Evaluation of the Tuberculosis Strain Typing Service in England

Jessica de Windt Mears

Doctor of Philosophy

UCL

Declaration

I, Jessica Mears, 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 response to rising tuberculosis notification rates in England, the tuberculosis strain typing service (TB-STS) was introduced in 2010. This complex intervention involves MIRU-VNTR (mycobacterial interspersed repetitive unit-variable number tandem repeats) typing of isolates from all tuberculosis patients for the prospective identification, reporting and investigation of strain typing clusters. A mixed-methods prospective evaluation was conducted, assessing the Structures, Processes, Outputs, Outcomes and Context of the TB-STS.

Structures were described and cost of the service estimated. Processes were explored using cross-sectional surveys, and semi-structured interviews assessed user experience. Outputs included the impact of the TB-STS on detection of false-positive *Mycobacterium tuberculosis* diagnoses, and its effect on contact tracing yield, diagnostic delay and cluster growth. These findings informed the estimation of the outcomes: a deterministic mathematical model to estimate the effectiveness of the TB-STS over 20 years and a cost-effectiveness analysis.

The TB-STS cost approximately £1m per year. Between the initial and follow-up surveys, knowledge of strain typing increased, perceived usefulness did not change, and time spent investigating tuberculosis transmission increased. Interviews identified broader benefits such as improved contact tracing and the research potential of a strain typing dataset. Between 2010 and 2012, 17,168 isolates were typed. The TB-STS detected 17 additional false-positive diagnoses and had no significant effect on contact tracing yield, diagnostic delay or cluster growth. Mathematical modelling suggested the TB-STS would not reduce tuberculosis incidence in the white UK-born population. However, in the non-white UK-born and non-UK-born populations, moderate reductions in tuberculosis incidence could be observed if detection of latent infection increases from 3%-13%/year or diagnostic delay decreases by one week. The TB-STS was not predicted to be cost-effective (£95,62/QALY).

The TB-STS in its current form was not effective or cost-effective; however, broader benefits justify its continuation. Recommendations are made for the TB-STS, future typing services and their evaluation.

Table of contents

List of tables

List of figures

List of boxes

Abbreviations

ARI	Annual risk of infection
CCDC	Consultant in Communicable Diseases
CI	Cluster Investigator
ETS	Enhanced Tuberculosis Surveillance
HIV	Human Immunodeficiency Virus
HPA	Health Protection Agency
HPS	Health Protection Services
HPU	Health Protection Unit
ICER	Incremental Cost-effectiveness Ratio
IGRA	Interferon Gamma Release Assay
IQR	Inter-quartile range
LaRS	Local and Regional Services
LTBI	Latent Tuberculosis Infection
MIRU-VNTR	Mycobacterial Interspersed Repetitive Units-Variable Number of Tandem Repeats
MRC	Medical Research Council
NHS	National Health Service
NICE	National Institute for Clinical Excellence
ONS	Office for National Statistics
PHE	Public Health England
PT	Preventative treatment
QALY	Quality Adjusted Life Year
Quango	Quasi autonomous non-governmental organisations
RCT	Randomised controlled trial
RMN	Regional Microbiology Network
RFLP	Restriction Fragment Length Polymorphism
SNP	Single nucleotide polymorphism
STM	Strain Typing Module
STROME-ID	Strengthening the Reporting of Molecular Epidemiology for Infectious Diseases
STPB	Strain Typing Project Board
TB	Tuberculosis
TB-STS	Tuberculosis Strain Typing Service
TST	Tuberculin Skin Test
UCL	University College London
WGS	Whole Genome Sequencing
WHO	World Health Organization
WTE	Working Time Equivalent

Glossary of key terms

Term	Definition
Cluster	Two or more *M. tuberculosis* isolates with indistinguishable strain types
Cluster investigation	An investigation to identify epidemiological links between TB patients whose isolates have indistinguishable strain types. A cluster investigation may consist of reviewing information from medical records, interviewing case managers and TB patients
Contact tracing	The identification and screening of contacts of TB cases that are at risk of being infected
Contact tracing yield	The proportion of contacts who have latent TB infection or have active TB disease
Diagnostic delay	The time between symptom onset and TB diagnosis or start of TB treatment
Epidemiologically confirmed cluster	Two or more people with TB who share definite epidemiological links (see below)
Epidemiological link	A connection between two TB patients involving people, places or time
Public health action	Any additional action that is taken in an attempt to interrupt TB transmission
Recent transmission	The transmission of TB between people that has occurred in the recent past (as opposed to reactivation of latent TB infection)

Outline of thesis

This thesis contains a number of studies that contribute to the evaluation of the tuberculosis (TB) strain typing service (TB-STS) in England.

Chapter 1 presents the background to the thesis. An overview is provided of the epidemiology and natural history of TB, TB control and TB strain typing. Existing strain typing services are described and the TB-STS in England is briefly introduced. Potential frameworks for the evaluation of the TB-STS are discussed. Finally, the aims and objectives of this thesis are presented.

Chapter 2 systematically reviews the literature to investigate the biases associated with interpreting the proportion of clustering estimated by strain typing data using MIRU-VNTR.

Chapter 3 details the methods used to address the objectives of the thesis.

Chapter 4 describes the Structures of the TB-STS.

Chapter 5 assesses the Processes within the TB-STS. The results of an initial and follow-up cross sectional survey of health protection and clinic staff are presented to describe the implementation and perception of the TB-STS amongst service providers and users. The user experience of the TB-STS is further explored using semi-structured interviews.

Chapter 6 describes the Outputs of the TB-STS, including false positive TB isolation detected by the TB-STS, cluster investigation activity and outcomes, and the impact of cluster investigations on contact tracing, diagnostic delay, the rate of cluster growth.

Chapter 7 presents the Outcomes of the TB-STS. It describes the results of the transmission model used to assess the effectiveness of the TB-STS and the cost-effectiveness model used to evaluate the cost-effectiveness of the TB-STS.

Chapter 8 contextualises the findings of this thesis. Findings are summarised and a discussion of the thesis follows, including the strengths and limitations of the evaluation, the interpretation of the findings and the relevance of the wider context and broader benefits. In this chapter recommendations are made for the TB-STS and future strain typing policy, as well as suggestions for molecular typing services in general and future evaluations of complex public health interventions.

My role in this thesis

I was the scientist and project coordinator for the evaluation of the TB-STS in England, which was chaired by my primary supervisor, Dr Pam Sonnenberg. The evaluation was steered by a group of experts from many disciplines, including my secondary and tertiary supervisors, Professor Ibrahim Abubakar and Professor Tim McHugh. I was based at PHE for the first nine months of the project, joined UCL as a part-time student and then moved to UCL full-time once funding for the PhD was secured. I will outline my role in the thesis here.

I designed, conducted, analysed and interpreted the results of a systematic review, with the advice of my supervisors and Professor Ted Cohen.

I identified an appropriate evaluation framework for the evaluation and presented this to the evaluation group at the inaugural meeting. Together with my supervisor, I led the initial discussions about the evaluation study design. I initiated and organised a fact-finding mission to the Netherlands to learn about the Dutch TB strain typing service. I developed and maintained relationships with members of the Public Health England TB Strain Typing Project Board in order for me to conduct interviews with them at different time points, to describe the TB-STS in detail.

I led the design of the initial and follow-up cross-sectional surveys, I piloted and conducted the surveys, after which I cleaned, analysed and interpreted the data. I designed and conducted the semi-structured interviews and analysed and interpreted the data. Through discussions with the cluster investigators I designed and built the cluster monitoring database and contributed to the design of the cluster outcome reporting form. I designed and conducted the false positive TB detection surveys, after which I cleaned, analysed and interpreted the data. I was involved in the initiation of the contact tracing yield, diagnostic delay and rate of cluster growth studies, leading the discussions around them. I cleaned and merged the datasets for the studies. I conducted the analysis and interpreted the data for the contact tracing yield and diagnostic delay studies. I conducted the exploratory analysis for the rate of

cluster growth study, was involved in the design and interpretation of the main analysis, which was conducted by Ted Cohen and Leonid Chindelevitch. I was involved in the discussions about the structure of the transmission model and the parameter development, and conducted analyses for certain parameters. I was involved in the discussions around the design and structure of the cost-effectiveness model, I collected the cost data, conducted analyses to define certain parameters and cross-checked the sensitivity analysis. I was involved in the discussion of how the transmission model would input into the cost-effectiveness model and was involved in the interpretation of the model outputs.

I drafted the recommendations of the evaluation and was part of the discussion with the expert steering group to finalise them. I wrote the discussion to this thesis.

Positionality

Unless otherwise stated, the research in the thesis was carried out by the author – Jessica Mears – who studied Public Health (MSc) at the London School of Hygiene and Tropical Medicine and Experimental Psychology (BSc) at the University of Oxford. The evaluator had the key generic public health skills and thinking to be able to bring the comprehensive evaluation of a complex intervention together into a coherent thesis. Given the range of skills required for such an evaluation, input from highly specialised professionals was necessary, for example, mathematical modeling, economics and clinical expertise were all needed for individual components of the work.

Originally employed by PHE as the TB-STS Evaluation Scientist, the position was converted to a PhD studentship after the first six months. The initial time at PHE was useful in developing relationships between those working in the TB-STS, having access to internal documentation, and understanding the service. The change in role from PHE scientist to PhD student enabled the positioning of the evaluation to move outside of PHE and into UCL.

The motivation to conduct this research was based on the decision that interventions and the allocation of public funds should be evidence based, and because the author and steering group (Box 1) had genuine equipoise; the available evidence did not suggest whether the TB-STS was going to be a cost-effective intervention or not. Having been recently appointed to the role at PHE and quickly moving to UCL, minimised potential conflicts of interest.

An external, independent, multi-disciplinary evaluation steering group was set up, consisting of a TB nurse, chest physician, Consultant in Communicable Diseases (based in an HPU), TB program directors, epidemiologists, public health specialists, microbiologist, mathematical modellers, and a health economist. The names, disciplines and affiliations of the steering group are listed below. The evaluation steering group acted as an advisory group for the evaluation of the TB-STS, and for two specific elements, the transmission dynamic and cost-effectiveness modelling, undertaking the analysis. They were also used to check concepts, proposed methodologies, and the interpretation of qualitative and quantitative analyses, and ensure the findings of the evaluation remained independent. The main outputs of the steering group were the evaluation interim report (December 2011) and final evaluation report presented to PHE (March 2013), which included the recommendations for the TB-STS. Members of the evaluation steering group were co-authors on the peer-reviewed publications, led by the candidate, subject to their contribution to the papers. All three PhD supervisors were members of the evaluation steering group.

Box 1 – The Evaluation Steering Group (name; expertise; institution)

Prof Ibrahim Abubakar; Public Health Epidemiologist; University College London, Medical Research Council, and Public Health England.

Prof Martien W Borgdorff; Public Health Epidemiologist; Public Health Service (GGD) Amsterdam and University of Amsterdam.

Prof Ted Cohen; Mathematical modeler and epidemiologist; Harvard School of Public Health.

Debbie Crisp; Lead Tuberculosis Nurse Specialist; George Elliot NHS Trust.

Dr John A Innes; Honorary Consultant Physician; Heart of England NHS Foundation Trust.

Dr Mike Lilley; Consultant in Communicable Disease; South Midlands and Hertfordshire Public Health England Centre.

Dr Joanne Lord; Health Economist; Brunel University.

Dr Helen Maguire; Epidemiologist; Public Health England.

Prof Tim D McHugh; Medical Microbiologist; University College London.

Dr Pam Sonnenberg; Epidemiologist and Chair of Evaluation Steering Group; University College London.

Dr Emilia Vynnycky; Mathematical modeler; Public Health England and London School of Hygiene and Tropical Medicine.

Publications and conference presentations resulting from this PhD

Three peer reviewed journal articles have been published from the work done as part of this PhD:

1. Mears J, Abubakar I, Crisp D, *et al.* Prospective evaluation of a complex public health intervention: lessons from an initial and follow-up cross-sectional survey of the tuberculosis strain typing service in England. *BMC Public Health* 2014; **14**: 1023.

2. Mears J, Abubakar I, Cohen T, McHugh TD, Sonnenberg P. Effect of study design and setting on tuberculosis clustering estimates using Mycobacterial Interspersed Repetitive Units-Variable Number Tandem Repeats (MIRU-VNTR): a systematic review. *BMJ Open* 2015; **5**: e005636.

3. Mears J, Vynnycky E, Lord J, *et al.* The Prospective Evaluation of the Tuberculosis Strain-Typing Service in England: A Mixed Methods Study. *Thorax* 2015; published online April 16.

Two abstracts were accepted for poster presentations:

1. Mears J, Vynnycky E, Lord J, et al. Evaluation of the Tuberculosis Strain Typing Service (TB-STS) in England. The Lancet Public Health Conference 29th November 2013, London, UK. Abstract published in The Lancet; 382, Supplement 3: S73.

2. Mears J, on behalf of the TB Strain Typing Evaluation Group, Evaluation of the National TB Strain Typing Service: Results from a Baseline Survey. National Health Protection Conference, 13th-14th September 2011, Warwick, UK

The accepted manuscripts, abstracts and posters can be found in Appendix 7.

Acknowledgements

I would like to acknowledge my primary supervisor, Pam Sonnenberg, who has provided support, inspiration, and critical thought throughout my PhD. Her supervision has spread across many aspects of my life, including motherhood, rule bending and priority setting. I am immensely grateful for her consistent support whilst balancing the competing demands of being brilliant.

Thank you to my secondary supervisor, Ibrahim Abubakar, for providing so much guidance, especially during my first year, and being an unlimited source of solutions throughout the PhD. It has been a pleasure to work with you. Thank you to Tim McHugh, my tertiary supervisory, who kept us all on the straight and narrow – bringing expertise on all things laboratory and post-graduate protocol. Thank you to all three supervisors for working so well together to support this PhD.

Thank you to the members of the TB-STS Evaluation Group, who provided multidisciplinary expertise and advice to the evaluation. You were all fantastic to work with. Emilia Vynnycky, Jo Lord, Ted Cohen and Martien Borgdorff were particularly generous with their time, expertise and critical thinking – thank you.

Thank you to the TB section at Public Health England, especially Laura Anderson and the Cluster Investigators; Tim Brown at the National Mycobacterium Reference Laboratory, Jason Evans at the Birmingham Reference Laboratory, and Debbie Osborne at the Newcastle Reference Laboratory; Jacqui White and Sue Dart from North Central London TB Services; and Gerrit Woltmann from Leicester TB Services, for providing data for the evaluation. Thank you to the health protection staff, physicians and TB nurses who responded to the cross-sectional surveys, the source laboratories who responded to the false positive TB isolation survey and the HPU interviewees, for being part of this evaluation.

Thank you to my examiners, Prof Bertie Squire and Dr James Lewis, who gave me an appropriately challenging and insightful viva. I really appreciated the time they

took to engage with this thesis and the formal and supportive viva experience created for me. Thank you also to Prof Judith Glynn who examined by upgrade and provided excellent guidance for the direction of the thesis.

I would like to acknowledge Public Health England and the UCL Impact Fund for supporting this PhD.

Finally, thank you for the inspiration and steadfastness from my fellow PhD students and the axis of pregnancy (Ellie and Vicky, you started it); thank you to Noah for being a wise and wonderful friend; to my family for being wildly supportive; and to Ted and Birdie for being Ted and Birdie.

Chapter 1. Introduction

This chapter sets the scene for the thesis. An overview is provided of the global burden of tuberculosis (TB), the epidemiology of TB in the UK, the natural history of TB, TB control and strain typing. The public health applications of strain typing are outlined, existing evaluations of strain typing services and potential evaluation frameworks are described. Finally, the aims and objectives of the thesis are presented. This chapter provides the background to the studies in this thesis and the literature presented here is referred to throughout the thesis.

—

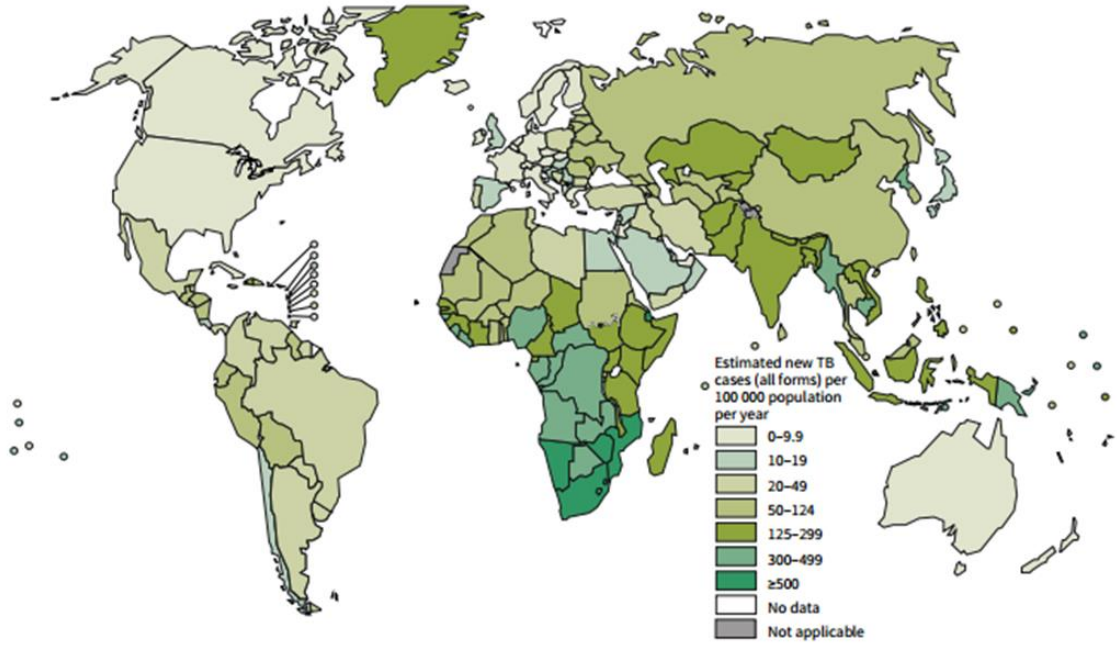
1.1 Tuberculosis (TB)

Global burden of TB

TB is a global disease, with an estimated 9 million cases in 2013 of which 1.5 million people died.[1] TB is the second leading cause of death from an infectious disease worldwide, after human immunodeficiency virus (HIV). TB mortality is unnecessarily high given that most deaths are preventable if people can access the healthcare required for diagnosis and treatment. Since TB was declared as a global public health emergency by the World Health Organization (WHO) in 1993, the TB mortality rate and incidence rate have been decreasing, but remain high with between and within regional variation (Figure 2).

