing/skintesting.htm.

39. Green Book. Tuberculosis 2011. [Accessed 23/09/2017]. Available from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/148511/Green-Book-Chapter-32-dh_128356.pdf.

40. Trajman A, Steffen RE, Menzies D. **Interferon-Gamma Release Assays versus Tuberculin Skin Testing for the Diagnosis of Latent Tuberculosis Infection: An Overview of the Evidence**. Pulmonary medicine. 2013;**2013**:601737.

41. CDC. Interferon-gamma release assays (IGRAs) - blood test for TB infection 2015. [Accessed 19/08/2017]. Available from: https://www.cdc.gov/tb/publications/factsheets/testing/igra.htm.

42. WHO. **Use of tuberculosis interferon-gamma release assays (IGRAs) in low- and middle-income countries: Policy statement.** Geneva, Switzerland: World Health Organization, 2011.

43. Aggerbeck H, Giemza R, Joshi P, Tingskov PN, Hoff ST, Boyle J, et al. **Randomised clinical trial investigating the specificity of a novel skin test (C-Tb) for diagnosis of M. tuberculosis infection**. PloS one. 2013;**8**(5):e64215.

44. Hoff ST, Peter JG, Theron G, Pascoe M, Tingskov PN, Aggerbeck H, et al. **Sensitivity of C-Tb: a novel RD-1-specific skin test for the diagnosis of tuberculosis infection**. The European respiratory journal. 2016;**47**(3):919-28.

45. WHO. Tuberculosis Fact Sheet 2015. Available from: http://www.who.int/mediacentre/factsheets/fs104/en/.

46. Cepheid. Xpert® MTB/RIF 2017. [Accessed 21/08/2017]. Available from: http://www.cepheid.com/us/cepheid-solutions/clinical-ivd-tests/critical-infectious-diseases/xpert-mtb-rif.

47. Kanabus A. Genexpert Test – TB diagnosis and resistance testing: Global Health Education; 2016. [Accessed 21/08/2017]. Available from: https://www.tbfacts.org/xpert-tb-test/.

48. WHO. Latent tuberculosis infection (LTBI) 2017. [Accessed 23/09/2017]. Available from: http://www.who.int/tb/challenges/ltbi/en/.

49. WHO. **Guidelines on the management of latent tuberculosis infection.** Geneva, Switzerland: World Health Organization, 2015.

50. National Institute for Health and Care Excellence. Tuberculosis 2016. [Accessed 21/08/2017]. Available from: https://www.nice.org.uk/guidance/ng33/chapter/Recommendations#active-tb.

51. WHO. Guidelines for treatment of drug-susceptible tuberculosis and patient care 2017. [Accessed 23/09/2017]. Available from: http://www.who.int/tb/publications/2017/dstb_guidance_2017/en/.

52. Public Health England. **Tuberculosis in England: 2016 report.** Public Health England, London: 2016.

53. Johnstone-Robertson S, Lawn SD, Welte A, Bekker LG, Wood R. **Tuberculosis in a South African prison - a transmission modelling analysis**. South African medical journal = Suid-Afrikaanse tydskrif vir geneeskunde. 2011;**101**(11):809-13.

54. Basu S, Andrews JR, Poolman EM, Gandhi NR, Shah NS, Moll A, et al. **Prevention of nosocomial transmission of extensively drug-resistant tuberculosis in rural South African district hospitals: an epidemiological modelling study**. Lancet (London, England). 2007;**370**(9597):1500-7.

55. Escombe AR, Oeser C, Gilman RH, Navincopa M, Ticona E, Martinez C, et al. **The detection of airborne transmission of tuberculosis from HIV-infected patients, using an in vivo air sampling model**. Clin Infect Dis. 2007;**44**(10):1349-57.

56. Escombe AR, Moore DA, Gilman RH, Pan W, Navincopa M, Ticona E, et al. **The infectiousness of tuberculosis patients coinfected with HIV**. PLoS medicine. 2008;**5**(9):e188.

57. Houk VN. **Spread of tuberculosis via recirculated air in a naval vessel: the Byrd study**. Annals of the New York Academy of Sciences. 1980;**353**:10-24.

58. CDC. Infection Control in Health Care Settings 2016. [Accessed 25/08/2017]. Available from: https://www.cdc.gov/tb/topic/infectioncontrol/default.htm.

59. Escombe AR, Moore DA, Gilman RH, Navincopa M, Ticona E, Mitchell B, et al. **Upper-room ultraviolet light and negative air ionization to prevent tuberculosis transmission**. PLoS medicine. 2009;**6**(3):e43.

60. Glynn JR, Vynnycky E, Fine PE. **Influence of sampling on estimates of clustering and recent transmission of Mycobacterium tuberculosis derived from DNA fingerprinting techniques**. Am J Epidemiol. 1999;**149**(4):366-71.

61. Murray M. **Sampling bias in the molecular epidemiology of tuberculosis**. Emerg Infect Dis. 2002;**8**(4):363-9.

62. Borgdorff MW, van den Hof S, Kalisvaart N, Kremer K, van Soolingen D. **Influence of sampling on clustering and associations with risk factors in the molecular epidemiology of tuberculosis**. Am J Epidemiol. 2011;**174**(2):243-51.

63. van Embden JD, Cave MD, Crawford JT, Dale JW, Eisenach KD, Gicquel B, et al. **Strain identification of Mycobacterium tuberculosis by DNA fingerprinting: recommendations for a standardized methodology**. J Clin Microbiol. 1993;**31**(2):406-9.

64. Supply P, Allix C, Lesjean S, Cardoso-Oelemann M, Rusch-Gerdes S, Willery E, et al. **Proposal for standardization of optimized mycobacterial interspersed repetitive unit-variable-number tandem repeat typing of Mycobacterium tuberculosis**. J Clin Microbiol. 2006;**44**(12):4498-510.

65. Oelemann MC, Diel R, Vatin V, Haas W, Rusch-Gerdes S, Locht C, et al. **Assessment of an optimized mycobacterial interspersed repetitive- unit-variable-number tandem-repeat typing system combined with spoligotyping for population-based molecular epidemiology studies of tuberculosis**. J Clin Microbiol. 2007;**45**(3):691-7.

66. Roetzer A, Diel R, Kohl TA, Ruckert C, Nubel U, Blom J, et al. **Whole genome sequencing versus traditional genotyping for investigation of a Mycobacterium tuberculosis outbreak: a longitudinal molecular epidemiological study**. PLoS medicine. 2013;**10**(2):e1001387.

67. Hatherell HA, Colijn C, Stagg HR, Jackson C, Winter JR, Abubakar I. **Interpreting whole genome sequencing for investigating tuberculosis transmission: a systematic review**. BMC medicine. 2016;**14**:21.

68. Walker TM, Ip CL, Harrell RH, Evans JT, Kapatai G, Dedicoat MJ, et al. **Whole-genome sequencing to delineate Mycobacterium tuberculosis outbreaks: a retrospective observational study**. The Lancet Infectious diseases. 2013;**13**(2):137-46.

69. Gardy JL, Johnston JC, Ho Sui SJ, Cook VJ, Shah L, Brodkin E, et al. **Whole-genome sequencing and social-network analysis of a tuberculosis outbreak**. The New England journal of medicine. 2011;**364**(8):730-9.

70. Maguire H, Dale JW, McHugh TD, Butcher PD, Gillespie SH, Costetsos A, et al. **Molecular epidemiology of tuberculosis in London 1995-7 showing low rate of active transmission**. Thorax. 2002;**57**(7):617-22.

71. Vluggen C, Soetaert K, Groenen G, Wanlin M, Spitaels M, Arrazola de Onate W, et al. **Molecular epidemiology of Mycobacterium tuberculosis complex in Brussels, 2010-2013**. PloS one. 2017;**12**(2):e0172554.

72. Kik SV, Verver S, van Soolingen D, de Haas PE, Cobelens FG, Kremer K, et al. **Tuberculosis outbreaks predicted by characteristics of first patients in a DNA fingerprint cluster**. Am J Respir Crit Care Med. 2008;**178**(1):96-104.

73. Houben RM, Dodd PJ. **The Global Burden of Latent Tuberculosis Infection: A Re-estimation Using Mathematical Modelling**. PLoS medicine. 2016;**13**(10):e1002152.

74. WHO Global Health Observatory. Prevalence of HIV among adults aged 15 to 49. Estimates by country (2000-2014) 2016. [Accessed 28/09/2015]. Available from: http://apps.who.int/gho/data/view.main.22500?lang=en.

75. WHO. TB comorbidities and risk factors 2017. [Accessed 05/10/2017]. Available from: http://www.who.int/tb/areas-of-work/treatment/risk-factors/en/.

76. WHO. Tuberculosis (TB): People who use drugs 2017. [Accessed 05/10/2017]. Available from: http://www.who.int/tb/areas-of-work/tb-hiv/substance-abuse/en/.

77. WHO. Tuberculosis (TB): Addressing the needs of vulnerable populations 2017. [Accessed 05/10/2017]. Available from: http://www.who.int/tb/areas-of-work/population-groups/en/.

78. McCormick A. **The notification of infectious diseases in England and Wales**. Communicable Disease Report Review. 1993;**3**(2).

79. Communicable Disease Surveillance Centre. **Tuberculosis in Wales Annual Report 2016: Data to the end of 2015.** Public Health Wales, 2016.

80. Health Protection Scotl. **Enhanced Surveillance of Mycobacterial Infections (ESMI) in Scotland: 2016 tuberculosis annual report for Scotland** 2016.

81. Public Health Agency Northern Ireland. Fall in Tuberculosis (TB) cases 2016. [Accessed 13/08/2017]. Available from: http://www.publichealth.hscni.net/news/fall-tuberculosis-tb-cases.

82. Public Health England. **Tuberculosis in England: 2015 report.** Public Health England, London: 2015.

83. Public Health England. **Tuberculosis in the UK: 2014 report.** Public Health England, London: 2014.

84. WHO. Definitions and reporting framework for tuberculosis - 2013 revision (updated December 2014) 2014. [Accessed 23/09/2017]. Available from: http://apps.who.int/iris/bitstream/10665/79199/1/9789241505345_eng.pdf.

85. Sonnenberg P, Glynn JR, Fielding K, Murray J, Godfrey-Faussett P, Shearer S. **How soon after infection with HIV does the risk of tuberculosis start to increase? A retrospective cohort study in South African gold miners**. Journal of Infectious Diseases. 2005;**191**(2):150-8.

86. Selwyn PA, Hartel D, Lewis VA, Schoenbaum EE, Vermund SH, Klein RS, et al. **A prospective study of the risk of tuberculosis among intravenous drug users with human immunodeficiency virus infection**. New England Journal of Medicine. 1989;**320**(9):545-50.

87. WHO. **Towards tuberculosis elimination: An action framework for low-incidence countries.** Geneva, Switzerland: World Health Organization, 2014.

88. Kwan CK, Ernst JD. **HIV and tuberculosis: a deadly human syndemic**. Clinical microbiology reviews. 2011;**24**(2):351-76.

89. Corbett EL, Charalambous S, Moloi VM, Fielding K, Grant AD, Dye C, et al. **Human immunodeficiency virus and the prevalence of undiagnosed tuberculosis in African gold miners**. Am J Respir Crit Care Med. 2004;**170**(6):673-9.

90. CDC. **Guidelines for preventing the transmission of *Mycobacterium tuberculosis* in health-care settings, 2005**. Morbidity and Mortality Weekly Report. 2005;**54**(17).

91. Crofts JP, Andrews NJ, Barker RD, Delpech V, Abubakar I. **Risk factors for recurrent tuberculosis in England and Wales, 1998-2005**. Thorax. 2010;**65**(4):310-4.

92. Houben RMGJ, Crampin AC, Ndhlovu R, Sonnenberg P, Godfrey-Faussett R, Haas WH, et al. **Human immunodeficiency virus associated tuberculosis more often due to recent infection than reactivation of latent infection**. International Journal of Tuberculosis and Lung Disease. 2011;**15**(1):24-31.