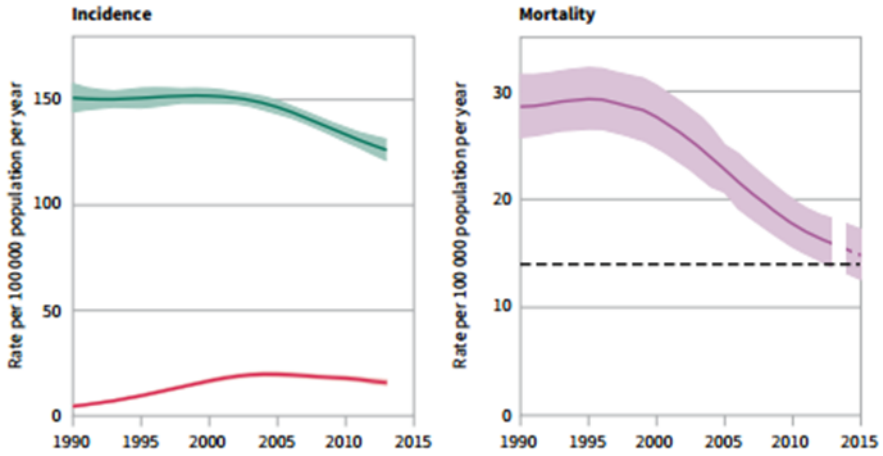
The TB incidence rate in the WHO European Region was estimated to be 39.4 per 100,000 in 2012, with an estimated 353,000 new cases.[2] This represents about 4% of the global burden of incident TB. 68,423 TB cases were notified in the European Region in 2012 (Figure 3). Of these, 27% (n=18,358) were of foreign origin (Figure 4). In the UK, the proportion of TB cases notified in 2013 that were foreign born was 72.5% (5518/288).[3]

Figure 1 – WHO estimates of TB incidence rates, 2013



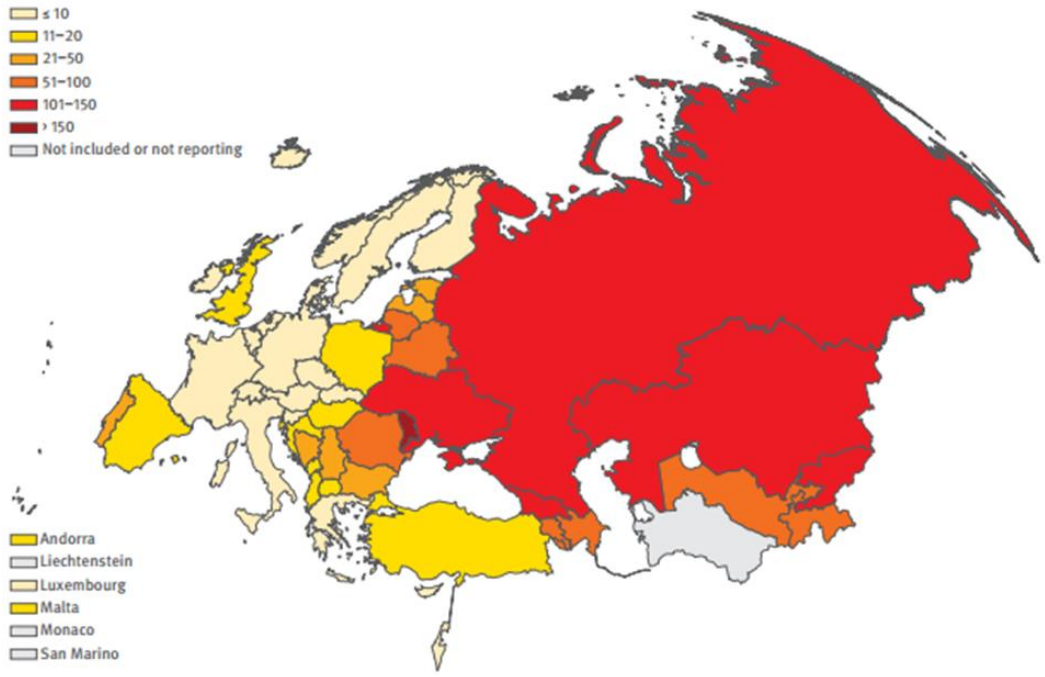
Source: Global Tuberculosis Report, World Health Organization 2014[1]

Figure 2 – Global trends in estimated rates of TB incidence and mortality



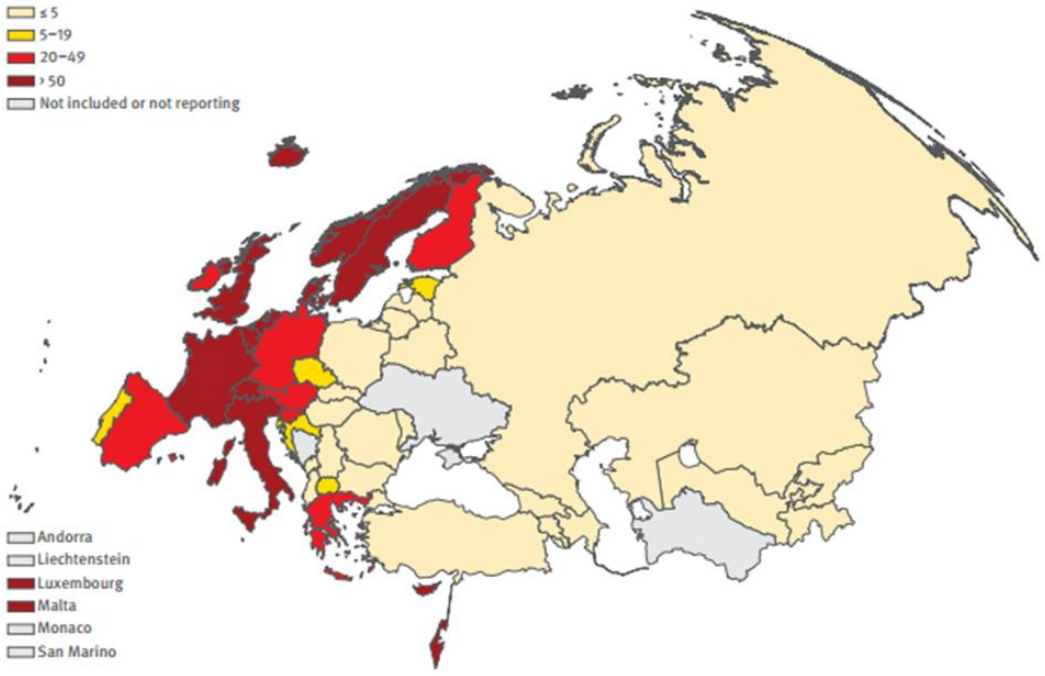
Left: Global trends in estimated overall TB incidence rate (green) and estimated incidence rate of HIV-positive TB (red). Right: Trends in estimated TB mortality rates 1990–2013 and forecast TB mortality rates 2014–2015. The horizontal dashed line represents the Stop TB Partnership targets of a 50% reduction in mortality rates by 2015 compared with 1990. Shaded areas represent uncertainty bands. Mortality excludes TB deaths among HIV-positive people.
Source: Global Tuberculosis Report, World Health Organization 2014[1]

Figure 3 – TB notification rates per 100 000 population, European Region, 2012



Source: European Centre for Disease Prevention and Control/WHO Regional Office for Europe. Tuberculosis surveillance and monitoring in Europe 2014.

Figure 4 – Percentages of notified TB cases of foreign origin among all TB cases, European Region, 2012

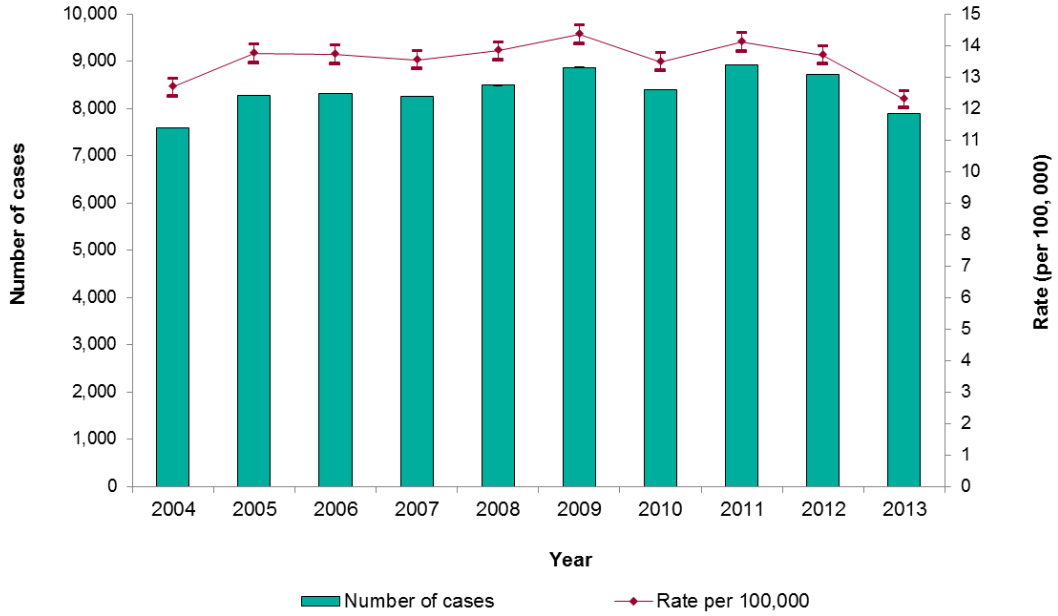


Source: European Centre for Disease Prevention and Control/WHO Regional Office for Europe. Tuberculosis surveillance and monitoring in Europe 2014.

Epidemiology of TB in the United Kingdom (UK)

TB has re-emerged as a serious public health problem in the UK over the last two decades with case notifications reaching 9000 in 2011, a rate of 14.1 per 100,000.[4] There has been a slight decrease in the numbers of TB cases in the last two years, with 7892 cases notified in 2013, a rate of 12.3 per 100,000.[3] Approximately half of which were pulmonary (52% of cases with known site of disease).

Figure 5 – TB case reports and rates in the UK, 2004-2013



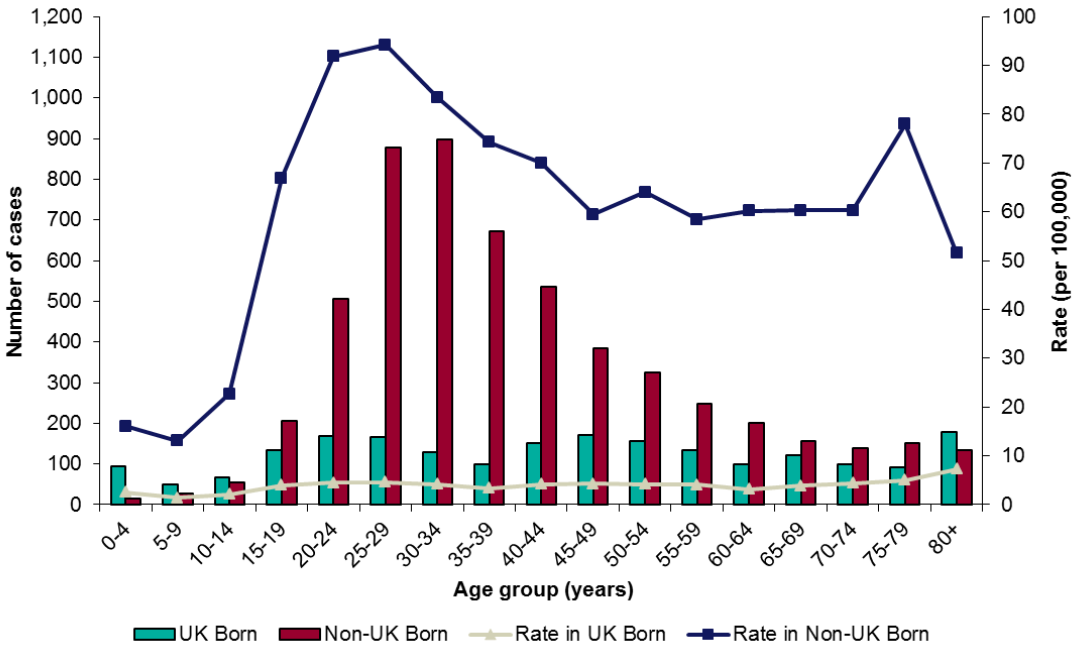
Source: Enhance Tuberculosis Surveillance (ETS), Enhanced Surveillance of Mycobacterial Infections (ESMI), Office for National Statistics (ONS)[3]

England accounted for 92.4% (7290/7892) of the cases in the UK in 2013 with a rate of 13.5 per 100,000, the highest proportion of which were in London (rate of 35.5 per 100,000). 59.2% of cases in England were culture confirmed and 71.3% of cases with pulmonary TB were culture confirmed.[3]

A majority of cases were male (58% in 2013), and most were aged between 15 and 44 years old (with a rate of 24.4 per 100,000 in this age group in 2013) (Figure 6).[3] Rates of TB were significantly higher in non-UK born cases compared to UK born

cases (70 per 100,000 compared to 4 per 100,000) (Figure 6). TB incidence rates are highest in those aged 75 and older in the UK born population, whereas in the non-UK born population the highest rates are in those aged 25 to 29 (rates of 6.6 and 94.3 per 100,000, respectively). In the last few years, the most common countries of origin of non-UK born cases are India, Pakistan and Somalia (30%, 20% and 5%, respectively in 2013).[3]

Figure 6 – TB case reports and rates by age group and place of birth, UK, 2013



Source: Enhance Tuberculosis Surveillance (ETS), Enhanced Surveillance of Mycobacterial Infections (ESMI),Labour Force Survey (LFS)
Data as at: May 2014. Prepared by: TB Section, Centre for Infectious Disease Surveillance and Control, Public Health England

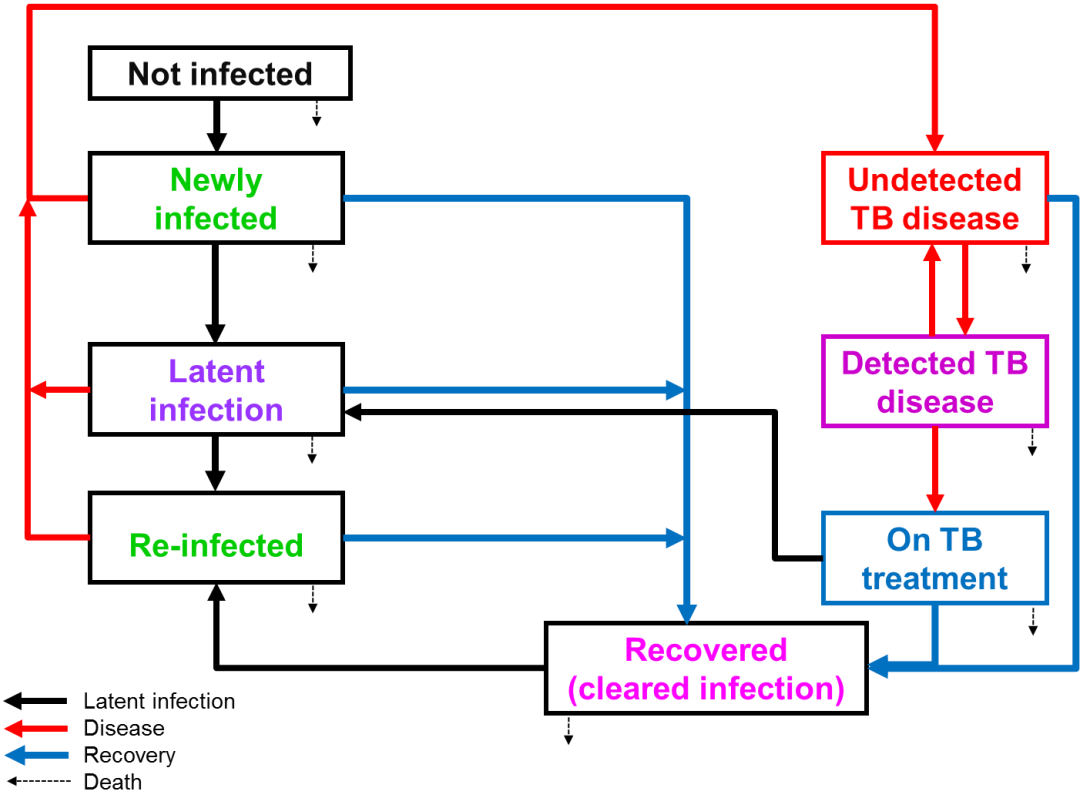
In 2013, 7% of cases in the UK had a previous diagnosis of TB.[3] Resistance to at least one first line antibiotic was identified in 7.8% of all cases, and 1.6% were multi-drug resistant (MDR; resistant to at least isoniazid and rifampicin).[3] These rates have been relatively stable over the last decade, with a slight increase between 2008 and 2011.

Approximately 10% of TB cases have at least one social risk factor for TB: current or history of problem drug misuse, alcohol misuse, homelessness or imprisonment.[3]

Natural history of TB

TB is caused by *Mycobacterium tuberculosis* complex. After inhaling aerosolised droplets of *M. tuberculosis* from a person with infectious TB, approximately 30% of people become infected.[5] In the remainder, the infection is cleared, which is likely due to host factors (such as the host's immune response which may eradicate the organism) or characteristics of the pathogen (such as the virulence of the strain). A simplified interpretation of the natural history of TB is illustrated in Figure 7.

Figure 7 – Natural history of TB



A person who is infected with *M. tuberculosis* may clear the infection and recover (follow the blue arrows in Figure 7), may become latently infected (black arrow), or develop active disease (red arrow). When an individual is infected with *M. tuberculosis* and the bacilli are in a dormant state, this is known as latent tuberculosis infection (LTBI), whereby the patient will show no clinical signs of TB.[6] This latent period can last a lifetime or the infection may reactivate to become active disease at any point, known as endogenous reactivation. Risk factors for reactivation include

age and a suppressed immune system.[7] The variable duration of latent infection of *M. tuberculosis* makes it difficult to identify transmission pathways and implement targeted control programmes.

Most immune competent individuals will not develop TB disease over their lifetime. The lifetime risk of progressing to active disease is 5-10%. It is highest shortly after infection (80% in the first two years) and decreases subsequently. The risk of developing disease is much higher for those who are immunocompromised, especially those with HIV, who have a 5-10% annual risk of progressing to active TB.[6]

People who are latently infected, or people who have recovered from a previous TB infection can be reinfected with the same or different strain of *M. tuberculosis*. This is known as exogenous infection, or reinfection. The variability between the time of infection and disease makes differentiating between endogenous and exogenous TB difficult, as there could be many years between infection from the index case and disease in the secondary case.[8] Molecular typing can be used as a tool to differentiate between relapse and reinfection (see section 1.3.2, page 21).[9]

Following treatment for active TB, an individual may clear the infection and recover, the treatment may fail and the individual may continue to have active disease, the individual may relapse at a later stage and develop recurrent TB. As disease does not infer immunity, the individual remains at risk for reinfection and subsequent disease.

1.2 TB control

The aims of TB control are to provide care to the infected individual that will reduce morbidity and mortality, and ensure the best possible prognosis, and prevent or reduce transmission in households and the community. Given what is known about the natural history of TB, identifying people who have recently been infected and treating them with chemoprophylaxis is a key intervention to prevent progression to active disease and the subsequent potential onward transmission. The main methods

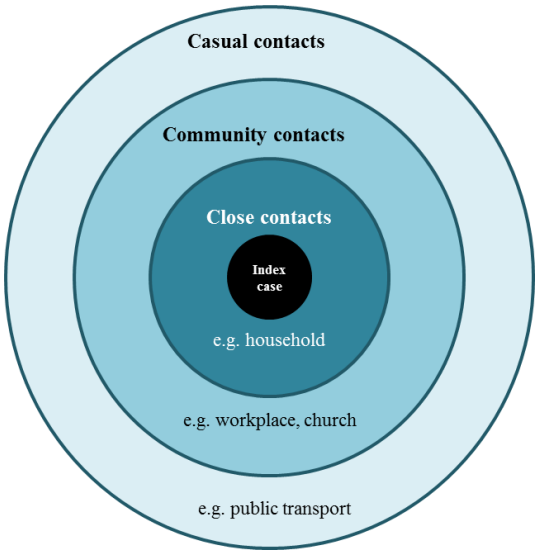
for TB control are through case finding (active and passive), treatment of cases and infected contacts, and vaccination with BCG.

Case finding

Active case finding is the active search for new cases of TB or LTBI, as opposed to passive case finding which relies on symptomatic cases presenting to the healthcare system. Cases of active TB are identified and treated, thereby preventing onward transmission, and people with LTBI are identified and treated prophylactically to prevent progression to disease.

The two most common forms of active case finding are contact tracing and screening (which is usually targeted to high risk groups, such as migrant entry screening and screening of hard to reach populations).[10,11] Contact tracing forms the basis of a traditional public health response to a case of TB in order to identify others who might have been infected, i.e. if there has been any onward transmission and in some instances to identify the source case, and to target individuals at high risk of TB infection or disease.[12] Contact tracing is based on the 'stone in the pond' principle,[13] whereby the close contacts of a TB patient are identified and invited to be screened (Figure 8). If a route of transmission has been identified and the need for further investigation is required, this process is then extended to a wider circle of contacts.

Figure 8 – The 'stone in the pond' principle for TB contact tracing



Once contacts have been identified through contact tracing, they are screened for TB. Screening involves testing for active TB or LTBI. Active TB is suspected based on symptoms and/or chest x-ray if the disease is pulmonary. In brief, laboratory diagnosis is confirmed by microscopy, culture or histopathology on specimens such as sputum, bronchial washings, cerebrospinal fluid or lymph nodes. Because of the high mycolic acid content of the mycobacterial cell wall, acid-fast bacilli can be stained and detected from a sputum sample using an arylmethane staining and acid-based de-colouring technique (such as the Ziehl-Neelsen stain). Diagnosis can also be confirmed using *in vitro* culture (Lowenstein-Jensen is the most commonly used egg-based formulation) or liquid culture (such as BACTEC Mycobacterium Growth Indicator Tube (MGIT) 960). Due to the long division time of *M.tuberculosis* (16 to 18 hours in optimal conditions), it may take two to four weeks to become culture positive.[14] More rapid diagnostic tests such as nucleic acid amplification tests are also available, but are expensive and do not replace the need for culture. The host's basic inflammatory response to infection with *M.tuberculosis* is the formulation of tuberculoid granulomata, which is diagnostic on histology.

Screening for LTBI can be done with a Tuberculin Skin Test (TST) or an interferon gamma release assay (IGRA) test. Contact tracing focuses on contacts from the household and non-household (such as the workplace and leisure settings). These two contact types are most easily explored through the traditional patient interview. However, assessing casual contacts, whom the patient may have never formally met before, poses a problem for traditional contact tracing as the patient is unlikely to name these people as their contacts.

Contact tracing is not restricted to TB incidents, but is used as an infection control strategy in response to other infectious diseases such as HIV, other sexually transmitted infections (STIs) and pandemic influenza. It is relatively easy to establish which contacts have been exposed to an STI because the contact required for transmission is easily defined. On the other hand, influenza is so infectious and has a short infectious period that contacts that have been exposed to the virus can be defined very inclusively – as anyone that one has spoken to in the last three days, for

example.[15] TB, on the other hand, can have a very long and variable incubation period and the infectiousness is relatively low. This makes defining a contact more difficult and the development of transmission models for TB challenging.