93. Cruciani M, Malena M, Bosco O, Gatti G, Serpelloni G. **The impact of Human Immunodeficiency Virus type 1 on infectiousness of tuberculosis: A meta-analysis**. Clinical Infectious Diseases. 2001;**33**:1922-30.

94. Carvalho AC, DeRiemer K, Nunes ZB, Martins M, Comelli M, Marinoni A, et al. **Transmission of Mycobacterium tuberculosis to contacts of HIV-infected tuberculosis patients**. Am J Respir Crit Care Med. 2001;**164**(12):2166-71.

95. Espinal MA, Perez EN, Baez J, Henriquez L, Fernandez K, Lopez M, et al. **Infectiousness of Mycobacterium tuberculosis in HIV-1-infected patients with tuberculosis: a prospective study**. Lancet (London, England). 2000;**355**(9200):275-80.

96. Huang CC, Tchetgen ET, Becerra MC, Cohen T, Hughes KC, Zhang ZB, et al. **The Effect of HIV-Related Immunosuppression on the Risk of Tuberculosis Transmission to Household Contacts**. Clinical Infectious Diseases. 2014;**58**(6):765-74.

97. Alpert PL, Munsiff SS, Gourevitch MN, Greenberg B, Klein RS. **A prospective study of tuberculosis and human immunodeficiency virus infection: Clinical manifestations and factors associated with survival**. Clinical Infectious Diseases. 1997;**24**(4):661-8.

98. Zenner D, Abubakar I, Conti S, Gupta RK, Yin Z, Kall M, et al. **Impact of TB on the survival of people living with HIV infection in England, Wales and Northern Ireland**. Thorax. 2015;**70**(6):566-73.

99. Elliott AM, Namaambo K, Allen BW, Luo N, Hayes RJ, Pobee JOM, et al. **Negative sputum smear results in HIV-positive patients with pulmonary tuberculosis in Lusaka, Zambia**. Tubercule and Lung Disease. 1993;**74**:191-4.

100. Yang ZH, Kong Y, Wilson F, Foxman B, Fowler AH, Marrs CF, et al. **Identification of risk factors for extrapulmonary tuberculosis**. Clinical Infectious Diseases. 2004;**38**(2):199-205.

101. Aderaye G, Bruchfeld J, Assefa G, Feleke D, Kallenius G, Baat M, et al. **The relationship between disease pattern and disease burden by chest radiography, M-tuberculosis load, and HIV status in patients with pulmonary tuberculosis in Addis Ababa**. Infection. 2004;**32**(6):333-8.

102. Rathman G, Sillah J, Hill PC, Murray JF, Adegbola R, Corrah T, et al. **Clinical and radiological presentation of 340 adults with smear-positive tuberculosis in The Gambia**. International Journal of Tuberculosis and Lung Disease. 2003;**7**(10):942-7.

103. Rangaka MX, Diwakar L, Seldon R, van Cutsem G, Meintjes GA, Morroni C, et al. **Clinical, immunological, and epidemiological importance of antituberculosis T cell responses in HIV-infected Africans**. Clin Infect Dis. 2007;**44**(12):1639-46.

104. Brock I, Ruhwald M, Lundgren B, Westh H, Mathiesen LR, Ravn P. **Latent tuberculosis in HIV positive, diagnosed by the M. tuberculosis specific interferon-gamma test**. Respiratory research. 2006;**7 (no pagination)**(56).

105. TB Alert. TB and HIV 2017. [Accessed 24/08/2017]. Available from: https://www.tbalert.org/about-tb/global-tb-challenges/tb-hiv/.

106. Sonnenberg P, Murray J, Glynn JR, Shearer S, Kambashi B, Godfrey-Faussett P. **HIV-1 and recurrence, relapse, and reinfection of tuberculosis after cure: a cohort study in South African mineworkers**. Lancet (London, England). 2001;**358**(9294):1687-93.

107. Ackah AN, Coulibaly D, Digbeu H, Diallo K, Vetter KM, Coulibaly IM, et al. **Response to treatment, mortality, and CD4 lymphocyte counts in HIV-infected persons with tuberculosis in Abidjan, Cote d'Ivoire**. Lancet (London, England). 1995;**345**(8950):607-10.

108. Meintjes G, Lawn SD, Scano F, Maartens G, French MA, Worodria W, et al. **Tuberculosis-associated immune reconstitution inflammatory syndrome: case definitions for use in resource-limited settings**. The Lancet Infectious diseases. 2008;**8**(8):516-23.

109. WHO. **Treatment of tuberculosis guidelines.** Geneva, Switzerland: World Health Organization, 2010.

110. Pozniak AL, Coyne KM, Miller RF, Lipman MCI, Freedman AR, Ormerod LP, et al. **British HIV Association guidelines for the treatment of TB/HIV coinfection 2011**. HIV Medicine. 2011;**12**:517-24.

111. WHO. Tuberculosis: TB and HIV 2017. [Accessed 31/01/2017]. Available from: http://www.who.int/tb/areas-of-work/tb-hiv/en/.

112. Yong Kim J, Shakow A, Castro A, Vande C, Farmer P. Tuberculosis control 2015. Available from: http://www.who.int/trade/distance_learning/gpgh/gpgh3/en/index4.html.

113. Wood R, Maartens G, Lombard CJ. **Risk factors for developing tuberculosis in HIV-1-infected adults from communities with a low or very high incidence of tuberculosis**. J Acq Immun Def Synd. 2000;**23**(1):75-80.

114. Zenner D, Conti S, Yin Z, Kall M, Kruijshaar M, Rice B, et al. **TB co-infection is associated with poor survival among HIV infected patients in England and Wales**. Thorax. 2013;**68**:A32-A.

115. WHO. HIV-Associated Tubeculosis: Fact Sheet 2016. [Accessed 25/07/2017]. Available from: http://www.who.int/tb/areas-of-work/tb-hiv/tbhiv_factsheet_2016_web.pdf?ua=1.

116. Pimpin L, Drumright LN, Kruijshaar ME, Abubakar I, Rice B, Delpech V, et al. **Tuberculosis and HIV co-infection in European Union and European Economic Area countries**. European Respiratory Journal. 2011;**38**(6):1382-92.

117. van der Werf MJ, Kodmon C, Zucs P, Hollo V, Amato-Gauci AJ, Pharris A. **Tuberculosis and HIV coinfection in Europe: looking at one reality from two angles**. AIDS. 2016;**30**(18):2845-53.

118. Karo B, Haas W, Kollan C, Gunsenheimer-Bartmeyer B, Hamouda O, Fiebig L, et al. **Tuberculosis among people living with HIV/AIDS in the German ClinSurv**

HIV Cohort: Long-term incidence and risk factors. BMC Infectious Diseases. 2014;**14 (1) (no pagination)**(148).

119. United Kingdom Collaborative HIV Cohort Study Group, Grant AD, Bansi L, Ainsworth J, Anderson J, Delpech V, et al. **Tuberculosis among people with HIV infection in the United Kingdom: opportunities for prevention?** Aids. 2009;**23**(18):2507-15.

120. WHO. WHO TB burden estimates 2017. [Accessed 23/07/2017]. Available from: http://www.who.int/tb/country/data/download/en/.

121. Podlekareva DN, Efsen AM, Schultze A, Post FA, Skrahina AM, Panteleev A, et al. **Tuberculosis-related mortality in people living with HIV in Europe and Latin America: an international cohort study**. The lancet HIV. 2016;**3**(3):e120-31.

122. Health Protection Agency. **Tuberculosis in the UK: Annual report on tuberculosis surveillance in the UK**. 2010.

123. Ahmed AB, Abubakar I, Delpech V, Lipman M, Boccia D, Forde J, et al. **The growing impact of HIV infection on the epidemiology of tuberculosis in England and Wales: 1999-2003**. Thorax. 2007;**62**(8):672-6.

124. Public Health England. **Tuberculosis in the UK: Annual report on tuberculosis surveillance in the UK**. 2013.

125. Yin Z, Brown AE, Hughes G, Nardone A, Gill ON, Delpech VC, et al. **HIV in the United Kingdom 2014 report: data to end 2013**. Public Health England, London. 2014.

126. Health Protection Agency. **HIV in the United Kingdom: 2011 report**. London: Health Protection Services, Colindale. 2011.

127. WHO. Framework towards TB elimination in low-incidence countries 2014. Available from: http://www.who.int/tb/publications/Towards_TB_Eliminationfactsheet.pdf.

128. Abubakar I, Stagg HR, Cohen T, Mangtani P, Rodrigues LC, Pimpin L, et al. **Controversies and Unresolved Issues in Tuberculosis Prevention and Control: A Low-Burden-Country Perspective**. Journal of Infectious Diseases. 2012;**205**:S293-S300.

129. ECDC. **The status of the HIV response in the European Union/European Economic Area, 2016.** Stockholm, Sweden: European Centre for Disease Surveillance and Control, 2017.

130. WHO. Management of Tuberculosis and HIV Coinfection. Clinical protocol for the WHO European region 2017. [Accessed 15/09/2017]. Available from: http://www.euro.who.int/__data/assets/pdf_file/0004/78124/E90840_Chapter_4.pdf.

131. Bourgarit A, Baron G, Breton G, Tattevin P, Katlama C, Allavena C, et al. **Latent tuberculosis infection screening and 2-year outcome in antiretroviral-naive HIV-infected patients in a low-prevalence country**. Annals of the American Thoracic Society. 2015;**12**(8):1138-45.

132. Cheallaigh CN, Fitzgerald I, Grace J, Singh GJ, El-Eraki N, Gibbons N, et al. **Interferon gamma release assays for the diagnosis of latent TB infection in HIV-infected individuals in a low TB burden country**. PLoS One. 2013;**8**(1):e53330.

133. Doshi S, Chen TF, Zapata J, Holzman RS, Zapata LC, Aberg JA, et al. **Risk factors for tuberculin skin test conversion among HIV-infected patients in New York City**. International journal of infectious diseases : IJID : official publication of the International Society for Infectious Diseases. 2012;**16**(7):e518-21.

134. French AL, Evans CT, Anastos K, Greenblatt RM, Hershow R, Huebner R, et al. **Incidence of tuberculin skin test conversion among HIV-infected and -uninfected women: results of a 6-year study**. J Acquir Immune Defic Syndr. 2006;**42**(5):592-6.

135. Gampper SN, George JA, Carter EJ, Jesdale BM, Flanigan TP, Mayer KH, et al. **Co-infection with Mycobacterium tuberculosis and HIV in high risk clinical care setting in Rhode Island**. AIDS care. 1998;**10**(2):221-9.

136. Girardi E, Palmieri F, Zaccarelli M, Tozzi V, Trotta MP, Selva C, et al. **High incidence of tuberculin skin test conversion among HIV-infected individuals who have a favourable immunological response to highly active antiretroviral therapy**. Aids. 2002;**16**(14):1976-9.

137. Kall MM, Coyne KM, Garrett NJ, Boyd AE, Ashcroft AT, Reeves I, et al. **Latent and subclinical tuberculosis in HIV infected patients: a cross-sectional study**. BMC Infectious Diseases. 2012;**12 (no pagination)**(107).

138. Luetkemeyer AF, Charlebois ED, Flores LL, Bangsberg DR, Deeks SG, Martin JN, et al. **Comparison of an interferon-gamma release assay with tuberculin skin testing in HIV-infected individuals**. Am J Respir Crit Care Med. 2007;**175**(7):737-42.

139. Mofenson LM, Rodriguez EM, Hershow R, Fox HE, Landesman S, Tuomala R, et al. **Mycobacterium tuberculosis infection in pregnant and nonpregnant women infected with HIV in the women and infants transmission study**. Archives of Internal Medicine. 1995;**155**(10):1066-72.

140. Pullar ND, Steinum H, Bruun JN, Dyrhol-Riise AM. **HIV patients with latent tuberculosis living in a low-endemic country do not develop active disease during a 2 year follow-up; a Norwegian prospective multicenter study**. BMC Infectious Diseases. 2014;**14 (1) (no pagination)**(667).

141. Schulte JM, Bryan P, Dodds S, Potter M, Onorato IM, O'Sullivan MJ. **Tuberculosis skin testing among HIV-infected pregnant women in Miami, 1995 to 1996**. Journal of Perinatology. 2002;**22**(2):159-62.

142. Stephan C, Wolf T, Goetsch U, Bellinger O, Nisius G, Oremek G, et al. **Comparing QuantiFERON-tuberculosis gold, T-SPOT tuberculosis and tuberculin skin test in HIV-infected individuals from a low prevalence tuberculosis country**. Aids. 2008;**22**(18):2471-9.

143. Aichelburg MC, Rieger A, Breitenecker F, Pfistershammer K, Tittes J, Eltz S, et al. **Detection and prediction of active tuberculosis disease by a whole-blood interferon-gamma release assay in HIV-1-infected individuals**. Clin Infect Dis. 2009;**48**(7):954-62.

144. Elzi L, Schlegel M, Weber R, Hirschel B, Cavassini M, Schmid P, et al. **Reducing tuberculosis incidence by tuberculin skin testing, preventive treatment, and antiretroviral therapy in an area of low tuberculosis transmission**. Clinical Infectious Diseases. 2007;**44**(1):94-102.

145. McLaughlin SI, Spradling P, Drociuk D, Ridzon R, Pozsik CJ, Onorato I. **Extensive transmission of Mycobacterium tuberculosis among congregated, HIV-**

infected prison inmates in South Carolina, United States. International Journal of Tuberculosis and Lung Disease. 2003;**7**(7):665-72.

146. Brassard P, Hottes TS, Lalonde RG, Klein MB. **Tuberculosis screening and active tuberculosis among HIV-infected persons in a Canadian tertiary care centre**. Canadian Journal of Infectious Diseases and Medical Microbiology. 2009;**20**(2):51-7.

147. Abgrall S, del Giudice P, Melica G, Costagliola D, Behalf F-AC. **HIV-associated tuberculosis and immigration in a high-income country: incidence trends and risk factors in recent years**. Aids. 2010;**24**(5):763-71.

148. Antonucci G, Girardi E, Raviglione MC, Ippolito G. **Risk factors for tuberculosis in HIV-infected persons: A prospective cohort study**. Journal of the American Medical Association. 1995;**274**(2):143-8.

149. Dooley SW, Villarino ME, Lawrence M, Salinas L, Amil S, Rullan JV, et al. **Nosocomial transmission of tuberculosis in a hospital unit for HIV- infected patients**. Journal of the American Medical Association. 1992;**267**(19):2632-5.

150. Girardi E, Antonucci G, Vanacore P, Libanore M, Errante I, Matteelli A, et al. **Impact of combination antiretroviral therapy on the risk of tuberculosis among persons with HIV infection**. Aids. 2000;**14**(13):1985-91.

151. Gupta RK, Lipman M, Brown A, Rice B, Pozniak A, Zenner D, et al. **Does long-term ART reduce TB rates to background population levels? Data from a national HIV cohort**. Topics in Antiviral Medicine. 2014;**22**:431-2.

152. Hasse B, Walker AS, Fehr J, Furrer H, Hoffmann M, Battegay M, et al. **Co-trimoxazole prophylaxis is associated with reduced risk of incident tuberculosis in participants in the swiss HIV cohort study**. Antimicrobial Agents and Chemotherapy. 2014;**58**(4):2363-8.

153. Jones JL, Hanson DL, Dworkin MS, DeCock KM, Thompsom M, Grochowski K, et al. **HIV-associated tuberculosis in the era of highly active antiretroviral therapy**. International Journal of Tuberculosis and Lung Disease. 2000;**4**(11):1026-31.

154. Manavi K, Hodson J. **An observational study on the incidence of tuberculosis among a cohort of HIV infected adults in a setting with low prevalence of tuberculosis**. HIV Clinical Trials. 2016;**17**(5):181-8.

155. Maniewski U, Payen MC, Delforge M, De Wit S. **Is systematic screening and treatment for latent tuberculosis infection in HIV patients useful in a low endemic setting?** Acta Clinica Belgica: International Journal of Clinical and Laboratory Medicine. 2016:1-4.

156. Markowitz N, Hansen NI, Hopewell PC, Glassroth J, Kvale PA, Mangura BT, et al. **Incidence of tuberculosis in the United States among HIV-infected persons**. Annals of Internal Medicine. 1997;**126**(2):123-&.

157. Miguez-Burbano MJ, Burbano X, Ashkin D, Pitchenik A, Allan R, Pineda L, et al. **Impact of tobacco use on the development of opportunistic respiratory infections in HIV seropositive patients on antiretroviral therapy**. Addiction biology. 2003;**8**(1):39-43.

158. Mohle-Boetani JC, Miguelino V, Dewsnup DH, Desmond E, Horowitz E, Waterman SH, et al. **Tuberculosis outbreak in a housing unit for human immunodeficiency virus-infected patients in a correctional facility: Transmission**

risk factors and effective outbreak control. Clinical Infectious Diseases. 2002;**34**(5):668-76.

159. Mor Z, Lidji M, Cedar N, Grotto I, Chemtob D. **Tuberculosis incidence in HIV/AIDS patients in Israel, 1983-2010**. PLoS ONE. 2013;**8 (11) (no pagination)**(e79691).

160. Moro ML, Errante I, Infuso A, Sodano L, Gori A, Orcese CA, et al. **Effectiveness of infection control measures in controlling a nosocomial outbreak of multidrug-resistant tuberculosis among HIV patients in Italy**. International Journal of Tuberculosis and Lung Disease. 2000;**4**(1):61-8.

161. Pettit AC, Mendes A, Jenkins C, Napravnik S, Freeman A, Shepherd BE, et al. **Timing of antiretroviral treatment, immunovirologic status, and TB risk: Implications for testing and treatment**. Journal of Acquired Immune Deficiency Syndromes. 2016;**72**(5):572-8.

162. Pettit AC, Jenkins CA, Stinnette SE, Rebeiro PF, Blackwell RB, Raffanti SP, et al. **Tuberculosis risk before and after highly active antiretroviral therapy initiation: Does HAART increase the short-term TB risk in a low incidence TB Setting?** Journal of Acquired Immune Deficiency Syndromes. 2011;**57**(4):305-10.

163. Pullar ND, Steinum H, Tonby K, Heggelund L, Leiva RA, Ofstad R, et al. **Low prevalence of positive interferon-gamma tests in HIV-positive long-term immigrants in Norway**. Int J Tuberc Lung Dis. 2014;**18**(2):180-7.

164. Rice B, Elford J, Yin Z, Kruijshaar M, Abubakar I, Lipman M, et al. **Decreasing incidence of tuberculosis among heterosexuals living with diagnosed HIV in England and Wales**. Aids. 2013;**27**(7):1151-7.

165. Rubinstien EM, Madden GM, Lyons RW. **Active tuberculosis in HIV-infected injecting drug users from a low-rate tuberculosis area**. Journal of Acquired Immune Deficiency Syndromes and Human Retrovirology. 1996;**11**(5):448-54.

166. Sackoff J, Bernard MA, Adams L, Jones JL. **HIV-associated tuberculosis in the era of highly active antiretroviral therapy**. International Journal of Tuberculosis and Lung Disease. 2001;**5**(5):489.

167. Sterling TR, Lau B, Zhang JB, Freeman A, Bosch RJ, Brooks JT, et al. **Risk Factors for Tuberculosis After Highly Active Antiretroviral Therapy Initiation in the United States and Canada: Implications for Tuberculosis Screening**. Journal of Infectious Diseases. 2011;**204**(6):893-901.

168. Sudre P, Hirschel B, Toscani L, Ledergerber B, Rieder HL, Battegay M, et al. **Risk factors for tuberculosis among HIV-infected patients in Switzerland**. European Respiratory Journal. 1996;**9**(2):279-83.

169. Taarnhoj GA, Engsig FN, Ravn P, Johansen IS, Larsen CS, Roge B, et al. **Incidence, risk factors and mortality of tuberculosis in Danish HIV patients 1995-2007**. BMC Pulmonary Medicine. 2011;**11 (no pagination)**(26).

170. Thomas P, Bornschlegel K, Singh TP, Abrams EJ, Cervia J, Fikrig S, et al. **Tuberculosis in human immunodeficiency virus-infected and human immunodeficiency virus-exposed children in New York City**. Pediatric Infectious Disease Journal. 2000;**19**(8):700-6.

171. Trieu L, Li JH, Hanna DB, Harris TG. **Tuberculosis Rates Among HIV-Infected Persons in New York City, 2001-2005**. American Journal of Public Health. 2010;**100**(6):1031-4.

172. Turkova A, Chappell E, Judd A, Goodall RL, Welch SB, Foster C, et al. **Prevalence, incidence, and associated risk factors of tuberculosis in children with HIV living in the UK and Ireland (CHIPS): A cohort study**. The Lancet HIV. 2015;**2**(12):e530-e9.

173. Aichelburg MC, Mandorfer M, Tittes J, Breitenecker F, Reiberger T, Rieger A, et al. **The association of smoking with IGRA and TST results in HIV-1-infected subjects**. Int J Tuberc Lung Dis. 2014;**18**(6):709-16.

174. Taarnhoj GA, Engsig FN, Ravn P, Johansen IS, Larsen CS, Roge B, et al. **Incidence, risk factors and mortality of tuberculosis in Danish HIV patients 1995-2007**. BMC pulmonary medicine. 2011;**11**:26.

175. del Amo J, Moreno S, Bucher HC, Furrer H, Logan R, Sterne J, et al. **Impact of Antiretroviral Therapy on Tuberculosis Incidence Among HIV-Positive Patients in High-Income Countries**. Clinical Infectious Diseases. 2012;**54**(9):1364-72.

176. Kirk O, Gatell JM, Mocroft A, Pedersen C, Proenca R, Brettle RP, et al. **Infections with Mycobacterium tuberculosis and Mycobacterium avium among HIV-infected patients after the introduction of highly active antiretroviral therapy**. American Journal of Respiratory and Critical Care Medicine. 2000;**162**(3 I):865-72.

177. Kruk A, Bannister W, Podlekareva DN, Chentsova NP, Rakhmanova AG, Horban A, et al. **Tuberculosis among HIV-positive patients across Europe: Changes over time and risk factors**. Aids. 2011;**25**(12):1505-13.

178. Brinkhof MWG, Egger M, Boulle A, May M, Hosseinipour M, Sprinz E, et al. **Tuberculosis after initiation of antiretroviral therapy in low-income and high-income countries**. Clinical Infectious Diseases. 2007;**45**(11):1518-21.

179. Nhamoyebonde S, Leslie A. **Biological differences between the sexes and susceptibility to tuberculosis**. Journal of Infectious Diseases. 2014;**209 Suppl 3**:S100-6.

180. UCL-TB. World TB Day Symposium 2017 Review 2017. [Accessed 28/10/17]. Available from: http://www.ucl.ac.uk/tb/WorldTBDay/wtbd-2017-review.

181. Van Asten L, Langendam M, Zangerle R, Aguado IH, Boufassa F, Schiffer V, et al. **Tuberculosis risk varies with the duration of HIV infection: A prospective study of European drug users with known date of HIV seroconversion**. Aids. 2003;**17**(8):1201-8.

182. Zenner D, Beer N, Harris RJ, Lipman MC, Stagg HR, van der Werf MJ. **Treatment of Latent Tuberculosis Infection: An Updated Network Meta-analysis**. Ann Intern Med. 2017;**167**(4):248-55.

183. WHO. Public health surveillance 2017. [Accessed 11/09/2017]. Available from: http://www.who.int/topics/public_health_surveillance/en/.