Indicators for the effectiveness of active case finding include the yield of active TB and LTBI identified through contact tracing, known as contact tracing yield, and the delay between the onset of symptoms and start of treatment, known as diagnostic delay. Contact tracing yield and diagnostic delay are described in turn.

Contact tracing yield

Contact tracing yield is defined as the proportion of contacts of an index case who have LTBI or have active TB disease. Fox *et al* (2013) conducted a review of the prevalence of LTBI and active TB, and the annual incidence of TB among contacts of patients with TB in low, medium, and high-income countries.[16] From high-income countries, they included 92 studies reporting the yield of LTBI (79,511 individuals with LTBI after screening 284,505 contacts; 28.1%, CI 24.3 to 32.4, I^2=99.5) and 87 studies reporting the yield of active TB (5058 active cases identified from 308,048 contacts; 1.4%, CI 1.1 to 1.8, I^2=98.7). Interpreting the results of a meta-analysis such as this is challenging due to the heterogeneity of the studies included. For example, this review included studies of different transmission risk groups (such as HIV-positive and drug-resistant TB patients). In addition, the definition of a contact varied largely between studies, with some only including 'close contacts', or 'household contacts', or 'casual contacts', and or contacts from congregate settings.

Another literature review, assessing the contact tracing yield amongst migrant and foreign-born index cases, included 70 studies and estimated the yield of LTBI to be 31.9% (IQR 16.9-36.9) and active disease to be 0.4 (IQR 0.0-2.2).[17]

To gather studies more relevant to the TB epidemiology and TB control programme in England, the references from Fox *et al*'s systematic review were examined[16] and a search of PubMed was conducted to identify more recently published articles from low TB incidence countries. The following search terms were used: "tuberculosis"

and “contact screening” or “contact tracing” or “contact investigation”. The identified articles are summarised in Table 1. There were four relevant studies from the UK, presenting data from Birmingham, Blackburn, London and South Glamorgan, spanning from 1982 to 2010. The contact tracing yield ranged from 3.0% to 7.9% for LTBI identified and 0.7% to 2.9% for active TB found. Other studies identified in the search were from France, Italy, the Netherlands and the USA.

The mean number of contacts identified and screened per index case was highest in the Netherlands (27.7 and 23.7, respectively for LTBI and active TB)[18] and lowest in Turin, Italy (4.3 and 4.0, respectively).[19] The highest yield of LTBI was in Maryland, USA (41% of those screened had LTBI) and the highest yield of active TB was in the Tower Hamlets in London (3.8% of contacts of pulmonary TB index cases screened had active TB)[20].[21] Yield of LTBI and active disease in the studies summarised in Table 1 ranged from 3% to 41% and 0.2% to 3.8%, respectively.

Table 1 – Summary of identified literature on contact tracing yield in low TB incidence countries

Author	Country	Study location	Study period	Index cases[a]	Contacts (mean per index case)	Contacts screened (coverage[b])	LTBI screening test[j]	LTBI (%)	Active TB (%)
Ansari[22]	UK	South Glamorgan	1992-94	103	732 (7.0)	707 (96.6)	H/X	21 (3.0)[d]	7 (1.0)
Ormerod[23]	UK	Blackburn	1982-90	649	.	7017	TST	.	50 (0.7)
Saunders[24]	UK	Birmingham	1990-2010	7365	46158 (6.3)	42613 (92.3)	TST/IGRA	1687 (4.0)	778 (1.8)
Saunders[24]	UK	Birmingham	1990-2010	5056[c,i]	35095 (6.9)	31852 (90.8)	TST/IGRA	2220 (7.0)	718 (2.3)
Underwood[21]	UK	Tower Hamlets	1997-99	227	.	646	TST/X	51 (7.9)	18 (2.8)
Underwood[21]	UK	Tower Hamlets	1997-99	144[c,i]	.	419	TST/X	29 (6.9)	16 (3.8)
Aissa[25]	France	Paris	2004-05	325	2009 (6.2)	1575 (78.4)	TST/X	410 (26.0)	15 (1.0)
Borraccino[19]	Italy	Turin	2001-08	1099	4759 (4.3)	4441 (93.3)	TST	1287 (29.0)	22 (0.5)
Mulder[18]	NL	NL	2006-07	642	17124 (27.7)	10391 (60.7)	TST/IGRA	841 (8.1)	.
Mulder[18]	NL	NL	2006-07	642	17124 (27.7)	15202 (88.8)	TST/IGRA	.	91 (0.6)
Sloot[26]	NL	Amsterdam	2008-11	235[c]	3743 (15.9)	2337 (62.4)	TST/X/IGRA	254 (10.9)	.
Sloot[27]	NL	Amsterdam	2002-11	610[c]	9332 (15.3)	4774 (51.2)	TST/X/IGRA	739 (15.5)	.
Behr[28]	USA	San Francisco	1991-96	1525[c,g]	.	11211	TST/X	3976 (35.5)	109 (1.0)
Davidow[29]	USA	5 TB programmes	1996	42[h]	724 (17.2)	494 (68.2)	TST	144 (29.1)	1 (0.2)
Golub[20]	USA	Maryland	2000-01	124	.	703	TST	288 (41.0)	16 (2.3)
Jereb[30]	USA	29 states	1999	9199	67585 (7.3)	56100 (83.0)	.[k]	13083 (23.3)	561 (1.0)
Marks[31]	USA	11 TB programmes	1996-97	1080[c,e]	.	6225[f]	TST	1725 (27.7)	134 (2.2)
Moran-Mendoza[32]	USA	British Columbia	1990-2000	3485	42593 (12.2)	33146 (77.8)	TST	.	228 (0.7)
Reichler[33]	USA	5 health departments	1996	349[c]	3824 (11.0)	2095 (54.8)	TST	655 (31.3)	24 (1.1)
Sprinson[34]	USA	California	1999-2000	2032	17774 (8.7)	15582 (87.7)	.[k]	4609 (29.6)	111 (0.7)

NL The Netherlands

[a] all TB, unless otherwise indicated; [b] coverage is the proportion of contacts identified that were actually screened (contacts screened/contacts identified*100); [c] pulmonary TB

[d] given chemoprophylaxis; [e] smear positive; [f] close contacts (does not include casual contacts); [g] paediatric TB; [h] cases in a workplace setting; [i] sub-analysis of pulmonary TB cases; [j] H Heaf test; X Chest x-ray; TST Tuberculin skin test; IGRA Interferon Gamma Release Assay; [k] the screening test was not reported by Jereb *et al* or Sprinson *et al*, but based on information in the article it can be inferred that TSTs were used

Diagnostic delay

Reducing diagnostic delay, the delay between onset of symptoms and diagnosis or treatment, is an important feature of a TB control service. During this time, when a symptomatic TB case is undiagnosed, they are at risk of transmitting the infection within the community and progressing to more severe disease. An effective TB control programme will provide early diagnosis and initiation of treatment, to improve the prognosis for the individual as well as preventing onward transmission. Diagnostic delay is used as a measure of the quality of a TB control programme. The total delay can be divided into patient delay (time between onset of symptoms and contact with the health service) and healthcare delay (time between contact with the health service and diagnosis or treatment).[35–37]

Although diagnostic delay can be useful, it is inherently biased as it relies on the patient's subjective assessment of the symptoms and their recall for when the symptoms started. Treatment start date is a more objective variable to estimate the end point of the diagnostic delay, but it is not recorded reliably in the UK surveillance dataset. Instead, the notification date is a more complete variable, but it may be subject to time lags as the administrative task of entering new cases onto the surveillance database may not be completed prospectively, but done retrospectively on a regular basis (e.g. monthly or quarterly). As a result, it is difficult to compare diagnostic delay across different countries as the data collection and notification biases will vary. In addition, there may be cultural differences in the subjective assessment of when symptoms, such as a cough, first started.

A brief review of the literature was conducted to explore the diagnostic delay experienced in the UK. PubMed was searched using the search terms (tuberculosis) and ("diagnostic delay" or "treatment delay" or "delayed diagnosis" or "delayed treatment") and (UK or England or London). Ten papers were identified that reported a measure of delayed diagnosis of TB patients in the UK. One additional report was identified from the grey literature.[3] Five studies were retrospective cohort studies, five studies were cross-sectional studies based on surveillance data, and one used

national surveillance data. Two systematic reviews containing three studies conducted in the UK were identified,[37,38] all three of which had already been identified by this search. Table 2 summarises the studies identified by this search.

The median diagnostic delay reported across the nine studies ranged from 37 days[39] to 126 days.[40] Excluding the studies that only included children[39] and women whose symptoms started during pregnancy,[41] the shortest diagnostic delay reported was 49 days.[42,43]

Table 2 – Summary of literature on diagnostic delay in the UK

Author	Study period	Study location	Study population[a]	Number of cases	Definition of diagnostic delay[b]	Median (IQR) diagnostic delay (days)
Abubakar[39]	1999 - 2006	England and Wales [b]	Children	3563	treatment	37 (13-89)
Craig[44] [c]	Jan 1999 - Aug 2002	Royal Free TB clinic		120	treatment	.
French[45]	2000-2005	England		40779	either	67 (30-131)
Lewis[40]	Jan 1996 - Oct 1997	Newham Chest Clinic		93	treatment	126 (4-1533)
Llewelyn[41]	Dec 1995 - May 1998	Northwick Park Hospital	Pregnant women	13	diagnosis	49 (14-210) [d]
Paynter[46]	April 2001 - March 2002	North Middlesex University Hospital		72	treatment	78 (39-159)
Public Health England[3]	2013	UK		3009	treatment	72 (36-132)
Rodger[42]	1998 - 2000	London		1355	treatment	49 (14-103)
Saldana[35]	2007	Thames Valley		273	treatment	73 (65-89)
Sultan[47]	Jan 2005 - Oct 2010	West Midlands		4840	treatment	81.5 (40-154)
Wares[43]	1991-1996	Blackburn Royal Infirmary		43	diagnosis	49 (28-91)

[a] The study population was all TB cases unless otherwise stated

[b] Definition of diagnostic delay was the time between onset of symptoms and: the start of treatment (treatment), the date of diagnosis (diagnosis), or whichever was available (either)

[c] Craig *et al* did not report the median total diagnostic delay. Instead, they reported separately the time between symptom onset and first presentation to primary care (46 (55%) < 3 months, 29 (35%) > 3 months) and time from first presentation to primary care to start of anti-TB treatment (62 (75%) ≤ 8 weeks, 21 (25%) > 8 weeks)

[d] The median and range are reported here (IQR was not reported in the paper)

TB Treatment

The standard treatment for active TB consists of six months of isoniazid and rifampicin supplemented in the first two months with pyrazinamide and ethambutol.[48] Any resistance to these first line drugs will be treated with a different combination of drugs, usually excluding the drug to which the organism is resistant, for up to 18 months. Multi-drug resistant tuberculosis (MDR-TB) (resistant to at least rifampicin and isoniazid) is treated with a combination of first and second line drugs. Second line drugs are divided into six main classes: aminoglycosides, polypeptides, fluoroquinolones, thioamides, cycloserine, and para-aminosalicylic acid. Four to six drugs from separate classes to which the organism is susceptible are selected to treat MDR-TB. Treatment is tailored to suit the characteristics of the patient and can last up to two years.

Prophylactic treatment for LTBI consists of either six months of isoniazid or three months of rifampicin and isoniazid. If a person has been in contact with a drug resistant case of TB, they may be treated prophylactically with six months of rifampicin. In the UK, prophylactic treatment is recommended for those aged 35 and below (because of increasing risk of hepatotoxicity with age), those HIV positive (any age), or healthcare workers (any age).[48]

BCG vaccination

The bacillus Calmette–Guérin (BCG) vaccine, a live vaccine derived from an isolate of *Mycobacterium bovis*, was first developed in the 1920s. The efficacy of the vaccine varies worldwide, but is most effective for TB meningitis.[49] This might be because BCG does not protect against disease when it is given to people who are already infected or who have been sensitised to environmental mycobacteria. Unlike the recently developed IGRA test, the tuberculin skin test cannot distinguish a positive response caused by *M. tuberculosis* infection from that caused by BCG vaccination or non-tuberculous mycobacterial infection.

Since 2005, the UK's vaccination strategy has been to target those at highest risk of exposure to TB:[48]

- People who have been in close contact with someone with pulmonary TB, who had a negative Mantoux test, have not been vaccinated before, and are aged 35 years or younger
- All infants (0-12 months) living in areas where TB incidence is at least 40 per 100,000 or with a parent or grandparent born in a country where TB incidence is at least 40 per 100,000
- Children who were not vaccinated as babies and are at high risk of TB
- People coming to live in the UK from areas where TB is widespread, if they have a negative Mantoux test result, have not been vaccinated before, and are aged 35 years or younger

1.3 Strain typing

The complex natural history of TB has left classical epidemiology with many unanswered questions including how to differentiate between recent transmission and reactivated previous infection; and to determine whether recurrent TB is due to relapse (endogenous) or re-infection (exogenous). Furthermore, there are unanswered questions about differences in the epidemiology of different lineages of TB, and the molecular evolution of *M. tuberculosis*.[8,50–52] Molecular epidemiology – "the use of molecular-typing methods for infectious agents in the study of the distribution, dynamics, and determinants of health and disease in human populations"[53] – has started to address these questions. In combination with classic epidemiology, differentiating *M. tuberculosis* isolates by strain type can improve our knowledge and understanding of the natural history of TB and TB transmission.

1.3.1 Strain typing methods

In the last decade the three main methods commonly used to distinguish between different strains of *M. tuberculosis* are insertion sequence IS*6110* restriction fragment length polymorphism (RFLP),[54] spacer oligonucleotide typing (spoligotyping)[55] and mycobacterial interspersed repetitive unit-variable number tandem repeats (MIRU-VNTR).[56] These methods are based on sequences of DNA that are found interspersed throughout the bacterial genome or at specific locations

on the genome.[8] More recently, whole genome sequencing (WGS) examines variation across most (~90%) of the bacterial genome at the nucleic acid level.[50,57]

The most appropriate method to detect recent transmission in a population will depend on whether the method can discriminate between strains that are not linked epidemiologically, whether the 'molecular clock' of the biomarker is slow enough to identify cases in the same chain of recent transmission and the genetic diversity of *M. tuberculosis* in the population.[8] A reliable molecular strain typing technique will be stable, rapid, reproducible, highly discriminative, easy to perform and interpret, and applicable to clinical material/samples.[50,58]

IS6110 RFLP

Until recently the most widely used molecular typing method, considered the gold standard technique, was IS*6110* RFLP typing. IS*6110* is an insertion sequence found in varying locations and copy numbers in the *M. tuberculosis* genome. The polymorphisms that arise from the different locations and copy numbers of the IS*6110* allow *M. tuberculosis* to be grouped into different lineages. The standardised method for IS*6110* RFLP typing results in a different banding pattern for each strain which can be compared using comparison algorithms[54] such as those incorporated into the *Bionumerics* software.[59]

IS*6110* RFLP does not meet all the criteria for a good molecular typing tool for TB:[8,60,61]

- The process is laborious with a slow turnaround time (requiring several weeks to culture enough DNA);
- A proportion of *M. tuberculosis* isolates contain no or very few copies of IS*6110* insertions (low copy number (LCN)) so they cannot be distinguished using this technique (five or fewer are not discriminatory)[62], requiring a second typing method; and
- The interpretation and comparison of the banding patterns relies on the subjective opinion of the researcher.

Spoligotyping

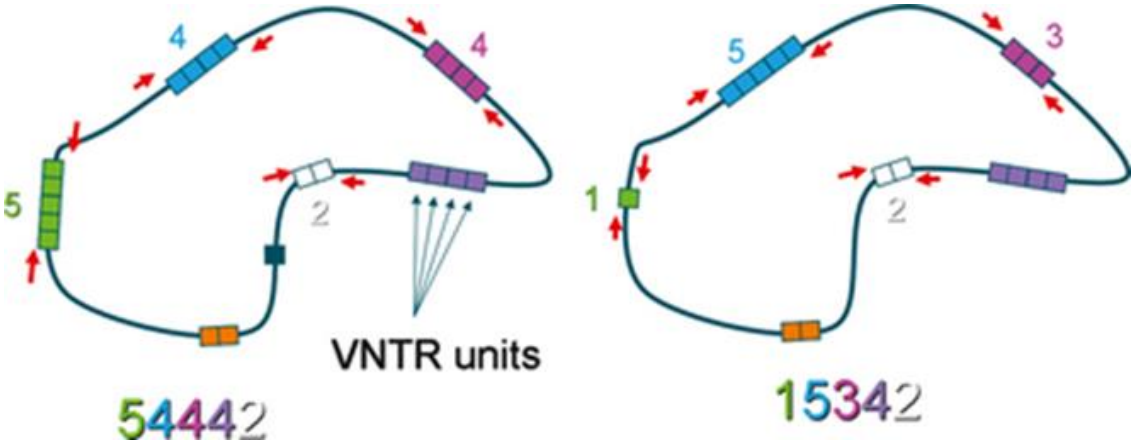
Spoligotyping is used in combination with other typing methods to provide a highly discriminatory typing system.[63,64] This method is based on the presence or absence of 43 unique, non-repetitive DNA 'spacer' sequences in a distinct chromosomal region, known as the Direct Repeat region, that are interspersed with 36-bp repetitive sequences (direct repeats).[55] The result is binary data based on the presence or absence of spacer sequences, which can be compared between strains. Spoligotyping can accurately discriminate between strains with a low copy number of IS*6110* and can easily detect lineages that are characterised by spacer deletions (such as Beijing strains which have 34 spacer deletions).[62,65]

The advantages of spoligotyping are that results are reproducible, easy to interpret and easily compared, and the procedure is rapid and requires small amounts of DNA. Disadvantages are that it has a lower discriminatory power compared to the other methods described here, and is subject to convergent evolution (whereby phylogenetically unrelated strains have the same spoligotyping patterns because of the irreversible deletion of spacer sequences).[66]

MIRU-VNTR

MIRU-VNTR typing is based on genetic elements found in the *M. tuberculosis* chromosome termed mycobacterial interspersed repetitive units (MIRU).[56] Supply *et al* described 41 MIRU-VNTR in the *M. tuberculosis* genome.[67] Each locus contains multiple copies of a repeated sequence; the variation in the number of copies at each site provides the basis for a unique identifier. Copy number is established by PCR amplification of the loci using primers specifically designed for the flanking region of each of the loci. The copy number can be expressed as a digit so that the copy numbers of all the loci typed can be recorded as a string of digits. The amplified loci are then analysed using computer software packages such as Bionumerics[59] or GeneScan[68] to build dendrograms for comparison.

Figure 9 – The principle of MIRU-VNTR typing of *M. tuberculosis.*



The size of each locus is measured using PCR amplification designed for the flanking region of each locus (indicated by the red arrows). Because the size of a unit is known, the number of repetitive units can be inferred by the size of the PCR product. As this is done for 24 loci, the strain type is presented as a 24-digit long numerical code.
Source: Borgdorff and van Soolingen, Clinical Microbiology and Infection, Volume 19, Issue 10, pages 889-901, 4 JUN 2013 DOI: 10.1111/1469-0691.12253

Evidence suggests that the discriminatory power of 24 MIRU-VNTR is slightly lower than IS*6110* RFLP,[69,70] but is more discriminatory for low IS*6110* copy strains.[71] Used in combination with spoligotyping, the discriminatory power of 24 MIRU-VNTR is comparable to IS*6110* RFLP in low and high TB incidence settings.[72,73]

Advantages of the MIRU-VNTR strain typing include the ease with which the numerical strain type can be interpreted and compared, the speed at which typing can be conducted, and the high throughput nature of the method. Disadvantages of MIRU-VNTR typing is that not all loci can be successfully amplified for every strain and that the 'molecular clock' of the biomarker is unclear. Early estimates suggested that the mutation rate of VNTR loci was extremely slow,[74] whereas more recent estimates are much higher,[75] resulting in some (unresolved) debate about whether recent transmission could be underestimated using this typing method.[76,77]

Whole Genome Sequencing (WGS)

The *M. tuberculosis* sequence was published in 1998[57] and since then sequencing methods have become more efficient and cost-effective. WGS identifies the whole

genome sequence from which Single Nucleotide Polymorphisms (SNPs), variations in single nucleotides arising from mutations on the genome, can be identified. Analysis of SNPs allows *M. tuberculosis* to be classified into strains based on SNP differences. SNP analysis shows how closely linked individuals are in time, and because backward mutations are rare, the direction of transmission can also be determined.[50,78–80] There is growing support for WGS to be used to detect chains of recent transmission and discriminate between reactivation and reinfection of TB.[81,82]

WGS has the highest discriminatory power for distinguishing between strains as it is based on the analysis of the genome at approximately 70% of nucleotide positions.[57] The 'molecular clock' of WGS SNP analysis is currently uncertain. There is some evidence that it is stable and the rate of change is rarely more than 5 SNPs in a three year period[78], whereas others have shown that the mutation rate to be more variable.[82,83] Studies have suggested that where MIRU-VNTR typing has clustered patients, WGS has divided these clusters, and where MIRU-VNTR typing has differentiated patients, WGS has found they are genetically linked.[78,81]

WGS is a useful instrument for understanding TB transmission and natural history, but its feasibility as a routine tool for public health is still being explored.[78] The cost of the technique and complexity of the bionformatic analysis needed are currently barriers to its wider use, requiring specialist software and highly skilled staff for data analysis and interpretation. It is expected that the cost will decrease over the next few years[50,84] and pipelines for analysis are being developed.[85]

1.3.2 Public health applications of strain typing

TB strain typing can be used for the following, each of which is discussed in more detail below:

- To estimate TB transmission;
- To prove or disprove suspected transmission;
- To identify risk factors for recent transmission;
- To interrupt transmission through targeted contact tracing;
- To distinguish between reactivation and reinfection; and

- To identify incidents of false positive TB isolation.