184. Aldridge RW, Shaji K, Hayward AC, Abubakar I. **Accuracy of Probabilistic Linkage Using the Enhanced Matching System for Public Health and Epidemiological Studies**. PloS one. 2015;**10**(8).

185. Moore HC, de Klerk N, Keil AD, Smith DW, Blyth CC, Richmond P, et al. **Use of data linkage to investigate the aetiology of acute lower respiratory infection hospitalisations in children**. Journal of paediatrics and child health. 2012;**48**(6):520-8.

186. Health & Social Care Information Centre. A Guide to Linked Mortality Data from Hospital Episode Statistics and the Office for National Statistics 2015. [Accessed 14/09/2017]. Available from: http://content.digital.nhs.uk/media/11668/HES-ONS-Mortality-Data-Guide/pdf/hes_ons__%20mortality_data_guide.pdf.

187. Glaziou P SC, Zignol M, Floyd K. **Methods used by WHO to estimate the global burden of TB disease.** Global TB Programme, World Health Organization, Geneva, Switzerland.

188. WHO/UNAIDS. Guidelines for second generation HIV surveillance. An update: know your epidemic 2013. [Accessed 29/09/2017]. Available from: http://www.who.int/hiv/pub/surveillance/2013package/module1/en/.

189. WHO. HIV/AIDS strategic information: Surveillance 2017. [Accessed 29/09/2017]. Available from: http://www.who.int/hiv/strategic/surveillance/en/.

190. ECDC/WHO Regional office for Europe. **HIV/AIDS surveillance in Europe 2015.** ECDC, Stockholm, Sweden.: 2016.

191. Public Health England. HIV surveillance systems 2008. [Accessed 29/09/2017]. Available from: https://www.gov.uk/guidance/hiv-surveillance-systems.

192. Public Health England. Guidance on notifying Tuberculosis (TB) cases 2010. Available from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/360263/Guidance_on_Notifying_Tuberculosis__TB__cases.pdf.

193. Mortimer JY, Salathiel J. **'Soundex' codes of surnames provide confidentiality and accuracy in a national HIV database**. Communicable Disease Report CDR Review. 1995;**5**:R183-R6.

194. van Smeden M, Naaktgeboren CA, Reitsma JB, Moons KG, de Groot JA. **Latent class models in diagnostic studies when there is no reference standard--a systematic review**. Am J Epidemiol. 2014;**179**(4):423-31.

195. AVERT. HIV and AIDS in the United Kingdom 2017. [Accessed 17/09/2017]. Available from: https://www.avert.org/professionals/hiv-around-world/western-central-europe-north-america/uk.

196. Gupta RK, Rice B, Brown AE, Thomas HL, Zenner D, Anderson L, et al. **Does antiretroviral therapy reduce HIV-associated tuberculosis incidence to background rates? A national observational cohort study from England, Wales, and Northern Ireland**. The lancet HIV. 2015;**2**(6):e243-51.

197. Department for Communities and Local Government. English Indices of Deprivation 2015. [Accessed 20/05/2016]. Available from: https://www.gov.uk/government/collections/english-indices-of-deprivation.

198. Welsh Govenment. Welsh Index of Multiple Deprivation (WIMD) 2015. [Accessed 20/05/2016]. Available from: http://gov.wales/statistics-and-research/welsh-index-multiple-deprivation/?lang=en.

199. Victora CG, Huttly SR, Fuchs SC, Olinto MTA. **The role of conceptual frameworks in epidemiological analysis: A hierarchical approach**. International Journal of Epidemiology. 1997;**26**(1):224-7.

200. Suthar AB, Lawn SD, del Amo J, Getahun H, Dye C, Sculier D, et al. **Antiretroviral therapy for prevention of tuberculosis in adults with HIV: a systematic review and meta-analysis**. PLoS medicine. 2012;**9**(7):e1001270.

201. Hue S, Brown AE, Ragonnet-Cronin M, Lycett SJ, Dunn DT, Fearnhill E, et al. **Phylogenetic analyses reveal HIV-1 infections between men misclassified as heterosexual transmissions**. AIDS. 2014;**28**(13):1967-75.

202. Skingsley A, Kirwan P, Yin Z, Nardone A, Hughes G, Tosswill J, et al. **HIV new diagnosies, treatment and care in the UK 2015 report: Data to end 2014.** Public Health England, London: 2015.

203. Public Health England, Health Protection Scotland, Public Health Wales, Public Health Agency Northern Ireland. **Shooting Up: Infections among people who inject drugs in the UK, 2015.** London: Public Health England, 2016 November 2016. Report No.

204. National AIDS Trust. **HIV and injecting drug use.** London: 2013.

205. Rice BD, Elford J, Yin Z, Delpech VC. **A new method to assign country of HIV infection among heterosexuals born abroad and diagnosed with HIV**. Aids. 2012;**26**(15):1961-6.

206. Temprano ANRS Study Group, Danel C, Moh R, Gabillard D, Badje A, Le Carrou J, et al. **A Trial of Early Antiretrovirals and Isoniazid Preventive Therapy in Africa**. The New England journal of medicine. 2015;**373**(9):808-22.

207. Fox-Lewis A, Brima N, Muniina P, Grant AD, Edwards SG, Miller RF, et al. **Tuberculosis screening in patients with HIV: An audit against UK national guidelines to assess current practice and the effectiveness of an electronic tuberculosis-screening prompt**. Int J STD AIDS. 2016;**27**(10):901-5.

208. White HA, Miller RF, Pozniak AL, Lipman MC, Stephenson I, Wiselka MJ, et al. **Latent tuberculosis infection screening and treatment in HIV: insights from evaluation of UK practice**. Thorax. 2017;**72**(2):180-2.

209. Aldridge RW, Zenner D, White PJ, Williamson EJ, Muzyamba MC, Dhavan P, et al. **Tuberculosis in migrants moving from high-incidence to low-incidence countries: a population-based cohort study of 519 955 migrants screened before entry to England, Wales, and Northern Ireland**. Lancet (London, England). 2016;**388**(10059):2510-8.

210. BHIVA. **British HIV Association guidelines for the treatment of HIV-1-positive adults with antiretroviral therapy 2015.** London: 2015.

211. Strategies for Management of Antiretroviral Therapy Study G, Lundgren JD, Babiker A, El-Sadr W, Emery S, Grund B, et al. **Inferior clinical outcome of the CD4+ cell count-guided antiretroviral treatment interruption strategy in the SMART study: role of CD4+ Cell counts and HIV RNA levels during follow-up**. The Journal of infectious diseases. 2008;**197**(8):1145-55.

212. Crofts JP, Kruijshaar ME, Delpech V, Ncube F, Abubakar I. **Tuberculosis and HIV co-infection in healthcare workers in England and Wales, 1999-2005**. Epidemiology and Infection. 2012;**140**(10):1873-9.

213. Haar CH, Cobelens FG, Kalisvaart NA, Van der Have JJ, Van Gerven PJ, Van Deutekom H. **HIV prevalence among tuberculosis patients in The Netherlands, 1993-2001: trends and risk factors**. The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease. 2006;**10**(7):768-74.

214. Vos AM, Meima A, Verver S, Looman CW, Bos V, Borgdorff MW, et al. **High incidence of pulmonary tuberculosis persists a decade after immigration, The Netherlands**. Emerg Infect Dis. 2004;**10**(4):736-9.

215. Kruijshaar ME, Abubakar I. **Increase in extrapulmonary tuberculosis in England and Wales 1999-2006**. Thorax. 2009;**64**(12):1090-5.

216. Naing C, Mak JW, Maung M, Wong SF, Kassim AI. **Meta-analysis: the association between HIV infection and extrapulmonary tuberculosis**. Lung. 2013;**191**(1):27-34.

217. Public Health England. Tuberculosis (TB): collaborative strategy for England 2015. [Accessed 11/07/2017]. Available from: https://www.gov.uk/government/publications/collaborative-tuberculosis-strategy-for-england.

218. NICE. **HIV testing: increasing uptake among people who may have undiagnosed HIV**. 2011.

219. Capocci S, Smith C, Morris S, Bhagani S, Cropley I, Abubakar I, et al. **Decreasing cost effectiveness of testing for latent TB in HIV in a low TB incidence area**. The European respiratory journal. 2015;**46**(1):165-74.

220. BHIVA. **UK National Guidelines for HIV Testing**. 2008.

221. Bradford-Hill A. **The environment and disease: association or causation?** Proceedings of the Royal Society of Medicine. 1965;**58**(5):295.

222. Public Health England. National HIV surveillance data tables 2016. [Accessed 21/06/2017]. Available from: https://www.gov.uk/government/statistics/hiv-annual-data-tables.

223. Public Health England. **HIV testing in England: 2016 report.** Public Health England, London, UK.: 2016.

224. WHO Global Health Observatory. Tuberculosis incidence data by country 2016. [Accessed 28/09/2015]. Available from: http://apps.who.int/gho/data/node.main.1320?lang=en.

225. UNAIDS. Global AIDS update 2016 2016. [Accessed 25/09/2017]. Available from: http://www.unaids.org/sites/default/files/media_asset/global-AIDS-update-2016_en.pdf.

226. Fortun J, Martin-Davila P, Gomez-Mampaso E, Gonzalez-Garcia A, Barbolla I, Gomez-Garcia I, et al. **Extra-pulmonary tuberculosis: differential aspects and role of 16S-rRNA in urine**. The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease. 2014;**18**(4):478-85.

227. Namme LH, Marie-Solande D, Hugo Bertrand MN, Elvis T, Achu JH, Christopher K. **Extrapulmonary tuberculosis and HIV coinfection in patients**

treated for tuberculosis at the Douala General Hospital in Cameroon. Annals of Tropical Medicine and Public Health. 2013;**6**(1):100-4.

228. Khan PY, Crampin AC, Mzembe T, Koole O, Fielding KL, Kranzer K, et al. **Does antiretroviral treatment increase the infectiousness of smear-positive pulmonary tuberculosis?** The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease. 2017;**21**(11):1147-54.

229. Vanhomwegen J, Kwara A, Martin M, Gillani FS, Fontanet A, Mutungi P, et al. **Impact of immigration on the molecular epidemiology of tuberculosis in Rhode Island**. J Clin Microbiol. 2011;**49**(3):834-44.

230. Gonzalez Diaz A, Battaglioli T, Diaz Rodriguez R, Goza Valdes R, Gonzalez Ochoa E, Van der Stuyft P. **Molecular epidemiology of tuberculosis in Havana, Cuba, 2009**. Tropical medicine & international health : TM & IH. 2015;**20**(11):1534-42.

231. Fok A, Numata Y, Schulzer M, FitzGerald MJ. **Risk factors for clustering of tuberculosis cases: a systematic review of population-based molecular epidemiology studies**. The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease. 2008;**12**(5):480-92.

232. Noppert GA, Yang Z, Clarke P, Ye W, Davidson P, Wilson ML. **Individual- and neighborhood-level contextual factors are associated with Mycobacterium tuberculosis transmission: genotypic clustering of cases in Michigan, 2004-2012**. Annals of epidemiology. 2017;**27**(6):371-6 e5.

233. Fenner L, Gagneux S, Helbling P, Battegay M, Rieder HL, Pfyffer GE, et al. **Mycobacterium tuberculosis transmission in a country with low tuberculosis incidence: role of immigration and HIV infection**. J Clin Microbiol. 2012;**50**(2):388-95.

234. Stucki D, Ballif M, Egger M, Furrer H, Altpeter E, Battegay M, et al. **Standard Genotyping Overestimates Transmission of Mycobacterium tuberculosis among Immigrants in a Low-Incidence Country**. J Clin Microbiol. 2016;**54**(7):1862-70.

235. Rodrigo T, Cayla JA, Garcia de Olalla P, Galdos-Tanguis H, Jansa JM, Miranda P, et al. **Characteristics of tuberculosis patients who generate secondary cases**. The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease. 1997;**1**(4):352-7.