Estimating transmission

Strain typing can distinguish between clustered cases and unique cases. Clustered cases – cases with indistinguishable strains of *M. tuberculosis* – can be used to estimate the amount of recent transmission occurring in a population by assuming that one case in each cluster is the source case and the rest are part of a subsequent chain of transmission.[86] Estimates of recent transmission provide a useful measure to evaluate the success of TB control programmes as you would expect recent transmission to decrease following a successful intervention.[87–89] Strain typing can also help identify and investigate outbreaks or chains of ongoing transmission.[86,90,91] This occurs where cases that are not epidemiologically linked through contact tracing, are clustered by strain type. Strain typing can also be used to disprove epidemiological links between cases where transmission is thought to have occurred when strains are indistinguishable.[28]

Potential biases in the interpretation of TB strain typing need to be considered (this topic is reviewed systematically in Chapter 2 (page 31)). There is a difference between the number of patients that are linked epidemiologically through contact tracing and the number of patients in molecular clusters defined by strain typing, making it unclear what the true rate of recent transmission is. Studies in the United States have found low concordance between epidemiological links between patients and molecular clusters.[28,86] This could represent the underestimation of transmission identified through contact tracing, or the overestimation of recent transmission identified by strain typing. An alternative explanation is that it may be impossible to establish transmission through interviews with patients.[92] The disparity between epidemiological information and molecular information illustrates the importance of combining the two sources of information to inform TB control activities.

There is evidence to suggest that rates of clustering by IS*6110* RFLP are influenced by the diversity of strains in the population, TB incidence, study design and characteristics of the population.[93] No studies have examined whether the same

biases occur with MIRU-VNTR. This is addressed in detail by the systematic review in Chapter 2.[94]

Proving or disproving suspected transmission

Strain typing can be used to prove or disprove suspected transmission identified through contact tracing or a suspected outbreak. If two cases have an epidemiological link (e.g. if they are colleagues), then indistinguishable strain types provide further evidence for transmission. If a case is suspected of being part of a larger outbreak, a distinguishable (not matching) strain type will disprove their involvement in the outbreak.

This tool is more useful in settings with a low TB incidence. This is because in high TB incidence settings where the proportion of clustering is very high, direct links between patients confirming suspected transmission are only found in a small proportion of instances. For example, in a gold-mining community in South Africa, in most instances miners with the same IS*6110* RFLP strain of TB did not share the same room, so transmission was not occurring in the dormitories as had been expected.[95] In these circumstances contact tracing becomes a rudimentary method to screen the population and cases are identified, not because of recent transmission from the index case, but because they have acquired TB within the population more broadly.[28] In low incidence areas, however, strain typing can more reliably be used to identify or prove suspected transmission.

Risk factors for recent transmission

Strain typing has enabled epidemiologists to identify risk factors for recent transmission, based on their risk factors for being in a molecular cluster. This makes it possible to target high risk groups with *ad hoc* TB control interventions. Fok *et al* conducted a systematic review of molecular epidemiological studies using IS*6110* RFLP typing to investigate risk factors for clustering.[96] Thirty-six studies from 17 different countries of low, medium and high TB incidence identified male sex, young adults, native (vs. foreign-born), urban residence (vs. rural), alcohol and drug abuse,

homeless, exposure in crowded settings (including prisons), and pulmonary TB as risk factors for clustering.

Interrupting transmission

Cluster investigations are the extended contact tracing and investigation activity surrounding a molecular cluster of patients. They can be used to: identify epidemiological links between clustered cases; identify non-traditional transmission settings (such as bars and churches);[97] identify unsuspected outbreaks;[98–101] or help to target contact tracing investigations based on the characteristics of the first two cases in a cluster.[102,103]

Conversely, strain typing can be used to evaluate contact tracing and its effectiveness. For example, strain typing has identified cases that were missed by traditional contact tracing, such as contacts that were from outside of the household or close contact group, or hard to reach groups (such as the homeless or drug users).[89,104] In addition, strain typing has disproved suspected transmission, showing that a proportion of TB in household contacts have different strains of TB, ruling out household transmission.[28,105]

Relapse and reinfection

Recurrent TB may be caused by relapse or exogenous reinfection. Differentiating between relapse and reinfection is important for TB case management and control strategies.[81,106–108] Where historical isolates from previous TB episodes are available it is possible to distinguish between relapse and reinfection in a patient, based on the assumption that reinfection is likely to result from *M. tuberculosis* with a strain type that is different from the previous infection (whereas relapse will be the same strain). Identifying reinfection is important as it shows that cases in high burden settings are arising from new infections rather than reactivation.[109] Reinfection is demonstrated by evidence of mixed infections where one patient is infected with more than one strain of *M. tuberculosis*.[110,111] The inference that *M. tuberculosis* infection does not result in sufficient acquired immunity against repeat infection has important implications for TB control. Rates of relapse and reinfection vary across different

countries, increasing with higher TB incidence and density of active disease,[112] and also across different subgroups. Patients with HIV are at increased risk of recurrent TB because of an increased risk of reinfection.[9,106,113]

False positive TB isolation

Strain typing can also be used to identify and confirm false positive *M. tuberculosis* isolates. Though not directly related to TB epidemiology, this is important for quality control and TB programming. False positive isolation may be the result of contamination, mislabelling or mishandling of clinical samples.[114] A false positive diagnosis of TB can cause physical and emotional distress to the patient and their contacts as a result of the TB treatment and the contact tracing and screening that will be carried out around the index case, all of which will have costs. In addition, false positive TB diagnosis may delay the diagnosis of the patient's true condition, which has important cost and health implications.[115] When multiple isolates are processed in the laboratory within a certain time period have the same strain type, possible false positive isolation can be investigated. It has been suggested that this should apply to isolates processed within seven days,[116] though the optimum time period will vary depending on the caseload of the laboratory. The adoption of MIRU-VNTR typing should help to improve the speed at which false positive TB diagnoses are detected.[117]

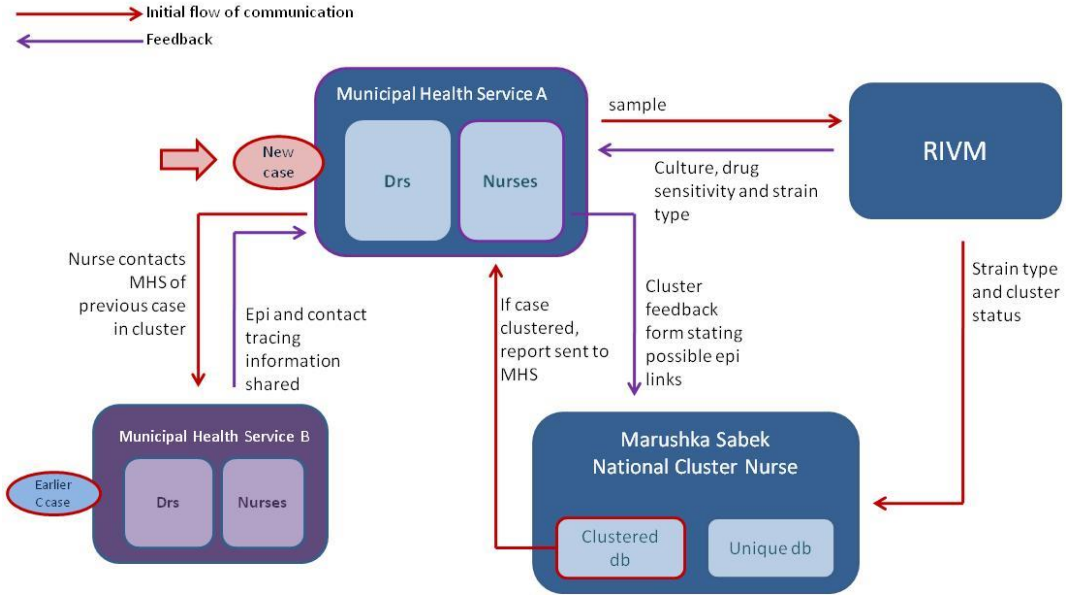
Rates of reported false positive isolation range from 0.1% to 65.9%,[118] but these estimates include studies initiated because of a suspicion that false positive isolation was occurring. It is, therefore, not surprising that the rates identified in these were higher. Population based studies with a sample size greater than 500 have a much smaller range of lower estimates (between 0.1% and 3%). Using IS*6110* RFLP typing the rate of cross-contamination in London has previously been estimated to be between 0.54% and 0.93% (depending on whether you include possible incidents of cross contamination or not),[118] which is lower than the rates observed in the USA (Jasmer *et al* observed 2% in California;[119] Braden *et al* observed 3.5% in Arkansas[120]; and Frieden 3% in NYC).[121] There is no UK-wide estimation of the rate of false positive TB isolation detected by MIRU-VNTR typing.

1.3.3 National TB strain typing services

Universal strain typing services are currently operating in the Kingdom of Denmark (Denmark, Greenland, Iceland and the Faroe Islands), the Netherlands, USA, Norway and Slovenia. The services are summarised in Table 3.

The first national strain typing service was initiated in the Netherlands in 1990, based on *IS*6110 RFLP typing of all TB isolates. Since 1995, the national cluster nurse routinely reports new cases in a cluster to the nurses whose patients are in the cluster. Standardised questionnaires about possible epidemiological links between the cases in the cluster are completed by additional information gathering and collaboration between the nurses. These questionnaires are returned to the national cluster nurse and are linked to the national TB surveillance database. This process is summarised in Figure 10. In the Dutch service there is no threshold for cluster investigation; all clusters are investigated. This approach seems to be manageable for TB nurses in the Netherlands (*personal communication with Marushka Sabek, the national cluster nurse, 30th September 2010*), where the burden of TB is much lower than in England – the number of cases in the Netherlands was 958 in 2012 (5.7 per 100,000), [122] compared to 8751 in England (13.9/100,000).[123]

Figure 10 – Flow of information in the strain typing service in the Netherlands



Based on personal communication with Marushka Sabek, the national cluster nurse, and Dick van Soolingen, head of the Mycobacterium Reference Laboratory (meetings in Rotterdam and Bilthoven in October 2010)

Table 3 – Summary of national TB strain typing services

Country	Typing method, year started	Reporting	Software	Cluster investigation	Key publications
Denmark	*IS*6110 RFLP, 1992 24 MIRU-VNTR, 2009	Routinely reported to the hospital department managing the patient with other laboratory results. Report to surveillance authorities when requested	No. Strain typing linked manually to clinical and epidemiological data	Ad hoc investigations initiated by the authorities in the five Danish regions	[124–126]
The Netherlands	*IS*6110, 1990 24 MIRU-VNTR, 2008	To national cluster nurse and municipal health service	Cluster forms are attached to each patient in the web-based surveillance system.	All clusters	[100,105,127–129]
Norway	*IS*6110, 1994 + Spoligotyping, 1999 24 MIRU-VNTR, 2012	Routinely reported to the laboratory that sent the MTB isolate and to the Department of Infectious Disease Epidemiology at the Norway Institute of Public Health. If transmission is suspected the public health doctor in charge is contacted.	Results are recorded in the national TB register along with the other information about each TB case	Clusters that fulfil the following criteria are investigated: ≥3 cases, diagnosed within 2 years, born in Norway or different countries of birth, and lived at least 6 months in Norway. Any cluster with children or MDR	[130–132]
Slovenia[a]	*IS*6110 RFLP, 2000 + Spoligotyping, 2003 24 MIRU-VNTR, 2009	.	.	.	[133,134]
USA	12 MIRU-VNTR, 2004 24 MIRU-VNTR, 2009	Strain types are submitted, maintained and accessed via the TB Genotyping Information Management System (TB-GIMS). Alerts about cluster priority are automatically generated	TB Genotyping Information Management System (TB-GIMS), 2010	Low, medium and high alert clusters based on expected geospatial concentrations of a strain in a specific county	[88,135–137]

[a]The reference laboratory in Slovenia was contacted via email to gather more information about their service, but no response was received.

The USA was the first country to start using MIRU-VNTR strain typing. In this setting, the strain typing information is reported through the Tuberculosis Genotyping Information Management System (TB-GIMS), a purpose-built software to collate and report information about strain typing and detect outbreaks. The outbreak detection is based on identifying greater than expected numbers of a single strain in one geographical area (county), compared with the national distribution of that strain, and sending alerts indicating low, medium or high risk of recent transmission. These alerts can help TB programs to identify outbreaks and prioritise strain typing clusters for cluster investigation. The National TB Genotyping Service is centred around the TB-GIMs which is accessed by the laboratories to upload the strain typing data, by the state public health teams to monitor and analyse the data, and by local TB programs to access data on individual cases or clusters. State and local TB programs work together to use the strain typing information for TB control and cluster investigations. A comprehensive guide to the application of strain typing to TB prevention and control was published to inform all TB program staff.[138]

1.3.4 The TB strain typing service in England

Following the steady rise of TB rates in the UK, the Chief Medical Officer published a TB Action Plan in 2004 that identified key control measures including the need for a national strain typing service as a component of TB control.[139] Three main requirements were identified for the implementation of a typing service:

1. Standardised laboratory methods to appropriately differentiate *M. tuberculosis* isolates;
2. The establishment of a central database containing both typing and epidemiological data; and
3. Operational guidelines on the public health use of molecular typing data.[140]

Prior to January 2010, strain typing was used on an *ad hoc* basis in all parts of England, apart from the West Midlands where 15 loci MIRU-VNTR had been carried out universally since 2003.[91,141–143] This meant that a 15 MIRU-VNTR strain type could be requested from the laboratories where the addition of patient strain

types would contribute greatly to public health decision making. Strain typing was used to identify outbreaks or disprove suspected transmission between cases.[91,99] For example, if two cases of TB were diagnosed in a school, the strain types would be requested to establish whether there had been transmission within the school or not. This would help to decide whether a large screening programme should be conducted in the school.

The TB Strain Typing Service (TB-STS) was launched in January 2010. It was initiated as a three-year project, running from October 2009 to December 2012. A prospective evaluation was commissioned by Public Health England (PHE) at the start of the service in order to inform PHE's strain typing strategy, i.e. whether the TB-STS should continue after the project period, and if so, how?

The TB-STS was based on prospective typing of *Mycobacterium tuberculosis* using 24 loci MIRU-VNTR. The results would be linked to existing national epidemiological and laboratory datasets, and reported in real time to front line teams. It was hypothesised that prospective universal strain typing and analysis of clusters could be used in real time to inform public health action.

The TB-STS was designed to support and strengthen the national TB control programme through the following aims (Figure 11):[140]

- To provide universal prospective strain typing of every first *M. tuberculosis* isolate from all TB patients in England with MIRU-VNTR typing;
- To generate the information needed to stop further transmission;
- To contribute towards contact investigations in real time; and
- To detect or confirm false positive cultures, thereby preventing the consequences of incorrect diagnoses due to laboratory contamination or error.

If the TB-STS were <u>not</u> implemented, then the following would be expected:

- Transmission will continue undetected for strain type clustered cases that are not identified as epidemiologically linked through traditional contact tracing;

- Where transmission is suspected public health action will be taken unnecessarily unless an ad hoc strain typing request is made; and
- False positive TB isolates will not be detected and patients with false positive TB will receive unnecessary treatment.

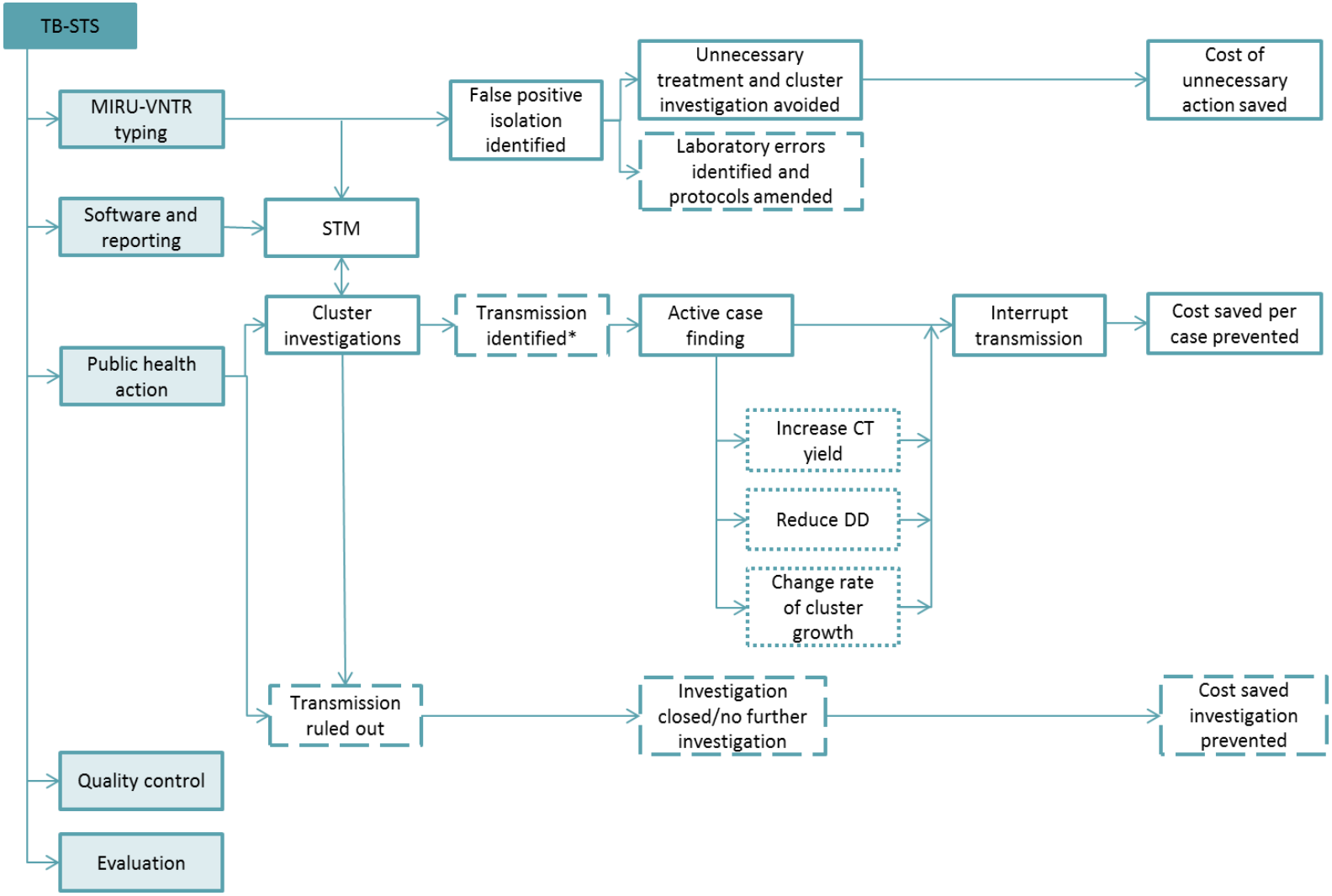


Figure 11 – The aims of the TB-STS

Shaded boxes are the five components of the TB-STS.
Dotted lined boxes are indirect outputs of the cluster investigations that were hypothesised to interrupt transmission.
Dashed lined boxes could not be adequately captured for this evaluation.
*Confirmed (or possible) epidemiological links identified
STM strain typing module (software to integrate strain typing with epidemiological and clinical data)
CT contact tracing
DD diagnostic delay

1.4 Evaluating strain typing as a tool for TB control

Evaluations explore whether interventions and policies do what they intend to do, and what impact they have. This is necessary because interventions are costly, funds are limited and for every intervention implemented, there is an opportunity cost – what alternative interventions could have been implemented with the same resources? Evaluations, therefore, enable decision-makers to set priorities. There is also the possibility that interventions may have unexpected, adverse effects. Evaluations can be useful for the design of future interventions and policies. There are five main evaluation paradigms: experimentalist, pragmatist, constructivist, pluralist and realist. These approaches are briefly summarised here.