236. Yuen CM, Kammerer JS, Marks K, Navin TR, France AM. **Recent Transmission of Tuberculosis - United States, 2011-2014**. PloS one. 2016;**11**(4):e0153728.

237. Public Health England. **TB strain typing and cluster investigation handbook: 3rd Edition (February 2014).** London: 2014.

238. Love J, Sonnenberg P, Glynn JR, Gibson A, Gopaul K, Fang Z, et al. **Molecular epidemiology of tuberculosis in England, 1998**. The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease. 2009;**13**(2):201-7.

239. Hamblion EL, Le Menach A, Anderson LF, Lalor MK, Brown T, Abubakar I, et al. **Recent TB transmission, clustering and predictors of large clusters in London, 2010-2012: results from first 3 years of universal MIRU-VNTR strain typing**. Thorax. 2016;**71**(8):749-56.

240. Gupta RK, Brown AE, Zenner D, Rice B, Yin Z, Thomas HL, et al. **CD4+ cell count responses to antiretroviral therapy are not impaired in HIV-infected individuals with tuberculosis co-infection**. Aids. 2015;**29**(11):1363-8.

241. Zaninotto P, Falaschetti E. **Comparison of methods for modelling a count outcome with excess zeros: application to Activities of Daily Living (ADL-s)**. J Epidemiol Community Health. 2011;**65**(3):205-10.

242. Tostmann A, Kik SV, Kalisvaart NA, Sebek MM, Verver S, Boeree MJ, et al. **Tuberculosis transmission by patients with smear-negative pulmonary tuberculosis in a large cohort in the Netherlands**. Clin Infect Dis. 2008;**47**(9):1135-42.

243. Kamper-Jorgensen Z, Andersen AB, Kok-Jensen A, Bygbjerg IC, Andersen PH, Thomsen VO, et al. **Clustered tuberculosis in a low-burden country: nationwide genotyping through 15 years**. J Clin Microbiol. 2012;**50**(8):2660-7.

244. Dahle UR, Eldholm V, Winje BA, Mannsaker T, Heldal E. **Impact of immigration on the molecular epidemiology of Mycobacterium tuberculosis in a low-incidence country**. Am J Respir Crit Care Med. 2007;**176**(9):930-5.

245. Barlow RE, Gascoyne-Binzi DM, Gillespie SH, Dickens A, Qamer S, Hawkey PM. **Comparison of variable number tandem repeat and IS6110-restriction fragment length polymorphism analyses for discrimination of high- and low-copy-number IS6110 Mycobacterium tuberculosis isolates**. J Clin Microbiol. 2001;**39**(7):2453-7.

246. de Beer JL, van Ingen J, de Vries G, Erkens C, Sebek M, Mulder A, et al. **Comparative study of IS6110 restriction fragment length polymorphism and variable-number tandem-repeat typing of Mycobacterium tuberculosis isolates in the Netherlands, based on a 5-year nationwide survey**. J Clin Microbiol. 2013;**51**(4):1193-8.

247. Office for National Statistics. Migration Statistics Quarterly Report: August 2017 2017. [Accessed 22/10/17]. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/internationalmigration/bulletins/migrationstatisticsquarterlyreport/august2017.

248. Chaisson RE, Schecter GF, Theuer CP, Rutherford GW, Echenberg DF, Hopewell PC. **Tuberculosis in patients with the acquired immunodeficiency syndrome. Clinical features, response to therapy, and survival**. The American review of respiratory disease. 1987;**136**(3):570-4.

249. NICE. Tuberculosis: Latent TB 2011. [Accessed 26/01/2017]. Available from: https://www.nice.org.uk/guidance/ng33/chapter/Recommendations#latent-tb.

9 Appendices

9.1 Supporting information for Chapter 2: Detailed search terms and search results from the systematic review

All searches were run on 4th May 2017

EMBASE

# ▲	Searches	Results
11	("american samoa" or andorra or antigua or barbuda or australia or austria or barbados or belgium or bermuda or bonaire or eustatius or saba or british or "virgin islands" or canada or "cook islands" or cuba or curacao or cyprus or czech or denmark or finland or france or germany or greece or grenada or hungary or iceland or ireland or EIRE or israel or italy or jamaica or jordan or luxembourg or malta or monaco or montserrat or netherlands or "new zealand" or NZ or niue or norway or oman or "puerto rico" or "saint kitts" or "nevis" or "saint lucia" or "st lucia" or "st vincent" or "saint vincent" or grenadines or "san marino" or seychelles or "sint maarten" or "st maarten" or slovakia or slovenia or sweden or switzerland or tokelau or turks or caicos or "united arab emirates" or UAE or UK or "united kingdom" or "great britain" or "england" or scotland or wales or "united states" or "united states of america" or USA or US or "wallis and futuna islands" or "west bank" or gaza).af.	27048910
12	(HIV or "human immunodeficiency virus" or "anti-retroviral" or "antiretroviral" or ART or retrovirus).ti.	269312
13	(tuberculosis or TB or "Mycobacterium tuberculosis").ti.	146450
14	2 and 3	6045
15	1 and 4	5103

Web of Science

Set	Results	
# 5	456	#4 AND #3 *Indexes=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI, CCR-EXPANDED, IC* *Timespan=All years*
# 4	4,969	#2 AND #1 *Indexes=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI, CCR-EXPANDED, IC* *Timespan=All years*
# 3	3,524,773	TOPIC: ("american samoa" or andorra or antigua or barbuda or australia or austria or barbados or belgium or bermuda or bonaire or “eustatius” or “saba” or "british" or "virgin islands" or canada or "cook islands" or cuba or curacao or cyprus or czech* or denmark or finland or france or germany or greece or grenada or hungary or iceland or ireland or EIRE or israel or italy or jamaica or jordan or luxembourg or malta or monaco or montserrat or netherlands or "new zealand" or NZ or niue or norway or oman or "puerto rico" or "saint kitts" or "nevis" or "saint lucia" or "st lucia" or "st vincent" or "saint vincent" or grenadines or "san marino" or seychelles or "sint maarten" or "st maarten" or slovakia or slovenia or sweden or switzerland or tokelau or turks or caicos or "united arab emirates" or UAE or UK or "united kingdom" or "great britain" or "england" or scotland or wales or "united states" or "united states of america" or USA or US or "wallis and futuna islands" or "west bank" or gaza) *Indexes=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI, CCR-EXPANDED, IC* *Timespan=All years*
# 2	442,239	TITLE: (HIV or "human immunodeficiency virus" or "anti-retroviral" or "antiretroviral" or ART or retrovirus) *Indexes=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI, CCR-EXPANDED, IC* *Timespan=All years*
# 1	98,477	TITLE: (tuberculosis or TB or "Mycobacterium tuberculosis") *Indexes=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI, CCR-EXPANDED, IC* *Timespan=All years*

PubMed

Search	Query	Items found
#5	Search **((((("american samoa" or andorra or antigua or barbuda or australia or austria or barbados or belgium or bermuda or bonaire or "eustatius" or "saba" or "british" or "virgin islands" or canada or "cook islands" or cuba or curacao or cyprus or czech* or denmark or finland or france or germany or greece or grenada or hungary or iceland or ireland or EIRE or israel or italy or jamaica or jordan or luxembourg or malta or monaco or montserrat or netherlands or "new zealand" or NZ or niue or norway or oman or "puerto rico" or "saint kitts" or "nevis" or "saint lucia" or "st lucia" or "st vincent" or "saint vincent" or grenadines or "san marino" or seychelles or "sint maarten" or "st maarten" or slovakia or slovenia or sweden or switzerland or tokelau or turks or caicos or "united arab emirates" or UAE or UK or "united kingdom" or "great britain" or "england" or scotland or wales or "united states" or "united states of america" or USA or US or "wallis and futuna islands" or "west bank" or gaza)))) AND ((tuberculosis[MeSH Terms]) AND ((hiv[Title]) OR ("human immunodeficiency virus"[Title]) OR (anti-retroviral[Title]) OR (antiretroviral[Title]) OR (ART[Title]) OR (retrovirus[Title]) OR (hiv[MeSH Terms])))**	2988
#4	Search **((tuberculosis[MeSH Terms]) OR (mycobacterium tuberculosis[MeSH Terms]) OR (tuberculosis[Title]) OR (tb[Title])) AND ((hiv[Title]) OR ("human immunodeficiency virus"[Title]) OR (anti-retroviral[Title]) OR (antiretroviral[Title]) OR (ART[Title]) OR (retrovirus[Title]) OR (hiv[MeSH Terms]))**	6226
#3	Search **(("american samoa" or andorra or antigua or barbuda or australia or austria or barbados or belgium or bermuda or bonaire or "eustatius" or "saba" or "british" or "virgin islands" or canada or "cook islands" or cuba or curacao or cyprus or czech* or denmark or finland or france or germany or greece or grenada or hungary or iceland or ireland or EIRE or israel or italy or jamaica or jordan or luxembourg or malta or monaco or montserrat or netherlands or "new zealand" or NZ or niue or norway or oman or "puerto rico" or "saint kitts" or "nevis" or "saint lucia" or "st lucia" or "st vincent" or "saint vincent" or grenadines or "san marino" or seychelles or "sint maarten" or "st maarten" or slovakia or slovenia or sweden or switzerland or tokelau or turks or caicos or "united arab emirates" or UAE or UK or "united kingdom" or "great britain" or "england" or scotland or wales or "united states" or "united states of america" or USA or US or "wallis and futuna islands" or "west bank" or gaza))**	11894404
#2	Search **(tuberculosis[MeSH Terms]) OR (mycobacterium tuberculosis[MeSH Terms]) OR (tuberculosis[Title]) OR (tb[Title])**	211218
#1	Search **(hiv[Title]) OR ("human immunodeficiency virus"[Title]) OR (anti-retroviral[Title]) OR (antiretroviral[Title]) OR (ART[Title]) OR (retrovirus[Title]) OR (hiv[MeSH Terms])**	241667

9.2 Supporting information for Chapter 2: Sociodemographic factors associated with latent tuberculosis infection

Lead author, year	Age	Ethnicity	Country of birth or residence	Sex	Pregnancy	Route of HIV infection	Drug misuse
Aichelburg, 2009	Age was not significantly associated with QFT-GIT positivity.	Black participants were more likely to be QFT-GIT positive than white	Risk of QFT-GIT positivity was increased for participants from other European countries (compared to Austria) and even higher among participants from African countries. The risk was positively correlated with TB incidence in country of birth.	Men were less likely to be QFT-GIT positive.	-	Route of HIV infection was not associated with QFT-GIT positivity.	Injecting drug use was not associated with QFT-GIT positivity.
Bourgarit, 2015	Age was not significantly associated with IGRA positivity.	-	Birth or residence in a country with TB incidence >100/100,000 was associated with IGRA-positivity in multivariable analysis.	Men were less likely to be IGRA-positive but this was not significant in multivariable analysis.	-	-	-
Brassard, 2009	-	-	Patients from Africa or Haiti had higher prevalence of TST positivity than patients from Canada.	-	-	There were significant differences in TST positivity by route of HIV infection, but they were not described.	-
Brock, 2006	QFT-IT-positive individuals were significantly younger than QFT-IT-negative individuals.	-	Long-term residence in a country with TB incidence >25/100,000 was positively associated with positive QFT-IT.	Sex was not significantly associated with QFT-GIT result.	-	-	Injecting drug use was not associated with QFT-GIT positivity.
Cheallaigh, 2013	Age was not significantly	-	Being from a country with TB incidence >100/100,000 was positively	Sex was not significantly	-	-	Injecting drug use was not

	associated with QFT-GIT or T-SPOT.TB positivity.		associated with QFT-GIT and T-SPOT.TB positivity.	associated with QFT-GIT or T-SPOT.TB positivity.			significantly associated with QFT-IT or T-SPOT.TB positivity
Doshi, 2012	Age was not associated with TST conversion.	No significant difference in rates of TST conversion for white, black or Hispanic people, but higher rates among Pacific islanders and Asians	Participants born outside of the mainland USA had higher rates of TST conversion. Not significant when Asians were excluded from the analysis.	Sex was not associated with the TST conversion rate.	-	-	Recreational drug use was not associated with TST conversion.
Elzi, 2007	Age was not associated with TST positivity.	-	Being from sub-Saharan Africa was strongly associated with TST positivity. Being from eastern Europe, east Asia or the Caribbean was also associated with TB positivity.	Women were less likely to be TST positive than men.	-	TST positivity was positively associated with acquiring HIV infection through heterosexual sex, injecting drug use or unknown route, compared to MSM.	-
French, 2006	-	-	-	All participants were women.	-	-	-
Gampper, 1998	-	African-Americans and Hispanics had higher rates of TST conversion than white	-	Sex was not associated with TST positivity.	-	Compared to men who have sex with men, TST positivity was higher for people who inject drugs or who have a partner who injects	-

		individuals.				drugs, and even higher if both criteria apply	
Girardi, 2002	Age was not associated with TST conversion.	-	-	Sex was not associated with TST conversion.	-	-	Injecting drug use was not associated with TST conversion.
Kall, 2012	Age was not associated with T-SPOT.TB positivity	Black Africans, black Caribbeans and Asians had higher proportions of T-SPOT.TB positivity than white British individuals.	Higher TB incidence in country of birth was positively associated with increased odds of T-SPOT.TB positivity.	Sex was not associated with T-SPOT.TB positivity.	-	-	-
Luetkemeyer, 2007	Age was not significantly associated with LTBI prevalence.	Ethnicity was not significantly associated with LTBI prevalence.	TB incidence ≥25/100,000 in country of birth was associated with higher prevalence of QFT-GIT positivity in multivariable analysis. There was no association between country of birth and TST positivity.	Sex was not significantly associated with LTBI prevalence.	-	-	Injecting drug use was not associated with LTBI prevalence.
McLaughlin, 2003	-	-	-	All participants were men.	-	-	-
Mofenson, 1995	There was a trend towards older patients being more likely to be TST-positive.	Black participants were more likely to be TST-positive, white participants were the least	-	All participants were women.	Pregnancy was not associated with TST positivity.	-	Higher rates of illicit drug use among TST-positive participants than TST-negative participants

		likely					
Pullar, 2014	Age was not significantly associated with QFT-GIT or TST positivity.	-	Being born in a TB-endemic country was positively associated with TST and QFT-GIT positivity but the risk decreased with time since migration.	Sex was not significantly associated with QFT-GIT or TST positivity.	-	-	-
Schulte, 2002	Age was not associated with TST positivity.	-	Foreign-born patients were more likely to be TST-positive.	All participants were women.	All participants were pregnant.	-	Injecting drug use was not associated with TST positivity
Stephan, 2008	Age was not associated with QFT-GIT, TST or T-SPOT.TB positivity.	Ethnicity was not associated with QFT-GIT, TST or T-SPOT.TB positivity.	Higher rates of TST positivity among participants from countries with high TB prevalence.	Sex was not associated with QFT-GIT, TST or T-SPOT.TB positivity.	-	Route of HIV infection was not associated with QFT-GIT, TST or T-SPOT.TB positivity.	Injecting drug use as route of HIV infection was not associated with QFT-GIT, TST or T-SPOT.TB positivity.

Lead author, year	Alcohol misuse	Homelessness	Imprisonment	Contact with TB or other close exposure to TB	Level of education	Income	Occupation
Aichelburg, 2009	-	-	-	-	-	-	-
Bourgarit, 2015	-	Collective living (in a group or institution) was positively associated with IGRA-positivity.	-	People with a contact with TB, or people who did not know if they had a contact with TB, had higher rates of IGRA positivity.	-	-	-
Brassard, 2009	-	-	-	-	-	-	-
Brock, 2006	Alcohol misuse was not associated with QFT-GIT positivity.	-	-	Exposure to a smear-positive TB case was positively associated with QFT-GIT positivity.	-	-	-
Cheallaigh, 2013	-	Homelessness was not significantly associated with QFT-IT or T-SPOT.TB positivity	Imprisonment was not significantly associated with QFT-IT or T-SPOT.TB positivity.	Having a close contact with TB was not significantly associated with QFT-IT or T-SPOT.TB positivity.	-	-	Occupation was not significantly associated with QFT-IT or T-SPOT.TB positivity.
Doshi, 2012	Alcohol dependence was not associated with TST conversion.	Those who have ever lived in a homeless shelter had higher rates of TST conversion.	Having ever been incarcerated was not associated with TST conversion.	-	-	-	-
Elzi, 2007	-	-	-	-	-	-	-
French, 2006	-	-	-	-	-	-	-

Gampper, 1998	-	-	-	-	-	-	-
Girardi, 2002	-	-	-	-	-	-	-
Kall, 2012	-	-	-	-	-	-	-
Luetkemeyer, 2007	-	Homelessness in the previous year was not associated with LTBI prevalence.	-	Living or working in a homeless shelter, prison or drug rehabilitation unit was associated with higher rates of TST positivity in multivariable analysis, but not QFT-GIT positivity. Self-reported contact with an active TB case was not associated with TST or QFT-GIT positivity.	-	-	-
McLaughlin, 2003	-	-	All participants were in prison. TST conversion rates were not significantly different in the dormitory of the source patient than in other dormitories over the same time period.	All participants were exposed to the source patient. Increased odds of TST conversion for patients on the same side of the dormitory as the source patient and patients exposed to the source patient for a longer time	-	-	-

				period.			
Mofenson, 1995	-	Homelessness and living in a shelter were both more common among TST-positive participants.	Imprisonment was not associated with TST positivity.	More TST-positive participants had a close contact with TB than TST-negative participants.	Level of education was not associated with TST positivity.	Income was not associated with TST positivity.	-
Pullar, 2014	-	-	-	Contact with a TB case was associated with higher odds of QFT-GIT and TST positivity. Visiting a TB-endemic country was not associated with QFT-GIT or TST positivity.	-	-	-
Schulte, 2002	-	-	-	Higher prevalence of TST positivity among participants exposed to a TB case. No TST positivity among people who had not been exposed.	-	-	-
Stephan, 2008	-	Being homeless or living in a shelter was not associated with QFT-GIT, TST or T-SPOT.TB	Imprisonment was not associated with QFT-GIT, TST or T-SPOT.TB positivity.	-	-	Receiving public social aid was not associated with QFT-GIT,	-

	positivity.		TST or T-SPOT.TB positivity.

IGRA: interferon-gamma release assay. LTBI: latent tuberculosis infection. QFT-GIT: quantiFERON-gold in-tube. TB: tuberculosis. TST: tuberculin skin test.

9.3 Supporting information for Chapter 2: Clinical factors associated with latent tuberculosis infection

Lead author, year	CD4 count	Other anergy / DTH	Viral load	Anti-retroviral therapy	BCG vaccination
Aichelburg, 2009	QFT-GIT result was not associated with median CD4 count but participants with CD4 <200 were less likely to be QFT-GIT-positive.	-	Median viral load was higher among QFT-GIT-positive participants than QFT-GIT-negative participants.	More QFT-GIT-negative participants were receiving ART than QFT-GIT-positive patients.	-
Bourgarit, 2015	Participants with higher CD4 count had non-significantly higher rates of IGRA (QFT-GIT or T-SPOT.TB) positivity.	-	No significant association between median viral load and IGRA-positivity	-	No significant association between BCG vaccination and IGRA positivity.
Brassard, 2009	Higher baseline CD4 count was associated with higher rates of TST positivity.	-	-	No significant association between ART and TST positivity.	-
Brock, 2006	No significant association between CD4 count and QFT-GIT positivity.	-	No significant association between viral load and QFT-GIT result.	No significant association between ART and QFT-GIT result.	-
Cheallaigh, 2013	No significant association between CD4 count and T-SPOT.TB, but increased CD4 count was positively associated with QFT-GIT positivity.	-	No significant association between viral load and QFT-GIT or T-SPOT.TB result.	No significant association between ART and QFT-GIT or T-SPOT.TB positivity.	-
Doshi, 2012	Participants whose CD4 count increased from baseline had increased odds of TST conversion, compared to participants whose CD4 count decreased.	-	No significant association between viral load and TST conversion	-	-
Elzi, 2007	CD4 count was positively associated with TST positivity.	-	Lower viral load was associated with TST positivity.	ART was negatively associated with TST positivity.	-
French, 2006	There was a non-significant positive correlation between CD4 count and	-	No significant association between viral load and TST	Rates of TST conversion decreased	-

	TST positivity.		positivity.	with time since initiation of ART.	
Gampper, 1998	-	-	-	-	-
Girardi, 2002	Increased CD4 count was associated with higher rates of TST conversion.	A positive delayed-type hypersensitivity testing result at baseline was not associated with TST conversion.	No significant association between viral load and TST conversion.	All patients were on ART.	-
Kall, 2012	No significant association between CD4 count and T-SPOT.TB positivity.	-	-	No significant association between ART and T-SPOT.TB positivity.	-
Luetkemeyer, 2007	Higher nadir CD4 count was associated with higher rates of TST and QFT-GIT positivity in univariable analyses, but not in adjusted analyses.	-	No significant association between viral load and TST or QFT-GIT positivity.	No significant association between ART and TST or QFT-GIT positivity.	No significant association between BCG vaccination and TST or QFT-GIT positivity.
McLaughlin, 2003	No significant association between CD4 count and TST positivity.	-	No significant association between viral load and TST positivity.	No significant association between ART and TST positivity.	-
Mofenson, 1995	-	-	-	-	-
Pullar, 2014	Nadir CD4 count >350 was associated with higher rates of TST and QFT-GIT positivity.	-	-	Patients on ART had lower rates of QFT-GIT and TST positivity, but the association was not significant.	-
Schulte, 2002	-	-	-	-	-
Stephan, 2008	QFT-GIT-positive participants had	-	No significant association	No significant	-

higher median CD4 count than QFT-GIT-negatives but there was no significant association between nadir CD4 count and QFT-GIT positivity. CD4 count was not associated with TST or T-SPOT.TB positivity.		between viral load and QFT-GIT, TST or T-SPOT.TB result.	association between ART and QFT-GIT, TST or T-SPOT.TB result.

Lead author, year	Previous TB disease	LTBI infection	Previous treatment for LTBI	AIDS-defining illness or other opportunistic infections	Other non-infectious co-morbidities	Time since HIV diagnosis / duration of follow-up time
Aichelburg, 2009	Previous active TB was more common among QFT-GIT-positive participants than QFT-GIT-negative participants. The median interval since previous disease did not differ by QFT-GIT status.	-	-	-	-	-
Bourgarit, 2015	Previous active TB was more common among IGRA-positive participants than IGRA-negative participants.	Higher proportion of IGRA-positive participants had a history of LTBI than IGRA-negative participants.	-	-	-	-
Brassard, 2009	-	-	-	-	-	-
Brock, 2006	Participants with a previous TB diagnosis had higher rates of QFT-GIT positivity.	-	-	No significant association between prior AIDS diagnosis and positive QFT-GIT.	Diabetes was not associated with positive QFT-GIT.	No association between the number of years since HIV diagnosis and QFT-GIT positivity.
Cheallaigh, 2013	-	-	-	-	-	-
Doshi, 2012	-	-	-	-	-	-
Elzi, 2007	-	-	-	-	-	-
French, 2006	-	-	-	-	-	-
Gampper, 1998	-	-	-	-	-	-
Girardi, 2002	-	-	-	Prior AIDS diagnosis	-	-

				was not significantly associated with TST conversion.		
Kall, 2012	-	-	-	-	-	-
Luetkemeyer, 2007	Higher rates of previous TB among QFT-GIT-positive participants than QFT-GIT-negative. No significant association between prior TB and TST positivity.	Prior TST positivity was associated with a positive TST result and a positive QFT-GIT result.	Higher rates of previous treatment for LTBI among participants who tested TST or QFT-GIT positive	-	-	-
McLaughlin, 2003	-	-	-	-	-	-
Mofenson, 1995	-	-	-	-	-	-
Pullar, 2014	Higher rates of QFT-GIT and TST positivity among patients with a previous TB diagnosis. Only half of those previously treated for active TB had a positive QFT-GIT.	-	-	Prior AIDS diagnosis was not significantly associated with IGRA positivity.	-	Decreasing risk of QFT-GIT positivity for each year since HIV diagnosis in univariable but not multivariable analysis. No association between TST positivity and time since HIV diagnosis.
Schulte, 2002	-	-	-	Prior AIDS diagnosis was not significantly associated with TST positivity.	-	No association between HIV being undiagnosed at the first prenatal visit and TST positivity.
Stephan, 2008	More QFT-GIT or T-SPOT.TB positive participants had a	-	-	Higher proportion of TST-positive participants had a prior AIDS diagnosis	-	No association between time since HIV diagnosis and TST positivity.

	previous TB diagnosis than QFT-GIT/T-SPOT.TB negative participants.		than TST-negative participants

AIDS: acquired immunodeficiency syndrome. ART: anti-retroviral therapy. BCG: Bacillus Calmette-Guérin. DTH: delayed-type hypersensitivity. IGRA: interferon-gamma release assay. LTBI: latent tuberculosis infection. QFT-GIT: quantiFERON-gold in-tube. TB: tuberculosis. TST: tuberculin skin test.