Experimental evaluation is based on the experimental research design whereby an intervention is introduced into one of two groups that are matched (either by randomisation or matching on potential confounders) and the groups are measured before and after the intervention.[144] This relies on a theory of causation that removes all other possible causative agents so that there is just one possible causal link; the intervention. This approach is not easily applied to interventions that exist in the real world and introduces the 'black box' problem, whereby it only produces a description of the outputs, rather than explaining why some does or doesn't work.

Pragmatist evaluation calls for evaluations to be useful to those that they are intended for; they are utilisation-focussed. There are four features of a pragmatic evaluation: utility, feasibility, propriety and accuracy.[145] This focus raises the problem of the policy-maker or the funder having too much influence over the evaluation, removing the equipoise and the objectivity. This approach also led the way for a text-book approach to evaluation, with a prescriptive step-by-step method.[146,147]

Constructivist evaluation turns the focus away from the outputs of an intervention and the policy makers, towards the processes and meaning – engaging all possible stakeholders to establish why or why not an intervention works.[148] This paradigm falls short because it does not allow for the objective assessment of an intervention or the asymmetry of power across the stakeholders.

Pluralist evaluation is an attempt to combine the experimental, the pragmatist and the constructivist paradigms. One form of this is the comprehensive evaluation, which ash three main activities:[149]

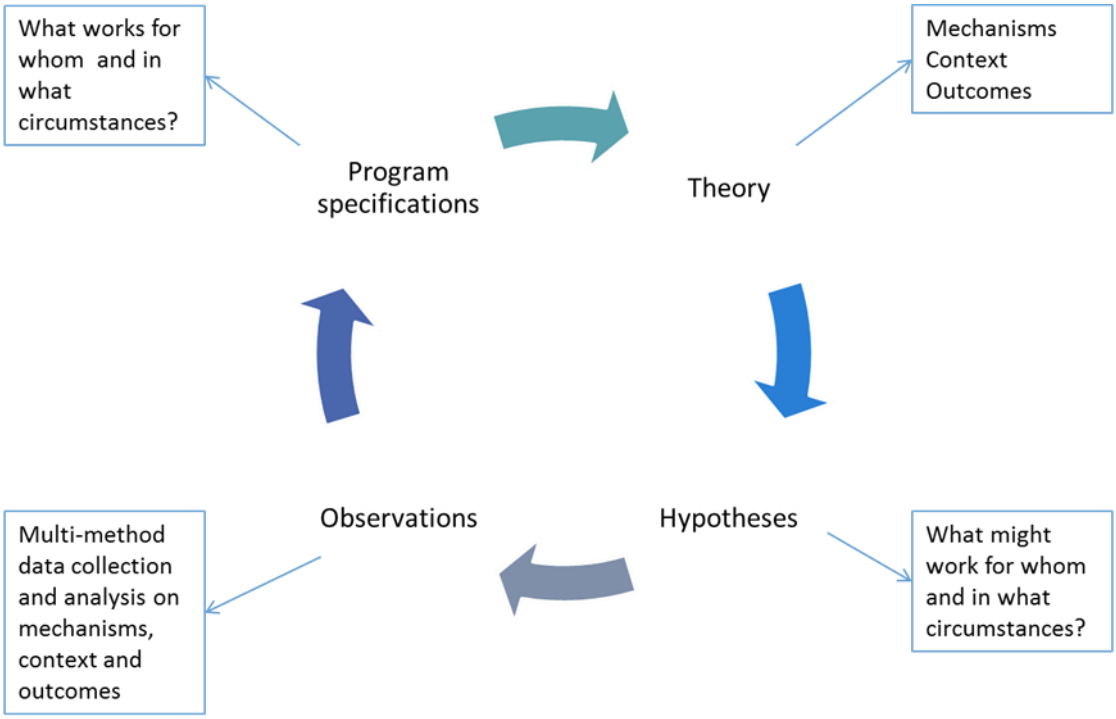
1. Analysis of the intervention design
2. Monitoring of the program implementation
3. Assessing program utility.

This approach is criticised for trying to encompass too much; being too broad and requiring too many resources to conduct properly.[150]

The final paradigm summarised here attempts to make the pluralist approach more realistic and applicable to the real world. Realist evaluation is based on a generative model of causation: that an action is causal only if its outcome is triggered by a mechanism acting in context (outcome=mechanism+context).[150] Pawson and Tilley (1997) argue that an evaluation should demonstrate if the program works, what it is about the program that works for whom and in what conditions. The realist evaluation is based on Wallace's (1971) wheel of science and argues for an evaluation cycle that includes:[151]

- Theories (based outcome=mechanism+context);
- Hypotheses (that hypothesise what might work for whom and in what circumstances);
- Observations (multi-method data collection and analysis of the mechanisms, context and outcomes); and
- Program specification (what works for whom and in what circumstances).

Figure 12 – Realist evaluation cycle



Adapted from: Pawson and Tilley (1997) p.85 [150]

Existing evaluations of TB strain typing services

TB surveillance is carried out in most countries to varying degrees of scope and quality.[1] However, molecular surveillance is a relatively new and expensive form of surveillance currently reserved for richer countries with lower TB burdens, and for research purposes. Therefore, molecular surveillance for TB occurs in very few countries; in 2014 the only countries with a universal molecular surveillance system for TB were the Netherlands,[100,127] the USA,[64,136] Denmark,[124,152] Norway[130,132] and Slovenia.[133,134] The public health value of such surveillance systems has been demonstrated through the retrospective evaluation of strain typing data.[87,102,116,127] A universal service in England was anticipated to add value to the TB control strategy.[153]

Evaluations of existing strain typing services have focussed on single elements of the service such as the discrimination of the typing method,[69] effectiveness of cluster investigations,[97,152] and the ability to identify cross contamination.[117] No evaluation

to date has looked at the entirety of the service and analysed its complexity, taking into account the resources and infrastructure, the processes involved, the multiple outputs and the long term outcomes.

Complex interventions

A national strain typing service is a complex intervention. Complex interventions are often defined as interventions with several interacting components.[154] The MRC's definition of a complex intervention includes:

- Number of, and interactions between, components within the experimental and control interventions
- Number and difficulty of behaviours required by those delivering or receiving the intervention
- Number of groups or organisational levels targeted by the intervention
- Number and variability of outcomes
- Degree of flexibility or tailoring of the intervention permitted

Complex interventions present difficult problems for evaluators due to the nature of their complexity – the design and delivery of the intervention is complex; the effect of the local context on the interventions is variable; and the length and complexity of the pathway between the intervention and its outcomes may be difficult to unwrap. In addition, there are practical difficulties in applying experimental research methods to service evaluation. A strain typing service is a complex intervention because it has many different parts that are organised and delivered by different groups (such as laboratories, public health teams, clinical teams); these groups accept and implement the service differently in different settings and across different parts of the country; and the causal chains between a strain typing service and its potential outcomes are complicated and difficult to capture.

To increase the number and improve the quality of evaluations, frameworks for evaluating complex interventions have been developed. The MRC published a 'Framework for the Development and Evaluation of RCTs for Complex Interventions to Improve Health' in 2000.[155] In 2006 they published an update based

on the experience collected by the scientific community, the need to include non-experimental research methods and to make it applicable to interventions outside of the health service.[154] The overall framework covers the development and piloting of an intervention, the evaluation and dissemination of findings. The evaluation framework has three main components: assessing effectiveness, process evaluation and assessing cost-effectiveness. It is acknowledged that an experimental research design is not always possible in the evaluation of a complex intervention, instead quasi-experimental or observational studies may be adequate alternatives. The framework advocates for a good theoretical understanding of the intervention in order to identify appropriate outcome measures, and suggests the use of surrogate outcomes measures.

The MRC evaluation framework has been criticised by 'realists' for not including an explanation of the mechanisms of change that might link the intervention with the outcomes and not examining how the intervention might interact with the context. Realist evaluation is based on a generative model of causation: that an action is causal only if its outcome is triggered by a mechanism acting in context (outcome=mechanism+context).[150] Pawson and Tilley (1997) argue that an evaluation should demonstrate if the programme works, what it is about the programme that works for whom and under what conditions.

Whilst the MRC framework advocates for the development of the intervention to be based on theory, it does not include the role of a theory in the evaluation process, which is crucial to other theory-driven frameworks.[156,157] Theory-driven frameworks are based on the thesis that understanding the theory underlying an intervention is necessary for understanding whether it works and how it works. Theory of Change is one such framework in which a hypothesised theory of how the intervention affects change is developed with stakeholders and can be tested empirically.[158] The advantages of Theory of Change is that it can better represent the complexity of an intervention as it makes explicit the causal pathways, but does not impose a structure, allowing for interactions, feedback loops and multiple pathways. Compared to the MRC Framework, disadvantages of Theory of Change is that there are multiple,

prescriptive steps that need to be followed which may not be appropriate for the evaluation context and it does not include a framework for the dissemination of results.[159,160]

These frameworks incorporate the processes involved in the development of the intervention as well as its evaluation. There is an assumption that the evaluation team have been a part of the intervention design team and that the design, implementation and evaluation of the intervention is (or can be) an iterative process. In situations where this is not the case, and the evaluation is either retrospective, or as an add-on to the implementation of the intervention, these frameworks may not be the most appropriate choice.

Donabedian's formative framework for evaluating the quality of medical care[161] provides a framework around which an evaluation can be designed, without prescribing any particular steps, nor assuming that the evaluators are involved in the design of the intervention from its initiation. The framework divides a service up by its Structures, Processes, Outputs and Outcomes. 'Structures' refers to the resources and inputs of the service; 'Processes' refers to the activity of the resources and the processes involved in the service; 'Outputs' refers to the products of the service; and 'Outcomes' refers to changes resulting from such outputs. An advantage of this framework is that it ensures that the whole of the service is considered in the evaluation, rather than just focussing on the Outcomes of the service and ignoring the Structures and Processes that produce those outcomes. In effect, by evaluating the Processes one is attempting to avoid so-called Type III errors – where one ends up evaluating an intervention that is not implemented properly.[162] Process evaluation can provide the insight and context necessary to interpret the outputs of a service, help to explain differences between the observed and expected outcomes, and identify ways of intervening to improve the service.[154] Subsequently, this enables the researchers to develop constructive recommendations based on the Structures and Processes of the service – elements of the service that can be directly influenced – as recommended in other evaluation guidelines.[154]

Evaluation frameworks tend to agree in many ways: the evaluation design should take into account who the evaluation is for and what kind of decisions it might influence; hypotheses about how the intervention will affect change should be made explicit; the appropriate methods available (of which there are likely to be multiple) should be used to understand if the intervention works, how it works, and in what contexts.

1.5 Aims and Objectives of this thesis

1.5.1 Aim

The aim of this thesis is to evaluate the TB-STS in England.

1.5.2 Objectives

The objectives of this thesis are:

1. To systematically review the literature to examine the effect of study design and setting on the estimation of the proportion of TB clustering by MIRU-VNTR strain typing

2. To describe the Structures of the TB-STS

3. To evaluate the Processes of the TB-STS
    a. To assess the implementation and perceptions of the TB-STS
    b. To explore the user experience of the TB-STS

4. To evaluate the Outputs of the TB-STS
    a. To describe the laboratory outputs of the TB-STS, including estimating the proportion of false positive TB isolation identified by the TB-STS
    b. To quantify the public health outputs of the TB-STS

5. To evaluate the Outcomes of the TB-STS
    a. To investigate the effectiveness of the TB-STS
    b. To assess the cost-effectiveness of the TB-STS

6. To contextualise the findings of the evaluation
    a. To discuss the strengths and limitations of the evaluation
    b. To make recommendations for future typing services and evaluations

Chapter 2. Effect of study design and setting on TB clustering estimates using MIRU-VNTR: A systematic review

One of the main applications of strain typing is to monitor clustering rates as an estimate of recent transmission over time (described on page 21). Published reviews have identified factors that might influence or bias clustering by IS6110 RFLP.[93,96] No study has repeated this analysis using more up-to-date typing methods, which is important for understanding the epidemiology of TB and to shape the application of molecular typing to improve TB control. In this chapter, the available literature reporting clustering using MIRU-VNTR typing was reviewed systematically to examine if the biases that occur using RFLP typing are also relevant for studies using MIRU-VNTR.

—

Studies using IS*6110* RFLP data show that the proportion of clustering observed can be affected by 1) study design (affecting the proportion of eligible cases that are included in the study); 2) features of the typing method (such as the ability to type isolates with low copy numbers); and 3) study setting (such as characteristics of the study population). For example, the proportion of clustering increases when the fraction of the total data sampled increases[96,163,164] and when study duration increases.[109]

MIRU-VNTR strain typing is increasingly being adopted worldwide,[64,100,127,134,152] yet unlike IS*6110* RFLP, the evidence for the interpretation of the findings such as the impact of study design and setting on clustering have not been reviewed. Although the two typing methods have been shown to have a similar discriminatory value, the markers evolve independently and at different rates, resulting in a difference in clustering between the two methods.[165] This suggests that there could

be differences in the way study design, typing method and setting affects clustering by the two methods. A systematic review was conducted to assess the evidence for the impact of study design and setting on the interpretation of TB transmission using clustering derived from MIRU-VNTR strain typing – as has been shown using IS*6110* RFLP typing.

2.1 Methods

Five electronic databases were searched (EMBASE, ISI Web of Science, CINHAL, Scopus and Medline (Ovid)) up to 20th October 2014. The search strategy combined the following terms with Boolean operators: Tuberculosis, strain typing, and transmission (Appendix 1). The search was limited to studies using the standard MIRU-VNTR method,[166] in humans only, and in English.

All titles and abstracts from each of the searches were examined. The full text of each paper was obtained and reviewed if the study reported MIRU-VNTR strain typing of *M. tuberculosis* complex isolates with at least 15 of the standardised 24 loci (ETR A, B, C, D, E; MIRU 2, 10, 16, 20, 23, 24, 26, 27, 39, 40; VNTR 424, 1955, 2163b, 2347, 2401, 3171, 3690, 4052, 4156).[56,70,166]

Studies using fewer than 15 loci were not included because the level of discrimination is inadequate for epidemiological use (n=121).[166] Studies that used loci different to the standardised 15 and 24 set were not included in the analysis in order to reduce the heterogeneity between studies (n=19). All publication types were included in this first screen to ensure that no relevant data were missed.

Reviews, letters, editorials, outbreaks or case reports (n=103) were excluded in the second screen. Studies that used incomplete sampling (e.g. studies using subsets of populations such as multidrug-resistant patients) (n=47) and studies that had a sample size of less than 50 (n=4) were also excluded.

A reviewer (Jessica Mears) extracted the following data items from all included studies using a form developed in Excel (Microsoft 2010): publication details (year,

authors, study country), study details (study duration, loci typed, secondary typing method, study population, whether participant consent was required (a characteristic of the study design that was used as proxy for sampling fraction, assuming that where consent was required the sampling fraction was low)), the number of clustered and unique isolates, and the covariates of interest: the maximum size of clusters; the proportion of clusters containing two cases; the proportion of the population that was culture positive; the proportion of culture positive isolates typed; risk factors for clustering; and the Hunter Gaston Discriminatory Index (HGDI)).[167] Ibrahim Abubakar extracted data from 10% (n=3) of the included papers for external validity, disagreements were discussed and a consensus agreed upon.

The main outcome measure – the proportion of TB isolates clustered by MIRU-VNTR strain typing – was calculated as the number of clustered isolates/number of clustered+unique isolates. Where there were uncertainties the two reviewers consulted each other.

Authors were contacted if TB incidence rate was not reported. Where no response was received WHO country estimates of TB incidence for the study year were used.[168] As so few studies reported the proportion coinfected with TB/HIV, these estimates for the study country were taken from an EU-wide survey and WHO country profiles.[169,170] Due to poor recording of the sampling fraction (the number of isolates typed/the total number of culture positive TB cases diagnosed during the study period (n=19)), whether the study required the consent of participants (yes/no) was included as a proxy for (low/high) sampling fraction. The risk of bias within each study was assessed using the STROME-ID checklist.[53]

Data were analysed in Stata version 12 (StataCorp. 2011. *Stata Statistical Software: Release 12*. College Station, TX: StataCorp LP.). Where studies reported data from more than one set of loci, the method with the highest discriminatory value was included (i.e. MIRU-VNTR 24 would be chosen over MIRU-VNTR 15, and MIRU-VNTR 15 plus Spoligotyping would be chosen over MIRU-VNTR 15 alone) (n=8). This review was not concerned with summary measures of clustering, but factors that

influenced clustering; therefore articles must have included at least one of the covariates. Continuous variables were transformed where the distribution was skewed. The proportion clustered was transformed using the Freeman Tukey transformation.[171] Study heterogeneity was assessed using a forest plot and the chi^2 test of heterogeneity. Univariable meta-regression analyses were carried out to determine the effect of the study design covariates on the proportion of clustered isolates. All covariates in the analysis were hypothesised to influence the proportion clustered *a priori*.

Sensitivity analyses were conducted to see the effect of removing studies reporting 0% clustering, with only extra-pulmonary TB cases, only *M.bovis* cases, studies using the 'old 12' MIRU loci as part of their 15 loci, and studies assessed as having a high likelihood of bias (STROME-ID score less than 20).

2.2 Results

The search identified 7274 references resulting in 27 studies (25 journal articles and 2 conference abstracts) included after deduplication and title/abstract/full text screening (Figure 13). The study setting and design characteristics of the included studies are shown in Table 4.

Figure 13 – Results of systematic search, screening and data extraction

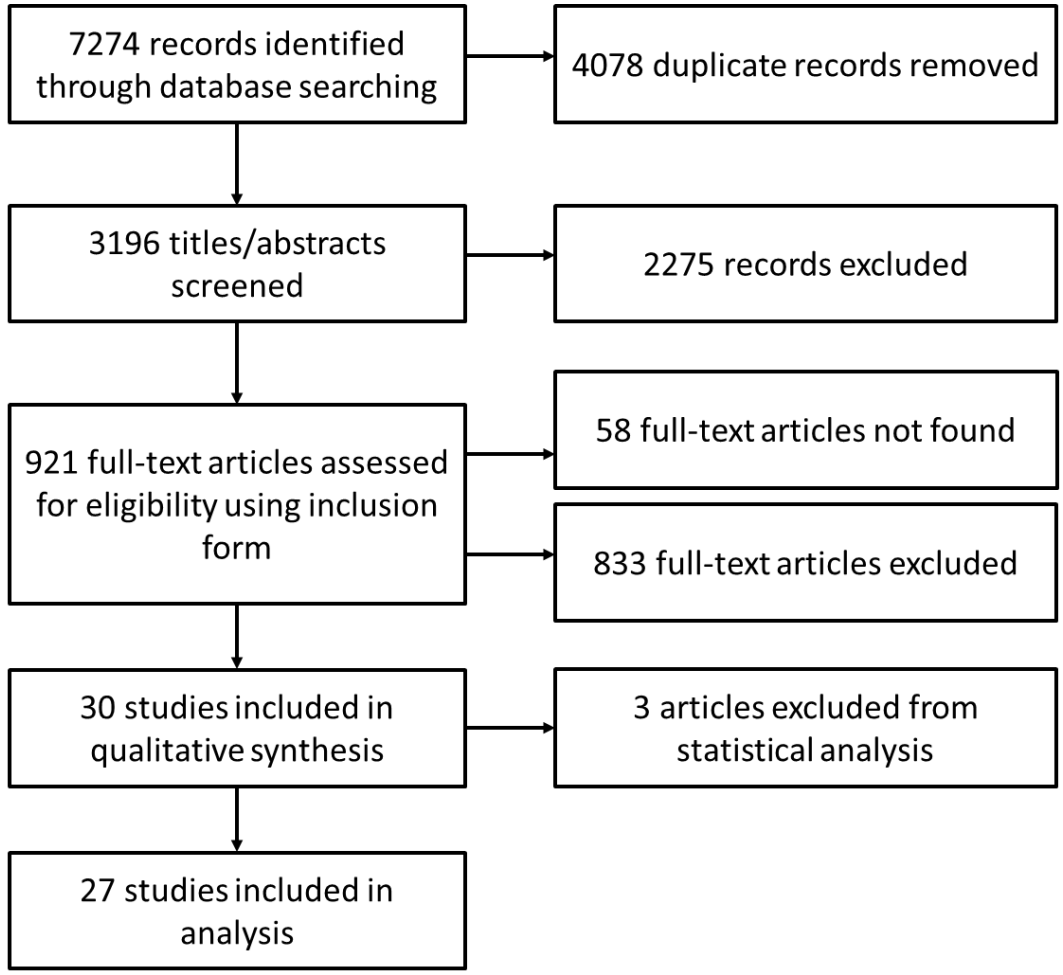


Table 4 – The study setting and design characteristics of the included articles

Ref	Study setting							Study design							S-ID [d]	Risk of bias [e]	Clustering (%)[f]
	Study area and country	TB incidence (per 100,000)	TB/HIV (per 100,000)[a]	Previous TB treatment (%)	Pulmonary TB (%)	Maximum cluster size	Clusters of size 2 (%)	Study duration (months)	Study size (clustered + unique isolates)	Culture positive in study population (%)	Culture positive isolates typed (%)	Typing method [b]	Loci typed [c]	Consent required			
172	New South Wales, Australia	6.7	0.2	0.0	63.7	.	.	36	1128	.	.	m24	N	no	32	low	20.1
173	Tabriz and Orumieh, Azarbaijan	26.0	.	5.2	87.0	5	81.8	12	156	.	94.5	m15	O	no	28	low	32.7
72	Brussels-Capital Region, Belgium	35.2	5.1	10.8	.	23	64.2	24	530	86.1	87.9	m24	N	no	32	low	29.6
174	Brussels-Capital Region, Belgium	35.2	5.1	.	100	.	.	39	802	81.8	84.7	m24s	N	no	25	low	28.8
175	Ontario, Canada	4.8	0.4	.	.	18	58.8	65	2016	.	.	m24s	N	no	31	low	23.1
176	Changping District, Beijing, China	.	0.3	.	100	0	.	30	318	31.5	94.6	m24	N	no	19	high	0.0
177	Croatia	19.0	0.1	.	.	45	48.3	36	1587	.	.	m15	N	no	19	high	62.8
178	Amhara region, Northwest Ethiopia	.	24.0	17.6	100	13	.	5	244	.	.	m24	N	yes	26	low	45.1
179	Finland	5.0	0.0	.	.	20	.	48	1048	75.4	99.4	m15s	.	no	29	low	33.9
73	Hamburg, Germany	12.7	.	.	.	.	45.5	12	154	78.2	91.1	m24s	N	no	34	low	22.1
180	Schleswig-Holstein, Germany	3.2	0.1	.	.	22	44.4	48	277	.	.	m24s	N	no	16	high	27.1
181	South West Ireland	15.3	3.3	.	82.7	12	.	36	171	79.5	96.1	m24s	N	no	36	low	27.5
182	South Tawara, Kiribati	370.0	.	4.1	100	25	55.6	24	73	45.4	98.6	m24s	N	yes	24	low	75.3
69	Netherlands	6.5	0.2	.	.	.	57.2	60	3978	.	100.1	m24	N	no	30	low	46.7
183	Kharkiv, Russia	94.0	3.8	63.3	100	10	50.0	3	98	.	100	m15	O	yes	19	high	31.6
184	Eastern province, Saudi Arabia	4.0	.	.	73.1	24	19.0	24	522	.	.	m24s	N	no	23	low	40.2
185	Singapore	40.5	1.2	.	.	21	48.0	24	1128	82.0	34.5	m24s	N	no	30	low	30.8
133	Slovenia	10.6	0.0	.	.	6	.	12	196	94.4	97.5	m24s	N	no	31	low	36.2
186	Almeria, Spain	26.0	6.0	.	.	8	.	27	281	.	81.9	m15	N	no	18	high	43.1
187	Sweden	4.8	0.1	.	.	10	.	36	406	.	.	m24s	N	no	22	low	21.2
188	Mubende, Uganda	.	86.0	31.1	87.8	11	70.0	6	67	21.5	90.5	m15s	N	yes	25	low	35.8
189	East Lancashire, UK	18.3	8.2	.	.	13	58.3	102	332	48.5	69.9	m15	O	no	23	low	42.8
190	UK	.	8.2	.	42.3	12	50.0	48	102	90.7	87.2	m15	O	no	32	low	30.4
191	London, UK	44.9	8.2	.	.	.	.	9	964	36.0	100	m24	N	no	.	.	37.0
192	Midlands, UK	15.0	8.2	.	.	.	.	48	4207	58.3	100	m15	O	no	.	.	61.2
193	Odessa and Nikolaev, Ukraine	80.4	3.9	34.2	100	.	.	4	225	.	.	m15	O	yes[g]	23	low	60.4
194	Hanoi, Vietnam	146.0	10.0	0.0	100	.	.	20	465	92.7	91.9	m15s	N	yes	31	low	55.3

Footnotes overleaf

[a] Estimates from of the prevalence of TB/HIV co-infection in the study country [169,170]
[b] 15=15 MIRU-VNTR loci (made up of the 'old 12' or 'new 12' defined in the footnote below), 24=24 MIRU-VNTR loci (ETR A, B, C, D, E; MIRU 2, 10, 16, 20, 23, 24, 26, 27, 39, 40; VNTR 424, 1955, 2163b, 2347, 2401, 3171, 3690, 4052, 4156), S=with Spoligotyping
[c] O= old 12 MIRU loci (MIRU 2, 4, 10, 16, 20, 23, 24, 26, 27,30, 31, 39, 40), N=new 12 MIRU loci (MIRU 10, 16, 26, 31, 40 + Mtub 04, 21, 39 + ETR A C + QUB 11b, 26)
[d] S-ID STROME-ID scores; individual studies score one for each element of checklist they had addressed
[e] Risk of bias was assessed using the STROME-ID checklist. Studies scoring <20 were categorised as have a high risk of bias.
[f] The proportion of clustering was calculated as the number of clustered isolates/number of clustered + unique isolates
[g] 11.3% did not consent to being part of the study. The other studies that required consent for isolates to be typed did not report the refusal rate

Studies were published between 2007 and 2014. The number of studies reporting each variable of interest is shown in Table 5. The clustering reported varied from 0% [176] to 62.8%.[177] In all studies, clustered isolates were defined as having identical strain types based on the MIRU-VNTR loci typed, with or without Spoligotyping. 17 studies included isolates from newly diagnosed TB cases, three studies reported including isolates from new and chronic cases of TB, and seven did not report this information. In addition, ten studies did not include repeat isolates from the same patient, one study included a repeat isolate from one patient, and the remaining 17 did not report whether repeat isolates were included or not. Furthermore, four studies included isolates with missing loci in the cluster analysis, whereas four excluded isolates with missing loci, and the remaining 20 did not report how they dealt with missing loci.

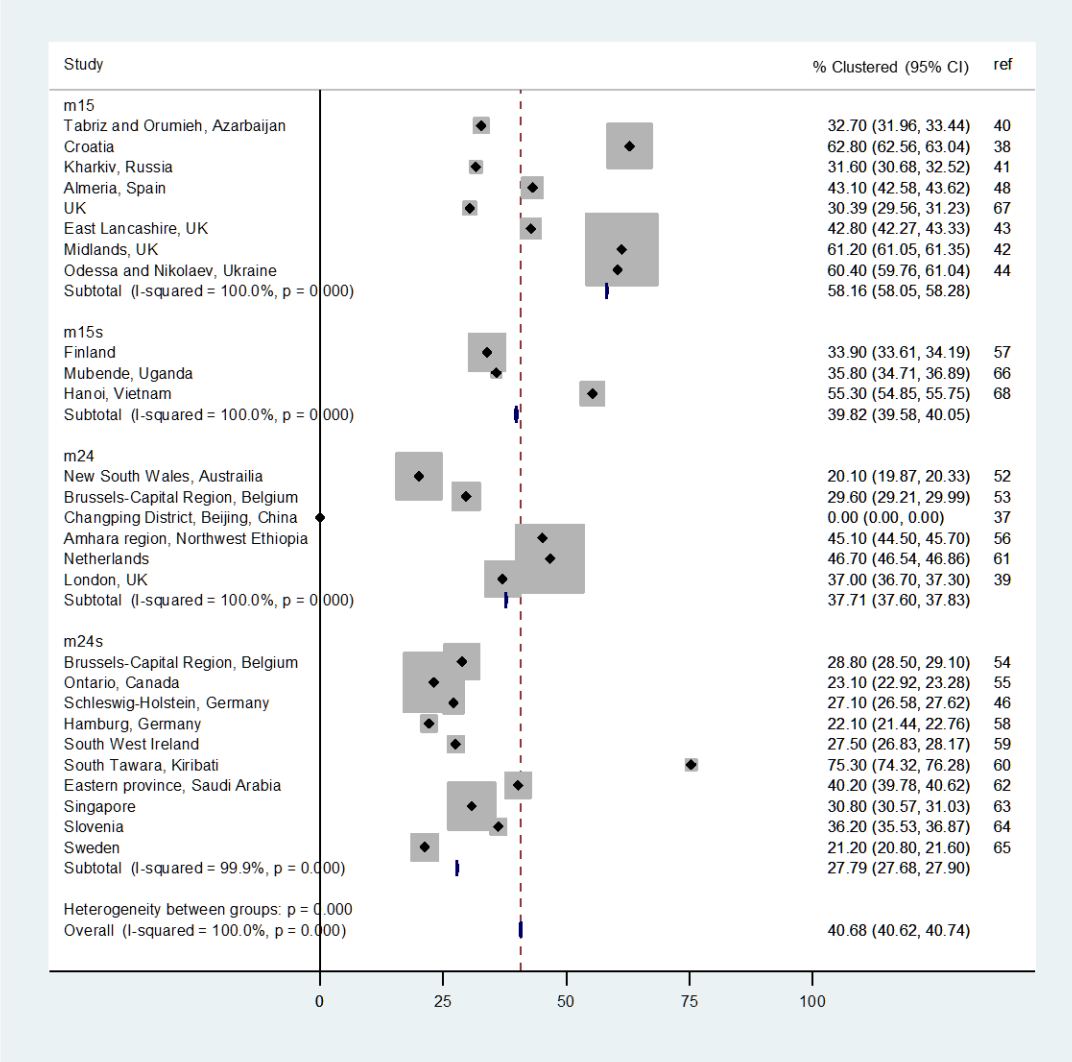
Table 5 – The number of studies that reported the variables of interest

	Reported	Missing
Study setting		
TB incidence	8	19
TB/HIV co-infection	5	22
Previous TB treatment	9	18
Proportion pulmonary TB	14	13
Maximum cluster size	19	8
% clusters with 2 cases	14	13
Study design		
Study duration	27	0
Study size	27	0
% population that is culture positive	15	12
% culture positive typed	19	8
24 loci (compared to 15)	27	0
Repeat isolates	12	15
Missing loci	8	19
Double alleles	1	26
Consent required	6[a]	21
Epidemiological information	6	21

[a] Only one study reported the consent rate

A forest plot shows the spread of clustering reported by number of loci and additional typing method (Figure 14). Significant heterogeneity was identified between the studies (p<0.001), suggesting that a meta-regression would be an appropriate analysis.

Figure 14 – Forest plot showing the proportion of clustering reported in each study by the number of MIRU-VNTR loci typed



The number of loci typed is categorised into 15 loci (m15), 15 loci with Spoligotyping (m15s), 24 loci (m24) and 24 loci with Spoligotyping (m24s). The study reference is shown in the right hand column.

The univariable meta-regression shows evidence for the proportion of clustering to decrease as the number of MIRU-VNTR loci typed increased from 15 to 24 (p=0.04; Table 6), accounting for 14% of the between study variation, and to increase when the study participants consented to being included in the study (p=0.03), accounting

for 14% of the between study variation. The proportion of clustering increased as the TB incidence in the population increased (p=0.007, Adj R^2 = 26.7). There was also evidence for the proportion of clustering to increase as the maximum cluster size increased (p=0.001), accounting for 48% of between study variation. There was no evidence of the other study design or study setting variables significantly influencing the proportion clustered. Though non-significant (p>0.05), the TB/HIV coinfection rate in the population explained 2% of the between study variation. Too few studies included information on the proportion of clusters containing two cases, proportion of the study sample with previous TB or with pulmonary TB, so these could not be included in the analysis (Table 5).

Table 6 – Univariable metaregression showing the coefficients for change in the proportion of clustering and the percentage of between-study variation explained by variables describing the study design and setting

	n	Coefficient [a]	CI	p	Adj R^{2} [b]
Study setting					
TB incidence	23	0.14	0.04-0.24	0.007	26.74
TB/HIV co-infection	23	0.04	-0.03-0.11	0.246	2.00
Maximum cluster size	19	0.20	0.09-0.30	0.001	48.20
Study design					
Study duration	27	-0.02	-0.09-0.06	0.677	-3.37
% population that is culture positive	15	0.34	-1.23-1.96	0.661	-5.92
% culture positive typed	19	0.22	-1.08-1.52	0.725	-5.41
Study size	27	0.03	-0.11-0.16	0.702	-3.31
24 loci (compared to 15)	27	-0.30	-0.59--0.01	0.04	13.58
Consent required	27	0.38	0.04-0.72	0.029	14.41

[a] Coefficients for the change in the proportion of clustering for each covariate. E.g. for a one unit increase in maximum cluster size, the proportion of clustering increases by 0.2.
[b] The proportion of between-study variation explained by the univariate meta-regression.

Sensitivity analyses to examine the effect of excluding studies reporting 0% clustering,[176] only *M.bovis* cases,[190] studies using the 'old 12' MIRU loci,[173,183,189,190,192,193] and studies assessed as having a high risk of bias,[176,177,180,186,195] did not generally change the results. The proportion of culture positive TB in the population remained insignificant but explained 2.6% of the between study variation when excluding 0% clustering (p=0.278 and Adj R^2=2.62). Similarly, the proportion of culture positive TB in the population remained

insignificant but explained 2.6% of the between study variation when excluding studies with the highest risk of bias (p=0.278 and Adj R^2=2.62). The number of loci typed became non-significant, but explained 9.6% and 10.5% of the between study variation when excluding studies using the 'old 12' loci and the highest risk of bias, respectively (p=0.106, Adj R^2=9.63; p=0.111, Adj R^2=10.51, respectively).

2.3 Discussion

This review identified 27 studies that met the inclusion criteria. The findings show that the interpretation of studies using MIRU-VNTR to estimate clustering is subject to bias relating to study design and setting; however, there were insufficient data available to fully explore this impact.

As expected, the proportion of clustering decreased with a greater number of MIRU-VNTR loci typed, with increasing TB incidence and with increasing maximum cluster size. Requiring consent to type patient isolates increased the proportion of clustering, which is not expected, given that the sampling fraction would be lower in these studies.

The other study design variables included in this analysis, such as study duration, did not significantly influence the proportion of isolates that were clustered, contrary to previous findings [93]. This is likely to be because of a lack of good quality evidence: of the 27 studies that met the inclusion criteria for the review, none reported all the variables of interest, reducing the power of the analysis and precluding multivariable meta-regression (Table 5). Importantly, key details of cluster analyses were not reported consistently across the studies, such as whether repeat isolates from the same patients were included, or typing profiles with missing loci were included, introducing new, unmeasured biases. In addition, the range of the variables may have been too limited to show any impact on clustering estimates. For example, the proportion of culture positive isolates typed ranged from 34.5% to 100%, with 17 of the 19 studies reporting this variable from 81.9% to 100%. Furthermore, most of the studies (17/27=63%) were from low TB burden settings and therefore may be

reflecting the rate at which imported cases have matching strain types by chance, rather than rates of recent transmission.

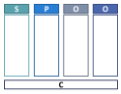
The sensitivity analysis suggested that, when excluding the studies with the greatest risk of bias, the culture-positivity in the population might explain a small amount of the between study variation. This is consistent with estimates of the influence of sampling on the proportion of clustering using *IS*6110 RFLP typing.[196] In the sensitivity analysis excluding studies that used the 'old 12' loci, the effect of the number of loci typed becomes non-significant. This is likely because studies using the 'old 12' accounted for six out of ten studies reporting 15 loci, reducing the number of studies and the power of the model.

This study is a timely evaluation of the impact of study design on estimates of TB clustering using MIRU-VNTR strain typing because it has been incorporated into national typing services globally.[64,197] The findings are relevant where strain typing is used to evaluate TB control systems across different settings because the proportion of clustering is influenced by the number of loci typed, the TB incidence and the maximum cluster size. Given that strain typing methods are advancing beyond MIRU-VNTR typing and that the application of whole genome sequencing to TB control and public health strategies has been demonstrated,[78,79,198,199] it is important that the biases in the analysis of such methods are explored and compared. Understanding how to design and compare research studies for public health will greatly improve the benefit gained from newer technologies.

The strength of this meta-analysis was limited by (a lack of) detail reported by the included studies. This review has highlighted the need for better quality reporting in primary studies to enable future reviews to be more robust. Recently published standards for reporting of molecular epidemiology for infectious diseases should improve the quality of reporting.[53] This review is further limited by our inability to access 58 of the title/abstract screened articles for full text screening.

The use of TB strain typing as a public health tool in TB control programmes is increasing globally. This review has identified a lack of good quality studies that can contribute to our understanding in interpreting the molecular typing of TB. It has also shown that the proportion of clustering derived from MIRU-VTNR typing is influenced by the number of loci typed, whether consent is required to type isolates, TB incidence in the study setting, and the maximum cluster size, highlighting these as important considerations in the design and interpretation of future studies.

Chapter 3. Methods



This chapter describes the methods employed in this evaluation. I begin by briefly describing the evaluation framework adopted. The methods used to address each component of the framework are then presented in turn. A mixed methods approach was employed, including quantitative and qualitative approaches, using both primary data collection and secondary data analysis.

As a guide to the reader, an icon is shown in the top right-hand corner of the title page of this and each subsequent chapter. This is based on the evaluation framework (see below and Figure 16 for an explanation) and will highlight which evaluation component is being addressed by each chapter.

—

3.1 Evaluation framework

The Donabedian framework[161] was chosen as the evaluation framework for this thesis for a number of reasons. Firstly, taking into account the design and implementation of the TB-STS (i.e. national intervention with a national rollout implementation plan), the role of the evaluator (i.e. external and independent, and not involved at the design stage of the intervention), the Donabedian framework is the most appropriate approach as it does not assume the evaluator has influence on the initial design of the intervention and, therefore, an integrated evaluation design (as the MRC framework intuits)[160]. Secondly, the framework takes into account the totality of the TB-STS and, because it does not include step-by-step instructions or a specific methodology for applying the framework (such as Theory of Change),[157] it is suitably flexible and pragmatic in order that the evaluation could be conducted within the constraints of the available data and resources. Thirdly, the Donabedian framework is simpler compared to other approaches outlined above. The application of a relatively simple framework to a complex intervention created scope for the

framework to be adapted for the specific evaluation (see below) and the inclusion of choice elements from other evaluation frameworks (e.g. the use of a causal diagram to illustrate the aims of the TB-STS – Figure 17 – is an approach advocated by Theory of Change).[157] Counter to this, the alternative, more complex evaluation frameworks do not lend themselves so well to adaptation (such as the inclusion of useful elements from different frameworks).