9.4 Supporting information for Chapter 2: Sociodemographic factors associated with active tuberculosis disease

Lead author, year	Age	Ethnicity	Country of birth or residence	Region	Sex
Abgrall, 2010	Age was not significantly associated with risk of TB among both migrant and non-migrant patients.	-	Risk of TB was higher among patients originating from sub-Saharan Africa and other migrants than patients born in France.	Among non-migrant patients, risk of TB was lower among residents of southern France and other/Reunion Island compared to residents of Paris area. Among migrant patients, there was no significant association between region of care and risk of TB.	Grouped with route of HIV infection
Aichelburg, 2009	-	-	Country of origin not significantly associated with active TB	-	-
Antonucci, 1995	Age was not significantly associated with risk of TB in an unadjusted analysis.	-	-	Place of residence (North/central/south Italy) was not significantly associated with risk of TB in unadjusted analysis.	Sex was not significantly associated with risk of TB in unadjusted analysis.
Brassard, 2009	Mean age of TB patients was significantly lower than entire cohort.	-	Canadian born less likely to have TB on univariable analysis (no figures shown)	-	Lower proportion of males had TB compared to entire cohort.
Dooley, 1992	Age was not significantly associated with risk of TB.	-	-	-	Sex was not significantly associated with risk of TB on crude analysis
Elzi, 2007	Age was not significantly associated with risk of TB.	-	Being from sub-Saharan Africa was associated with a higher risk of TB.	-	Sex was not significantly associated with risk of TB.
Girardi, 2000	Age was not significantly associated with risk	-	Country of birth was not significantly associated with risk of TB.	Place of residence (North/central/south Italy) was not significantly associated with risk of	Sex was not significantly associated with risk

	of TB.			TB.	of TB.
Grant, 2009	Age was not significantly associated with risk of TB.	Compared to white participants, Black African and Other ethnic groups had higher risk of TB.	-	-	Grouped with route of HIV infection
Gupta, 2015	Participants between 36-55 years had lower risk of TB compared to those 16-35 years. No significant risk in participants 56 years or more.	Compared to whites, Black African, south Asian and other ethnicities had higher risk of TB.	-	-	Sex was not significantly associated with risk of TB.
Hasse, 2014	Participants >40 years had significantly higher risk of TB in both ART sub-groups (no previous ART and current ART).	-	Compared to those from Europe/North America/Australia, participants from other regions had higher risk of TB in both ART sub-groups.	-	-
Jones, 2000	Age was not significantly associated with risk of TB.	Ethnicity was not significantly associated with risk of TB.	Compared to US-born, foreign-born participants had a significantly higher risk of TB.	-	Women had a lower risk of TB disease than men.
Karo, 2014	Age was not significantly associated with risk of TB in both ART subgroups (never started cART and patients on cART)	-	Persons from SSA had higher risk of TB compared to those from Germany in both ART sub-groups	-	Sex was not significantly associated with risk of TB in both ART sub-groups.
Manavi, 2016	Age was not significantly associated with TB incidence.	Ethnicity was not significantly associated with TB incidence.	Compared to participants from low-incidence countries those from medium-incidence countries had significantly	-	-

			higher risk of TB, and the risk was even greater for patients from sub-Saharan Africa.		
Maniewski, 2016	Age was not significantly associated with risk of TB on crude analysis.	Among patients with TB, 83% were Africans, and 14% were Caucasian while among those without TB 45% were Africans, and 48% were Caucasian.	-	-	Sex was not significantly associated with TB on crude analysis.
Markowitz, 1997	-	Blacks had a higher risk of TB on univariable analysis. On adjusted analysis, ethnicity was not significantly associated with risk of TB.	-	Compared to residents of the Western USA, the risk was not significantly higher for patients from the Midwest, but was higher for patients from the eastern USA.	Sex was not significantly associated with risk of TB.
McLaughlin, 2003	-	All TB patients were black, compared to 89% of the study population.	-	-	-
Miguez-Burbano, 2003	-	-	-	-	-
Mohle-Boetani, 2002	-	-	-	-	All men
Mor, 2013	Increasing age was significantly associated with higher risk of TB.	-	Analysis restricted to non-Israel born participants. Risk of TB higher with increasing interval (in years) from arrival in Israel to HIV diagnosis.	-	Males had significantly higher risk of TB.
Moro, 2000	Age was not significantly associated with risk	-	-	-	Sex was not significantly associated with risk

	of MDR-TB.				of MDR-TB.
Pettit, 2011	Age was not significantly associated with risk of TB on crude analysis.	Higher proportions of Black participants among TB cases. Among TB cases, 39% were Blacks compared to 29% among non-TB cases.	Higher proportions of foreign-born with TB cases. Among TB cases, 21% were foreign-born compared to 5% among non-TB cases.	-	Significantly higher proportion of males among TB cases. Among TB cases, 91% were males while among non-TB cases 76% were males.
Pettit, 2016	Age was not significantly associated with risk of TB on crude analysis.	Higher proportions of Blacks among those with TB	Higher proportions of foreign-born among TB cases. Among TB cases, 19% were foreign-born, compared to 5% among non-TB cases.	-	Similar proportion of males among TB cases and non-TB cases.
Rice, 2013	Age was not significantly associated with risk of TB.	Compared to Whites, Black Africans, Indian/Pakistani/Bangladeshi, and Other ethnic groups had a significantly higher risk of TB. No significant risk for Black Caribbean.	No association between country of birth and risk of TB. Participants infected with HIV abroad (outside UK) had a significantly higher risk of TB.	-	Men had a significantly higher risk of TB.
Rubinstien, 1996	-	-	-	-	26/27 patients that developed TB were males.
Sackoff, 2001	-	-	-	-	-
Sterling, 2011	Age was not significantly associated with risk of TB.	Compared to white participants, black African Canadians, Hispanic and 'other' participants had higher risk of TB in	-	-	Sex was not significantly associated with risk of TB.

		multivariable analysis. The risk of TB was higher for black African Americans, but the significance was borderline in a multivariable model.			
Sudre, 1996	Age was not significantly associated with risk of TB.	-	Compared to participants from Switzerland, those from Portugal, eastern Europe, Brazil, Africa, other resource-poor countries had a higher risk of TB. TB risk of participants from other industrialised countries was similar to those from Switzerland. There was no significant risk among participants with unknown country of origin	-	Sex was not significantly associated with risk of TB.
Taarnhoj, 2011	Participants < 40 years had a higher risk of TB on univariable analysis. Age was not significantly associated with risk of TB on multivariable analysis.	-	Compared to participants from Denmark, those from Africa or Asia had a significantly higher risk of TB. No significant risk among patients from Europe or Other	-	Males had a higher risk of TB on univariable analysis. Sex was not significantly associated with risk of TB on multivariable analysis.
Thomas, 2000	-	Higher proportions of Blacks had TB. Among TB cases, 71%, 24% and 4% were Black, Hispanic and	-	-	Higher proportion of females among TB cases. Among TB cases, 67% were females compared to 49% among non-TB

		Other/unknown respectively; while among non-TB cases 57%, 36% and 7% were Black, Hispanic and Other/unknown.			cases.
Trieu, 2010	TB incidence was highest among participants 30-39 years.	Compared to non-Hispanic White participants, non-Hispanic Black, Hispanic and Asian had a higher TB incidence.	Overall, compared to US-born, TB incidence was higher among foreign-born.	-	Overall, TB incidence was not substantially different between men and women.
Turkova, 2015	Compared to children 5-9 years at baseline, children < 5 years had a lower risk of TB, while children 10+ years did not have any significant risk.	No significant association	Children born in the UK had a significantly lower risk of TB than children born abroad.	No significant association between region (London compared to the rest of the UK) and risk of TB on univariable analysis	No significant association between sex and risk of TB on univariable analysis.

Lead author, year	Route of HIV infection	Drug misuse	Alcohol misuse	Imprisonment	Level of education
Abgrall, 2010	Among non-migrant patients, risk of TB was higher among IDU men, heterosexual men, other men, IDU women, heterosexual women and other women than in homosexual men. Among migrant patients, there was no significant association between sex/HIV transmission group and risk of TB.	-	-	-	-
Aichelburg, 2009	-	-	-	-	-
Antonucci, 1995	HIV transmission category was not significantly associated with risk of TB in unadjusted analysis	-	-	-	-
Brassard, 2009	MSM less likely to have TB on univariable analysis (no figures shown)	-	-	-	-
Dooley, 1992	-	Injecting drug use was not significantly associated with risk of TB.	-	-	-
Elzi, 2007	Route of HIV infection was not significantly associated with risk of TB.	-	-	-	-
Girardi, 2000	Route of HIV infection was not significantly associated with risk of TB.	-	-	-	-
Grant, 2009	Compared to male homosexual participants, female heterosexual, male heterosexual, and female	-	-	-	-

	other participants had a higher risk of TB.				
Gupta, 2015	Compared to MSM, heterosexual, IDU and maternal-to-child routes of infection were significantly associated with a higher risk.	-	-	-	-
Hasse, 2014	-	Injecting drug use was associated with a higher risk of TB.	-	-	-
Jones, 2000	Compared to male-male sex, IDU and heterosexual persons had higher risk of TB.	-	-	-	-
Karo, 2014	Route of HIV infection was not significantly associated with risk of TB.	-	-	-	-
Manavi, 2016	-	-	-	-	-
Maniewski, 2016	-	-	-	-	-
Markowitz, 1997	Compared to homosexual men, IDUs had a higher risk of TB on univariable analysis. On adjusted analysis, there was no significant association between transmission category and risk of TB	-	-	-	No significant association between level of education and risk of TB.
McLaughlin, 2003	-	-	-	All participants were prison inmates.	-
Miguez-Burbano, 2003	-	All patients had a history of previous or current drug use. Long term tobacco users had higher risk of TB	-	-	-
Mohle-Boetani, 2002	-	-	-	All participants were	-

				prison inmates.	
Mor, 2013	Compared to MSM, heterosexual participants (from both HIV-epidemic and non-epidemic countries) and IDUs had a significantly higher risk of TB.	-	-	-	-
Moro, 2000	-	-	-	-	-
Pettit, 2011	Significantly higher proportions of heterosexuals among TB cases. Among TB cases, 44% had heterosexual contact, while among non-TB cases 23% had heterosexual contact. No significant difference in proportions of IDUs and MSMs in those with TB and those without TB.	-	-	-	-
Pettit, 2016	-	Injecting drug use was significantly more common (32%) among TB cases than PLHIV without TB (17%).	-	-	-
Rice, 2013	All participants were heterosexual.	-	-	-	-
Rubinstien, 1996	-	All participants were injecting drug users.	-	-	-
Sackoff, 2001	-	-	-	-	-
Sterling, 2011	-	People who inject drugs had a significantly higher risk of TB.	-	-	-
Sudre, 1996	HIV transmission route was	-	-	-	-

	not significantly associated with risk of TB.				
Taarnhoj, 2011	Compared to homosexuals, heterosexual and IDU participants had a higher risk of TB.	-	Alcohol abuse was not significantly associated with risk of TB.	-	-
Thomas, 2000	HIV-infected children with TB were more likely to have mothers who acquired HIV through injecting drug use than heterosexual sex or other routes of transmission.	-	-	-	-
Trieu, 2010	-	-	-	-	-
Turkova, 2015	-	-	-	-	-

IGRA: interferon-gamma release assay. LTBI: latent tuberculosis infection. QFT-GIT: quantiFERON-gold in-tube. TB: tuberculosis. TST: tuberculin skin test.