As explained above, the Donabedian framework[161] was chosen in order to take into account the totality of the TB-STS, to have the flexibility to conduct the evaluation within the constraints of the available data and resources, and to apply a relatively simple framework to a complex intervention. This framework has been adapted to take into account the context across the Structures, Processes, Outputs and Outcomes (Figure 15). As described previously, 'Structures' refers to the resources and inputs of the TB-STS such as the creation of service-specific roles; 'Processes' refers to the activity of the resources and the processes involved in the TB-STS such as the laboratory processes and reporting; 'Outputs' refers to the products of the TB-STS activities such as cluster investigations; and 'Outcomes' refers to changes resulting from the Outputs such as a change in TB transmission. 'Context' refers to the need to take into account the political, economic and technological context in order that the findings be interpreted appropriately and in a way that is relevant to the policy decision-makers (as advocated by realist evaluation theory).[200]

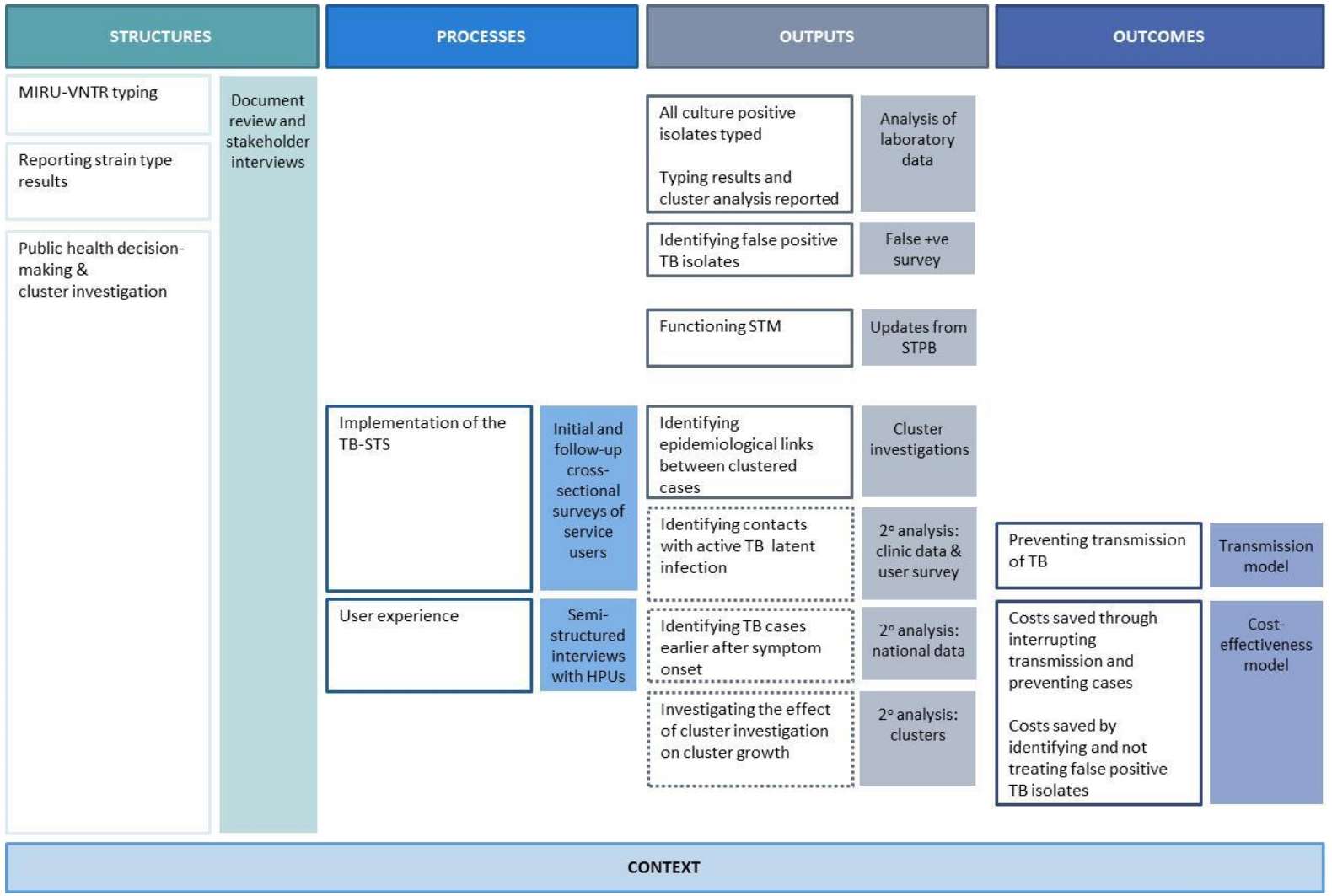
The evaluation consists of the following components (Figure 16): a descriptive component (Structures); cross-sectional surveys of user-perspectives and a qualitative component involving semi-structured interviews with Health Protection Units (HPUs) (Processes); a description of the laboratory outputs and an investigation into false positive TB isolation (laboratory Outputs); an exploration of the outcomes of cluster investigations (public health Outputs); a transmission model and a cost-effectiveness model (Outcomes). The Context is considered in the discussion (Chapter 8). These methods are detailed in the remainder of this chapter.

An overview of the evaluation of the TB-STS, including the framework and methods used, is shown in Figure 16. The timeline for the implementation of the TB-STS and the data collection timeline for each of the evaluation components are shown in Figure 17.

Figure 15 – The evaluation framework used here is an adaptation of the Donabedian framework that divides an intervention into the Structures, Processes, Outputs and Outcomes

This is an adaptation of the Donabedian framework in which the intervention is divided into Structures, Processes, Outputs and Outcomes. In addition, the Context is considered across the intervention.

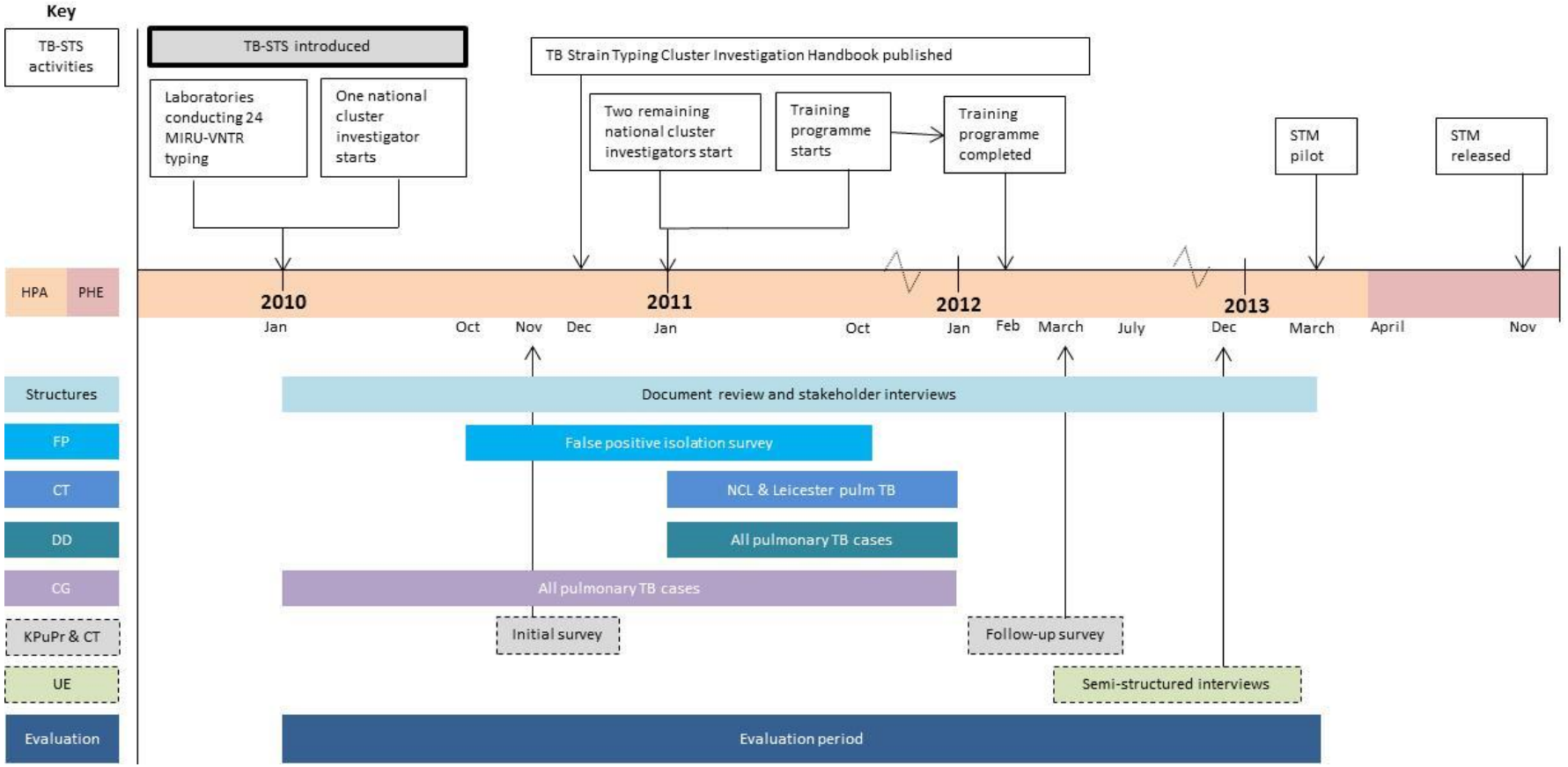
Figure 16 – The evaluation model of the TB-STS



The Structures, Processes, Outputs and Outcomes of the TB-STS are shown in the white boxes with coloured outlines. To their right, the methods used to describe, measure or estimate these elements are shown in coloured boxes. Dotted lined boxes are indirect outputs of the cluster investigations that were hypothesised to interrupt transmission.

STM Strain typing module is the software system for the integration of strain typing data with epidemiological and clinical data; STPB Strain Typing Project Board were the PHE board responsible for the implementation of the TB-STS

Figure 17 – Evaluation data collection in the context of the implementation of the TB-STS



STM strain typing module; FP false positive TB isolation; CT contact tracing yield; NCL North Central London; DD diagnostic delay; CG cluster growth; KPuPr knowledge, perceived usefulness and practices; UE user experience

3.2 Describing the TB-STS

To describe the TB-STS – how it is structured and how it functions – the following activities were carried out:

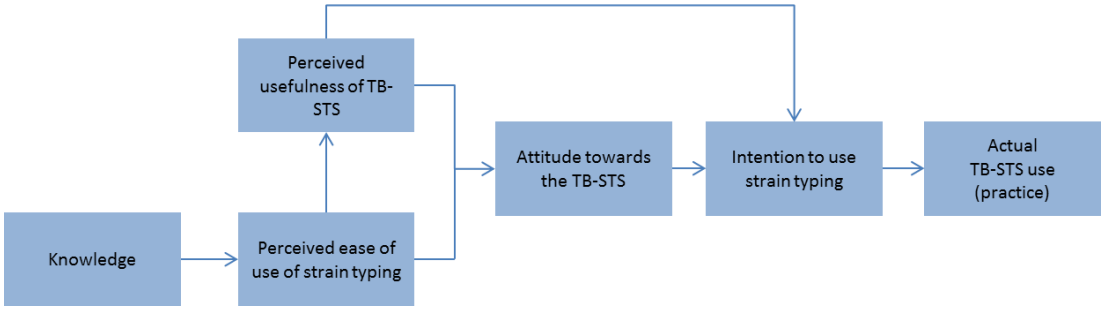
- a search was conducted of the PHE website to find published documents and web pages;
- the PHE intranet was searched for documents produced internally;
- meetings were attended and minutes read from the TB Strain Typing Project Board; and
- discussions/interviews with members of the TB Strain Typing Project Board and the PHE Finance team were conducted by email, telephone and face-to-face over the course of the evaluation.

3.3 Initial and follow-up cross-sectional surveys

The successful implementation of the TB-STS depends not only on laboratory scientists to type the isolates, but also on public health staff to use the cluster reports to inform their decisions, and clinical staff to gather additional information about the clustered patients. Consequently, the engagement of public health and clinical staff with the service is likely to have a direct impact on its implementation and subsequent impact.

The technology acceptance model is a theory for how people will come to accept and use a new technology.[201] It suggests that one's perception of usefulness and ease of use will influence one's attitude towards the technology, which will influence their behavioural intention, and subsequently whether they actually adopt the technology. The technology acceptance model is used here because it suggests that someone's perceived usefulness of a technology can influence their intention to use it, regardless of their attitude towards it. For example, if the TB-STS is perceived to be useful, it may be used, even if people don't like the service. If one assumes that perceived ease of use of strain typing is influenced by an individual's knowledge, then the theory can be illustrated as shown in Figure 18.

Figure 18 – The technology acceptance model, as modified for application to the TB-STS



An initial and follow-up web survey of public health and clinical staff working in TB control was carried out to establish the extent to which the TB-STS had been implemented by measuring the effect of the service on changes in knowledge, perceived usefulness and uptake, assessing the impact of the service implementation on workload, and measuring changes in contact tracing yield over time. The results of this two-part survey were used to guide the semi-structured interviews described below, and helped to develop parameters for the cost-effectiveness analyses.

Objectives of the initial and follow-up surveys:

1. Knowledge: To assess changes in the awareness of the TB-STS and knowledge of strain typing
2. Perceived usefulness: To explore changes in perceived usefulness of the TB-STS
3. Practices: To investigate the practices associated with the TB-STS and capture any changes over time in
    a. Laboratory turnaround times
    b. The extent to which the TB-STS has been implemented
    c. How strain typing is used and its impact on TB control
    d. TB-associated workload
4. Estimate the yield of contacts with TB disease or LTBI identified through contact tracing, and capture any changes over time (for these methods see page 75).
5. General comments: To collect and explore any further comments that survey respondents wanted to share with the researcher

3.3.1 Study Design

An initial survey was conducted in November 2010 prior to the release of the Handbook, using a web-based survey questionnaire (www.objectplanet.com/opinio). A follow-up survey was conducted in March 2012, following the completion of the public health training (see 'training' on page 98). Figure 17 (page 57) shows the timing of the surveys in relation to the implementation of the TB-STS.

For these surveys, it was assumed that prior to the Handbook publication in December 2010, the TB-STS would not have had an impact on clinical staff and health protection professionals and their work in TB control, and therefore the situation could be compared to the *ad hoc* service provided prior to the TB-STS (see page 56). Put simply, the publication of the Handbook marked the initiation of the TB-STS for public health application.

The target population were all public health staff, chest/respiratory physicians and TB nurses working in TB control in England at the time of both surveys. Questions were asked about the knowledge (awareness of the service, training, resources and self-reported knowledge), perceived usefulness of the service, and practice (if and how strain typing is accessed and used, and its associated workload). In addition, nurses were asked about five recent index cases and their contacts to estimate the number of contacts with TB disease and LTBI identified through contact tracing (termed contact tracing yield). This is explained in more detail in the methods for contact tracing yield (page 73). There was an option for open-ended responses at the end of the survey for any general comments. The survey was piloted with a nurse, a physician and a public health specialist to ensure that the questions were clear and interpreted appropriately. All questions and possible responses are available in Appendix 2 and published.[202]

For most of the survey, individual respondents were of interest (e.g. their personal knowledge and opinions of strain typing for TB control and any change in workload following the introduction of the TB-STS), rather than the opinion of a TB clinic, TB

service or HPU. A sample size of 100 was calculated to provide adequate statistical power (80%) to detect a proportion change from 90% to 99% for key implementation questions. The initial survey was emailed to all users of the TB notification system (Enhanced Tuberculosis Surveillance; ETS),[203] and to staff responsible for TB control in HPUs who were asked to pass it on to their local TB teams. Therefore, the sampling frame could not be enumerated. The follow-up survey was emailed to respondents to the initial survey.

3.3.2 Analysis

Participants that responded to both surveys were included in the statistical analysis. Responses from people working at national, regional or PCT-level (including cluster investigators) and people working in Wales were excluded because the survey questions were tailored to people working in HPUs and clinics (not in regional and national offices), PCTs were reorganised in 2012 and did not exist at the time of the follow-up survey and Wales was not part of the TB-STS. These people received the survey because they were registered ETS users.

HPU was used as a proxy for geographical area on which TB incidence rates were based because it was (correctly) anticipated that the geographic jurisdiction would remain following the creation of Public Health England in April 2013, ensuring it would be a comparable unit at the follow-up survey. The TB incidence of the HPU area in which respondents worked was used as a unit of analysis throughout the survey (defined as low, medium and high incidence based on an annual notification rate of <10/100,000 population, 10 to 19/100,000 and ≥20/100,000 respectively). TB incidence categories were defined based on the mean TB incidence rate in England and Wales during 2007-2009 (14.95 per 100,000 population). Mean TB incidence was calculated using the mid-year population estimates from 2008.[204,205]

The knowledge, perceived usefulness and practices of public health and clinical staff working on TB control were compared across the initial and follow-up surveys by calculating and comparing medians and inter-quartile ranges (IQR), and means and

standard deviations (SD), and using two-sample t-tests, chi^2 tests, Wilcoxon rank sum tests, and logistic regression, where appropriate. Calculations excluded item non-responses. Analyses were conducted overall, by professional category and the TB incidence of the HPU area in which respondents worked (low, medium and high incidence defined above). Statistical analyses were considered significant at the 95% level ($p \leq 0.05$). The characteristics of non-responders and responders were compared using logistic regression for binary or ordered variables where appropriate.

Knowledge

To explore the awareness and knowledge public health and clinical staff had about strain typing and its uses, responses to questions across the two surveys regarding whether respondents had heard of the TB-STS, had access to strain typing data, had received any training, or had access to resources or tools to help them interpret and use strain typing data were compared. They were then asked to rate their knowledge of how to interpret the MIRU-VNTR data on a scale of one to five, one representing no knowledge and five representing excellent knowledge. These questions attempted to establish the extent to which the TB-STS training strategy had been implemented and whether professionals working in TB felt well-equipped to use strain typing.

Perceived usefulness

Survey respondents were asked to rate how useful they found the strain typing information: very useful, quite useful, not very useful or useless.

Practices

Practices (a): Laboratory turnaround times

One of the key elements for an efficient service is the time it takes for the strain typing result to reach the local public health teams. Nurses and physicians were asked "how many weeks after a sample is sent to the lab do you receive the individual strain type?" Health protection staff were asked "do you usually access/get hold of strain typing data before or after you have finished contact tracing?" It was not possible to compare the distribution of responses and mean number of weeks

reported across the two surveys because none of the respondents included in the panel data responded to the question in the initial survey. Therefore, given the item non-response, the follow-up data alone are presented.

Practices (b): How strain typing is used and its impact on TB control

The questionnaire explored the way in which clinical and health protection professionals used the TB-STS, i.e. their strain typing-related practices. This was an opportunity to begin to understand the Processes within the TB-STS, the way strain typing data are used and how it impacts the management and control of TB. Questions were asked about how strain typing was accessed and used, and whether it influenced the TB control activities of the different professional groups.

Practices (c): TB-related workload

TB-related workload was measured in order to compare the workload before the implementation of the TB-STS and capture any change in workload that might be attributed to the TB-STS. Survey respondents were asked a variety of questions to capture their workload associated with TB. The questions posed to nurses, physicians and health protection staff varied according to their specific roles. Estimated workload was used to define the human resource cost of the TB-STS for the cost-effectiveness model.

Nurses were asked about their TB caseload including: the number of new active TB cases, the number of TB cases that needed contact tracing, the number of contacts they had screened in the last month and how many hours this took. Physician TB workload was measured by the number of new TB cases started on treatment by the physician each month and the frequency of TB incident meetings attended by the physician. The workload of health protection staff working in HPUs was estimated by the number of cluster investigations in the HPU in the last month and the proportion of time spent on cluster investigations.

Estimates of contacts screened with active disease and LTBI

For methods see page 75.

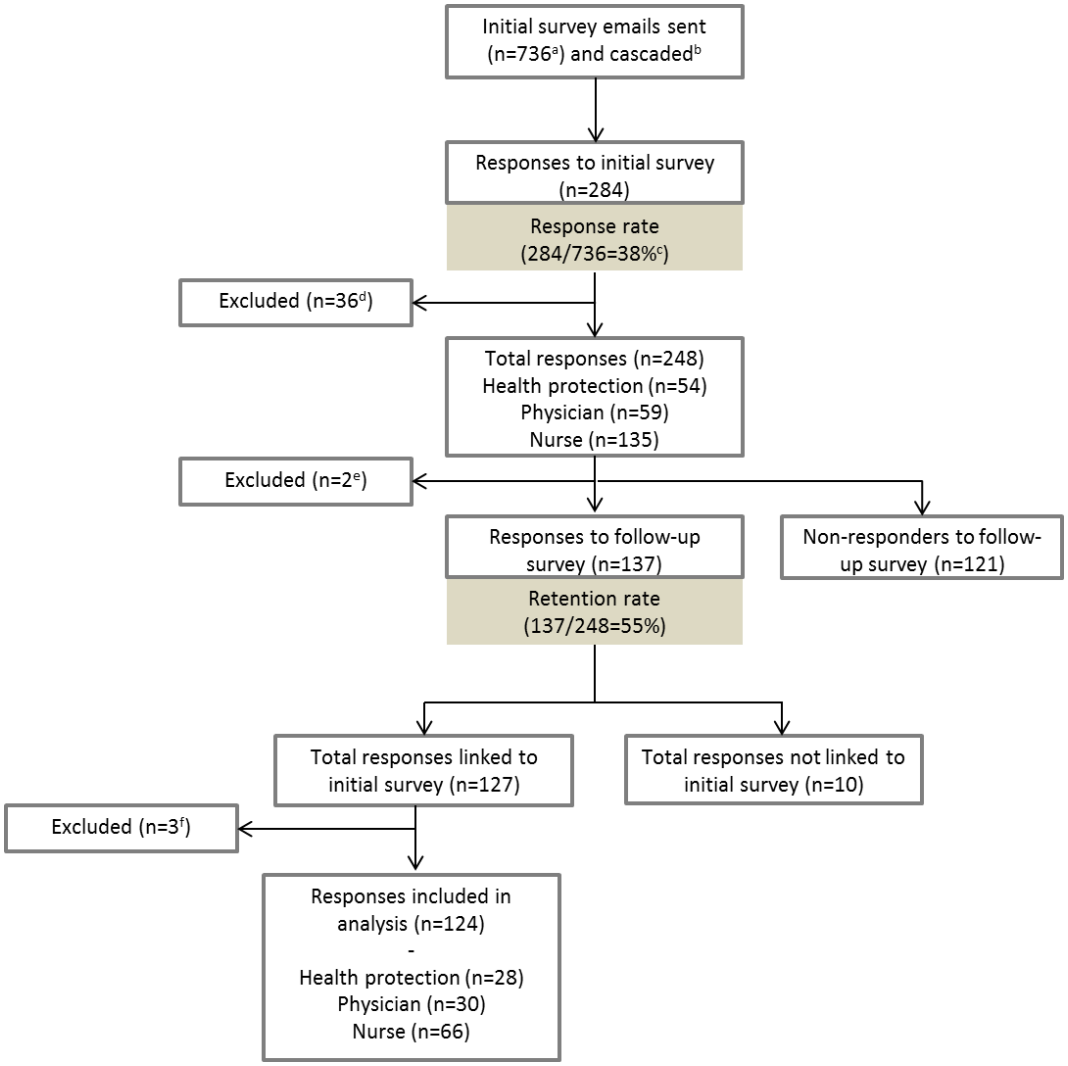
Free text for general comments

The follow-up survey provided an opportunity to collect general comments that respondents felt were not covered by the set questions. This was achieved by including free text space at the end of the questionnaire with the following prompt: "If you have any other comments about how you use (or don't use) strain typing data and how useful you find it please write in the box below." The comments were used to inform the development of the semi-structured interviews described below.