9.5 Supporting information for Chapter 2: Clinical factors associated with active tuberculosis disease

Lead author, year	Year of diagnosis	CD4 count	Viral load	Anti-retroviral therapy	Previous TB disease
Abgrall, 2010	Cohort entry at later periods was associated with higher risk of TB in both migrant and non-migrant sub-groups.	CD4 count <500 was associated with increasing risk of TB among both migrant and non-migrant patients.	VL 10,000+ was associated with increasing risk of TB among both migrant and non-migrant patients.	Compared to being ART-naïve, the risk of TB was significantly decreased for patients on mono/dual therapy. For patients on cART, the risk was also significantly decreased, and was lower in patients who had been on cART >6 months compared to <6 months.	-
Aichelburg, 2009	-	Patients with active TB had significantly lower median CD4 count but not lower nadir CD4 count.	Patients with active TB had significantly higher HIV-1 RNA levels	Patients with active TB were less likely to receive ART	-
Antonucci, 1995	-	CD4 count < 350 was associated with increasing risk of TB on multivariable analysis.	-	No significant association between ART and active TB.	History of TB was associated with higher risk of TB on multivariable analysis.
Brassard, 2009	-	No significant association between CD4 count and risk of TB.	-	No significant association between ART and active TB, however only 3 patients were receiving ART.	-
Dooley, 1992	-	-	-	-	-
Elzi, 2007	-	Increasing CD4 count was significantly associated with lower risk of TB.	Increasing HIV RNA level was significantly associated with higher risk of TB.	ART was associated with lower risk of TB	-
Girardi, 2000	-	Participants with CD4 count <200 had higher risk of TB.	-	Participants o 2 or more ART drugs had a lower risk of TB.	-
Grant, 2009	No significant association between year of cohort entry and risk of TB.	CD4 count <500 was associated with increasing risk of TB.	VL <50 was associated with a lower risk of TB.	Being on ART for 2 or more years was associated with a lower risk of TB.	-

Gupta, 2015	-	Compared to CD4 700 or more, having CD4 <500 was significantly associated with increasing risk of TB. Having unknown CD4 count was associated with a higher risk of TB. There was no significant difference in risk for participants with CD4 500-699.	Participants with VL 50+ or unknown VL had a higher risk of TB.	All patients were on ART.	-
Hasse, 2014	-	Increase in CD4 count was associated with a lower risk of TB in both ART sub-groups.	-	-	-
Jones, 2000	-	Participants with CD4<100 and missing CD4 count had a higher risk of TB compared to those with CD4 250+.	VL was not significantly associated with risk of TB (but VL available only for a subset).	Being on HAART and other ART were associated with lower risk of TB.	-
Karo, 2014	-	CD4 <200 was associated with higher risk of TB in both ART sub-groups. TB risk higher among participants that never started ART.	VL 5 or more was significantly associated with a higher risk of TB in both ART subgroups. TB risk higher among participants that never started ART.	Multivariable analysis stratified by ART status. But crude TB incidence higher among participants that never started ART.	-
Manavi, 2016	-	Compared to participants with CD4 count >500, those with CD4 count 200 or less had significantly lower TB incidence.	VL 40 or more was significantly associated with a higher TB incidence.	Duration of ART was not significantly associated with TB incidence.	-
Maniewski, 2016	-	Median CD4 nadir significantly lower among persons that developed TB, compared to those that did not develop TB	No significant association in median VL and TB.	No significant association between ART and TB on crude analysis.	-
Markowitz, 1997	-	Incidence of TB higher among participants with CD4 cell count < 200.	-	-	-

McLaughlin, 2003	-	No significant association between CD4 count and risk of TB.	No significant association between VL and risk of TB.	No significant association between ART and risk of TB.	-
Miguez-Burbano, 2003	-	No significant association between CD4 count and risk of TB.	-	-	-
Mohle-Boetani, 2002	-	No significant association between CD4 count and risk of TB.	-	-	No significant association between prior TB disease and risk of TB.
Mor, 2013	HIV diagnosis after periods of ART introduction (1997-2010) was associated with a significantly higher risk of TB.	-	-	No figures on ART, but risk of TB higher during the ART period, compared to pre-ART period.	-
Moro, 2000	-	Patients that developed MDR-TB had significantly lower median CD4 %.	-	-	-
Pettit, 2011	-	Median baseline CD4 count and CD4% were significantly lower among TB cases.	Median baseline VL was significantly higher among TB cases.	On multivariable analysis, compared to participants not on HAART at baseline, the risk of TB reduced after at least 180 days of HAART. Among those on HAART, risk of TB was higher within first 180 days of HAART.	-
Pettit, 2016	-	Median baseline CD4 count significantly lower among TB cases.	Median baseline VL significantly higher among TB cases	HAART (irrespective of duration) was associated with significant lower risk of TB on multivariable analysis. Accounting for duration of HAART, risk of TB was lower after being on HAART for > 6 months.	-
Rice, 2013	-	MV. CD4 count 100+ significantly associated with lower risk of TB.	-	TB incidence was higher among participants who were not on ART.	-
Rubinstien, 1996	-	-	-	-	-
Sackoff, 2001	-	-	-	-	-
Sterling, 2011	-	Decreasing CD4 count was	No significant	No significant association between	-

		significantly associated with an increasing risk of TB.	association between VL and risk of TB.	HAART initiation or HAART type and risk of TB. No significant association between duration of HAART and risk of TB.	
Sudre, 1996	-	Decreasing CD4 count was significantly associated with an increasing risk of TB.	-	-	-
Taarnhoj, 2011	No significant association between year of HIV diagnosis and risk of TB.	CD4 count 200 or less was associated with a higher risk of TB.	No significant association between VL and risk of TB.	Start of HAART (time-updated) was significantly associated with a lower risk of TB.	-
Thomas, 2000	-	A significantly higher proportion of patients with TB had moderate or severe immunosuppression than patients without TB.	-	-	-
Trieu, 2010	-	-	-	-	-
Turkova, 2015	No significant association between period of cohort entry (HIV diagnosis) and risk of TB.	-	Baseline and time-updated VL were not significantly associated with risk of TB on univariable analysis.	Baseline and time-updated ART were not significantly associated with risk of TB on univariable analysis.	-

Lead author, year	LTBI infection and anergy	Previous treatment for LTBI	Contact with TB or other close exposure to TB	AIDS-defining illness or other opportunistic infections	Time since HIV diagnosis / duration of follow-up time
Abgrall, 2010	-	-	-	Prior AIDS was associated with higher risk of TB among non-migrant patients. There was no significant association between prior AIDS and risk of TB among migrant patients. .	The risk of TB was highest in the first 6 months of follow-up and decreased over time in both migrant and non-migrant patients, although the decline for migrants was less steep than for non-migrants.
Aichelburg, 2009	-	-	-	There were not enough patients to assess the relationship between previous AIDS-defining illness and active TB. Patients diagnosed with active TB had a higher number of clinical symptoms of tuberculosis at baseline.	-
Antonucci, 1995	TST-positive and anergic patients had significantly higher risk of TB compared to tuberculin-negative non-anergic in multivariable analysis.	-	-	No significant association between CDC clinical class and risk of TB on multivariable analysis.	-
Brassard, 2009	-	No patient on LTBI	-	-	Patients with TB had

		treatment developed active TB compared to 2/3 not on LTBI treatment on univariable analysis.			significantly shorter time periods between HIV diagnosis and 1st clinic visit compared to the entire cohort in univariable analyses.
Dooley, 1992	-	-	Patients exposed to a TB case whilst in the HIV hospital unit had significantly greater risk of developing TB than unexposed patients.	AIDS-status was not significantly associated with risk of TB.	-
Elzi, 2007	Participants with positive TST results had higher incidence of TB overall and when stratified by region of origin.	None of patients treated developed TB	-	-	-
Girardi, 2000	TST-positive participants had a higher risk of TB.	-	-	CDC clinical class was not significantly associated with risk of TB.	-
Grant, 2009	-	-	-	-	-
Gupta, 2015	-	-	-	-	-
Hasse, 2014	-	-	-	Co-trimoxazole use associated with lower risk of TB in both ART sub-groups.	-
Jones, 2000	-	Preventive treatment was not significantly associated with risk of TB.	-	Persons with clinical AIDS had higher risk of TB.	-
Karo, 2014	-	-	-	-	TB incidence higher in first year of follow-up than in the second year.
Manavi, 2016	-	-	-	-	-

Maniewski, 2016	-	-	-	-	No significant association between follow-up time and TB.
Markowitz, 1997	Compared to PPD-negative participants, PPD-positive and converters had higher TB incidence.	-	-	-	-
McLaughlin, 2003	-	-	Persons on same side of dorm with TB case had a higher risk of TB, and had on average been exposed for less time than control patients.	-	-
Miguez-Burbano, 2003	-	-	-	-	-
Mohle-Boetani, 2002	No significant association between prior infection and risk of TB. No significant association between anergy and risk of TB.	Higher but not significant proportion of controls had preventive therapy.	Persons spending 20 hours or more in the communal day room had higher risk of TB, whilst those with a TV in their cell had lower risk of TB. There was no association between having an inmate with TB in the next cell and risk of TB.	-	-
Mor, 2013	-	-	-	-	-
Moro, 2000	-	-	-	-	-
Pettit, 2011	-	-	-	-	-
Pettit, 2016	-	-	-	-	-
Rice, 2013	-	-	-	-	-
Rubinstien, 1996	TB incidence was higher among participants with known TB	Among participants with known LTBI, TB incidence was higher among participants	-	-	-

	infection compared to those without documented TB infection.	who had not previously received isoniazid.			
Sackoff, 2001	TB incidence was higher among participants who were TST-positive than participants who were anergic or had unknown TST status. There were no TB cases among TST-negative participants.	Among TST-positive persons, TB incidence reduced with increasing duration of INH treatment.	-	-	-
Sterling, 2011	-	-	-	-	-
Sudre, 1996	-	-	-	-	Compared to patients with less than 6 months of follow-up, patients with 25-36 moths and > 36 months of follow-up had higher risk of TB.
Taarnhoj, 2011	-	-	-	-	-
Thomas, 2000	-	-	-	-	-
Trieu, 2010	-	-	-	-	-
Turkova, 2015	-	-	-	Compared to WHO immunological severe stage (time-updated), participants with stages none had lower risk of TB, while mild and advanced stages did not have any significant risk. No significant association	-

							between baseline AIDS-status and WHO stage and risk of TB.

AIDS: acquired immunodeficiency syndrome. ART: anti-retroviral therapy. BCG: Bacillus Calmette-Guérin. DTH: delayed-type hypersensitivity. IGRA: interferon-gamma release assay. LTBI: latent tuberculosis infection. QFT-GIT: quantiFERON-gold in-tube. TB: tuberculosis. TST: tuberculin skin test.