3.3.3 Survey responses

There were 248 responses to the initial survey, 137 responses to the follow-up survey (55% retention), and for 124 there were responses to both the initial and follow-up surveys (Figure 19). Respondents to the initial survey who did not respond to the follow-up survey were not significantly different to those that responded to both: no particular profession, full-time/part-time position or TB incidence was more (or less) likely to respond to the follow-up survey, and there was no significant difference between the proportion of people who had heard of the TB-STS or had access to strain typing at the time of the initial survey (Table 7). Respondents were from all nine regions of England and covered 24 (of 26) HPUs, representing areas of low, medium and high TB incidence.

Figure 19 – Responses to the initial and follow-up surveys



[a] The email was sent to all users of the Enhanced TB Surveillance database. This included all administrative staff as well as well as staff working at national, regional and Primary Care Trust level, which were not relevant to this survey.
[b] It is not known how many people received the email via through the HPU cascade.
[c] This response rate is an underestimation because of the denominator used.
[d] Respondents working at national, regional or PCT-level (n=27) and those from Wales (n=9) were excluded from this analysis.

Table 7 – Characteristics of non-responders and responders to the follow-up survey[a]

		Non-responders to the follow-up survey		Initial & follow-up survey responses*	
		n	%[b]	n	%[b]
Total		121		124	
Profession	HPU	23	19.0	28	22.6
	Physician	29	24.0	30	24.2
	Nurse	69	57.0	66	53.2
TB incidence[c]	Low	50	42.0	56	45.2
	Medium	32	26.9	33	26.6
	High	37	31.1	35	28.2
Work time	Full-time	87	77.0	95	79.2
	Part-time	26	21.5	25	20.2
Heard of the TB-STS, initial survey		100	84.7	105	85.4
Access to strain typing, initial survey		99	81.8	90	72.6

[a] Using the information reported in the initial survey
[b] There were no significant differences between characteristics of non-responders and responders (logistic regression was used to test the association of each characteristic with the outcome non-responder/responder, $p>0.05$)
[c] Area where respondents worked is defined as low, medium and high TB incidence defined as <10/100,000, 10-19/100,000, ≥20/100,000 population, respectively

3.4 Semi-structured interviews

The initial and follow-up surveys were useful in assessing the implementation of the TB-STS, but they did not give much insight into the user experience. Free-text comments written in the follow-up survey suggested that to help understand how the service was being used and the benefits and barriers to the service, the experiences of those using the TB-STS could be further explored.

Semi-structured interviews were used to understand more about the public health component of the TB-STS and the overall value of the TB-STS for health protection.

Rationale for using semi-structured interviews

The process undertaken to describe the TB-STS and the initial and follow-up surveys identified some areas for more in-depth exploration. The most appropriate approach to engage directly with service users and to learn about their different experiences with the TB-STS was through semi-structured interviews that could be conducted over the phone, at a time convenient to the interviewee.

Semi-structured interviews provided an opportunity for more in-depth information to be gathered (compared to data collected from the survey); the interviewee is able to

influence the topic, making the process responsive to the ideas and opinions of the respondent and allowing for unanticipated issues to emerge; further questions can be asked to understand the perspective of the respondent; and it allows a more natural conversation to take place.

Data collection

Semi-structured interviews were carried out with the main TB strain typing contact at each HPU. The aim of the interviews was to explore the different processes that are employed around the country to process, disseminate and act upon strain typing information.

The main TB strain typing contact at each HPU was identified through the national and regional CIs. Using the comments from the follow-up surveys, pilot interviews were designed and carried out with the TB strain typing leads at two HPUs during October 2012 to inform the development of the semi-structured interview guide (Box 2). The interview topic guide was adapted iteratively throughout the interview process, as new topics of interest emerged. The guide included questions on how and when cluster reports are received, what the HPU does with the information in the cluster report and suggestions for improving the service.

During November and December 2012 the remaining 24 HPU leads were interviewed. Telephone interviews lasted 20-30 minutes. 17 of 24 interviews were recorded (no recording devices were available for the other seven for which notes were taken during the interviews).

Box 2 – Interview guide for interviews with HPU strain typing leads

Interview guide for interviews with the HPU strain typing lead:

1. What is your role in the HPU? Job/profession?
    a. Are you sent cluster reports (from lab and/or cluster investigators)?
    b. How many people receive cluster reports (from lab and/or cluster investigator)?
2. From whom do you receive cluster reports? Labs? Cluster investigator?
3. What do you do with a cluster report? Outcome reporting form?
    a. When do you receive it? Is this too early/too late/the right time to be useful to you? Who is involved in looking at it?
    b. Does the information within it affect your decision making? What is it most useful for?
    c. If you act on it, do you use the strain typing handbook as a guide? Are the thresholds appropriate for your area?
    d. What does 'cluster investigation' mean to you? What does it involve? Who decides when an investigation should be launched (cluster review meetings)?
    e. Who does it involve? Do you use the questionnaires (why not)?
    f. How much time does it take?
    g. What do you do with the information?
    h. How does it affect your decision making (local outbreaks identified? Helped to solve a known outbreak?)?
        i. Please give examples of when it did
        ii. Please give examples of when it didn't
4. Are there ways you think the service would work better or the strain typing would be more useful?
    a. E.g. Would it be better if you used the strain typing when you felt you needed it, rather than you being advised by cluster investigators when to investigate?
    b. Are enough resources available to use the strain typing as you would wish?

Analysis

Transcripts and notes from the semi-structured interviews were analysed using a thematic analysis based on the Framework approach.[206] This approach was chosen because the analysis can be based on the predefined themes from the initial and follow-up surveys, as well as allowing for new themes to emerge from the interview data. This approach includes the following steps: familiarisation with the data (i.e. reading over the transcripts and notes); developing a thematic framework; coding the transcripts according to the framework and revising the framework as more data are analysed; once all data were coded, themes and sub-themes were refined and a descriptive analysis presented. These steps were followed with a bottom-up analytical approach, so that the shape of the analysis could inform the remaining interviews. The analysis was conducted using Excel.

The analysis was carried out alongside the interviews so that themes identified in the analysis could inform the remaining interviews. The themes were quality-checked, validated and discussed with the PhD supervisors and relevant members of the evaluation steering group throughout the analysis. Themes were revised where appropriate and additional themes were added when they arose. The interviews that had already been analysed were re-visited to examine how the data related to the recently added themes. Once all 26 interviews had been coded into themes, more detailed sub-themes were identified and charted. The HPUs were grouped into low (n=14), medium (n=7) and high (n=5) TB incidence areas for the analysis.

3.5 False positive TB isolation survey

False positive isolation may be a result of contamination in the clinic or the laboratory, mislabelling or mishandling of clinical samples.[114] One of the drivers for the initiation of the TB-STS was that the service would aid the identification of false positive isolation, saving the costs associated with the unnecessary treatment of false positive cases. Strain typing can identify (and confirm) suspected false positive TB isolates by showing that two or more isolates processed at a similar time in the source or reference laboratory have the same strain type. It can also confirm or refute a suspected incident of false positive isolation that was identified through another channel such as clinical suspicion, by showing that they have indistinguishable or different strain types and preventing the investigation of a false positive outbreak.

Data on suspected false positive isolates were collected from all three reference laboratories between October 2010 and 2012 (see Figure 17, page 57) based on their individual laboratory false positive isolation protocols. The protocol developed by the Birmingham laboratory was shared between the laboratories but not adopted formally.[207] A data collection form was developed and completed quarterly to establish the number of isolates typed in the previous quarter, the number suspected of being false positive isolates or not because of the strain typing data, and the number confirmed to be false positive TB isolation (or not) because of the strain typing data (Figure 20).

Figure 20 – False positive isolation data collection form sent quarterly to the three reference laboratories

A1	Total number of TB isolates strain typed in past three months	
B1	Total number of isolates queried or suspected of being false positive TB isolates	
B2	Number queried because of strain typing data	
B3	Number queried for other reasons (not due to strain typing data)	
C1	Total number of isolates found to be false positive TB isolates	
C2	Number for which strain typing data contributed to the confirmation of false positive TB status	
C3	Number for which strain typing data **did not** contribute to the confirmation of false positive TB status	

Each incident of false positive isolation suspected between October 2010 and October 2012 was recorded alongside the outcome of the reference laboratory's communication with the source laboratory. The outcome of the communication was categorised as possible, unconfirmed and unlikely cross-contamination. A short email questionnaire was designed for each of these outcome categories (Appendix 3). Email questionnaires were sent to the source laboratory to establish whether they were aware of the incident before they were contacted by the reference laboratories, if they carried out an investigation, and if so, the outcome of their investigation and its impact on their laboratory protocol. These data were used as a parameter in the cost-effectiveness model of the TB-STS.

3.6 Capturing cluster investigation activity and outputs

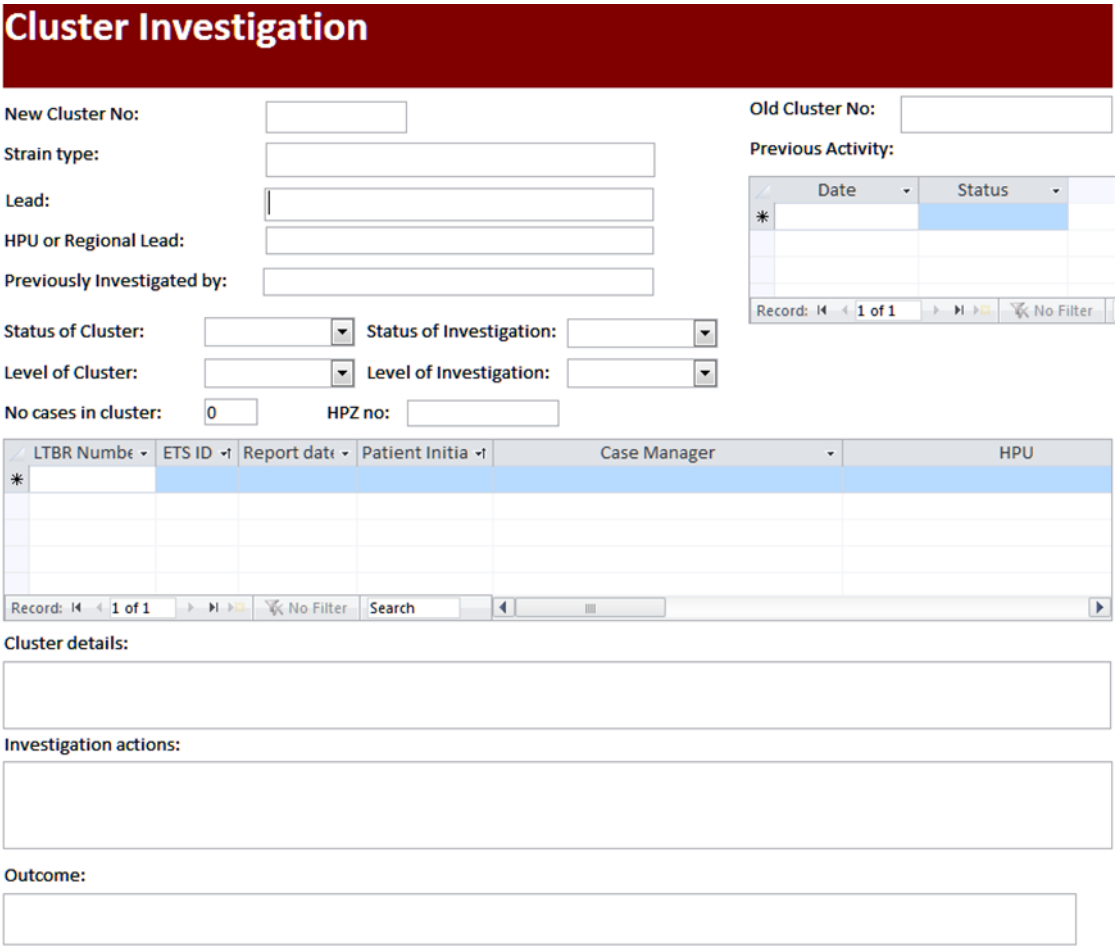
Public health Outputs of the TB-STS include the public health and TB control activities carried out, or not carried out, as a result of the strain typing information. A measurable output is the number of cluster investigations carried out and the outcomes of those investigations. Cluster investigations are described in Chapter 4 (page 92). The next few paragraphs describe the different data collection strategies explored to capture cluster investigation activity and their outcomes: the strain typing

module; the cluster monitoring database; cluster outcome reporting forms; and the database of national clusters. Personal identifiers were removed for the analysis of these data.

3.6.1 The cluster monitoring database

The software for collating and reporting the strain typing and epidemiological data (the STM) had not been developed in time (see Chapter 4, page 91), therefore a cluster monitoring database was developed to collect primary data specifically for this evaluation. This enabled CIs to make a record of each cluster that was investigated and the outcomes of those investigations (Figure 21). The cluster monitoring database collected information on all the clusters that were formally investigated at a local (HPU), regional and national level. It recorded the number of cases per cluster, the date the cluster started to be investigated, the date it closed, the action taken during or as a result of the investigation, and the ETS ID numbers of the cases in the cluster (which allows it to be linked to the national dataset). Data were input following the receipt of cluster questionnaires from HPUs and any other communication with HPUs. The cluster monitoring database enabled CIs to keep a record of the investigations they were initiating and was a shared file, accessible by all three CIs. The database was developed using Microsoft Access.

Figure 21 – Cluster monitoring database: cluster entry form



3.6.2 Cluster outcome reporting forms

Direct measures of the outcomes of interest from cluster investigations are the number of additional epidemiological links identified between clustered patients and the number of secondary cases identified as a result of the strain typing information. Cluster outcome reporting forms were developed to capture the outcomes of local and regional cluster investigations. To report on the activities and outcomes of cluster investigations, these forms were completed and sent to the CIs every six months until the investigation closed. The information collected on the forms is reported and described.

3.6.3 Database of national clusters

The national TB cluster team developed an additional database of national clusters that contains more detail. The database records all clusters of two or more patients that are from more than one region. It holds information on whether there are any epidemiological links between the cases and whether the links were found following a cluster investigation. A descriptive analysis was conducted.

3.7 Indirect measures of the impact of the TB-STS on TB control: secondary data analyses

Due to the restricted cluster investigation activity and outcome data (described above), alternative approaches to measure the impact of the TB-STS on TB control are presented in the following section. These alternative methods and data sources are explored based on the assumption that strain typing will enable targeted cluster investigations and as a result one would expect to identify a greater number of LTBI and cases of TB disease earlier. It was hypothesised that this would lead to three indirect outcomes (see the diagram of the aims of the TB-STS, Figure 11 on page 31):

1. An increased yield of LTBI and TB disease identified through contact tracing when the index case is in a cluster that is investigated (compared to when the index case is in a cluster that is not investigated);
2. A reduction in the time between symptom onset and diagnosis, termed diagnostic delay; and
3. A change in the rate of cluster growth following a cluster investigation (compared to no investigation).

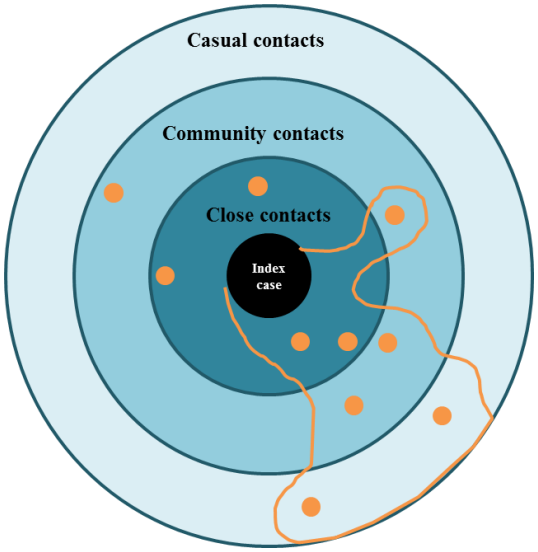
All data were cleaned in Excel (2010) and Access (2010).

3.7.1 Contact tracing yield

Contact tracing yield is the proportion of contacts who have LTBI or active TB disease. It was hypothesised that the TB-STS would increase contact tracing yield. If a case is in a cluster and the cluster is investigated then one might expect more TB

disease and LTBI to be identified, as a result of the additional information about the transmission occurring within that cluster. Figure 22 shows how more targeted contact tracing (because of the additional information collected in a cluster investigation) could produce a higher yield of active TB and LTBI, rather than following the traditional stone in the pond principle (Figure 8, page 8).

Figure 22 – More targeted contact tracing (shown in orange) as a result of the TB-STS might lead to a greater contact tracing yield (as compared with approach illustrated with the blue/green concentric circles)



To explore this, the outcomes of contact tracing were investigated using two different datasets. The first used data collected from nurses in the initial and follow-up survey to summarise the number of contacts screened per index case and the outcomes of the screening (see page 58 for a description of the surveys, and see below for a description of the data collected, and Appendix 3 for the survey questions). The second data source combined data from North Central London Sector and Leicester TB clinics and was used to analyse the impact of cases being part of a cluster and/or part of a cluster that was being investigated on the contact tracing yield (see below). Means, medians and proportions are all presented, as contact tracing yield is a key parameter for the cost-effectiveness model. Figure 17 (page 57) shows the data collection time periods in the wider context of the TB-STS implementation and evaluation.

Overall estimates of contact tracing activity and yield: survey data

Using the initial and follow-up survey (page 53), the aim was to estimate the yield of contacts with TB disease or LTBI identified through contact tracing, and capture any changes over time.

Nurses were asked to detail up to five recent TB index cases whose contacts had been screened. They were asked for the following information for each case: date of birth, site of disease (pulmonary or extra-pulmonary), smear test result, culture result, when the strain type was available to the nurse, the number of contacts identified, and the number of contacts screened. They were also asked to report the number of contacts that had active TB disease and the number of contacts that had LTBI.

Means, medians, standard deviations, inter-quartile ranges or ranges were presented and results were compared across the initial and follow-up surveys using the Wilcoxon rank sum test. The proportion of contacts screened, with active disease and LTBI and the exact confidence intervals based on the Binomial distribution were calculated and compared across surveys using the chi^2 test of significance and the Fisher's exact test of significance where one or more cells has an expected frequency of less than five.

The impact of cluster status and cluster investigations on contact tracing yield: clinic data

Data collected from North Central London (NCL) Sector and Leicester TB Services that contained information on the number of contacts screened per patient and the outcome of that screening in July 2012 for pulmonary TB cases diagnosed in 2011, were used in this analysis. These data were linked to the ETS and the cluster database. For each case reported in these areas the number of their contacts that were screened was calculated and the outcomes of the screening were summarised (e.g. three contacts with LTBI and zero with active disease). The number of contacts screened, the number with active disease and the number with LTBI were calculated

and compared between cases that were unique (not clustered), clustered and part of a cluster investigation, and clustered but not part of a cluster investigation.

Medians and inter-quartile ranges (IQR) for the contact tracing yield for pulmonary index cases are presented by clustering and whether the cluster was investigated. The Wilcoxon-rank sum test was used to compare differences between the groups. A sensitivity analysis was conducted assuming that index cases with missing contact tracing information yielded no cases of active disease or LTBI. This was based on the assumption that those with positive results would be more likely to be recorded. The proportions of contacts screened, with active disease and LTBI and the exact confidence intervals based on the Binomial distribution were calculated and compared across unique/clustered and investigated/not investigated clusters using the chi^2 test of significance and the Fisher's exact test of significance where one or more cells had an expected frequency of less than five.

3.7.2 Diagnostic delay

Diagnostic delay is the time between symptom onset and case notification. It was hypothesised that the diagnostic delay would be reduced following the introduction of the TB-STS because cluster investigations would lead to undiagnosed TB cases being actively identified earlier. A case-control design was used to compare the diagnostic delay observed with cluster investigations and without. Diagnostic delay was defined as the number of days between onset of symptoms and the date of notification. Date of notification was used instead of date of treatment as it was a more complete field.

Cases were defined as pulmonary TB cases diagnosed in 2011 that were part of a cluster that was investigated and were diagnosed after the cluster investigation was initiated. There were two comparison (control) groups: a) pulmonary cases diagnosed in 2011 that were part of a cluster that was investigated and were diagnosed before the cluster investigation was initiated; b) pulmonary TB cases diagnosed in 2011 that were part of a cluster that was not investigated. The first two cases from each cluster

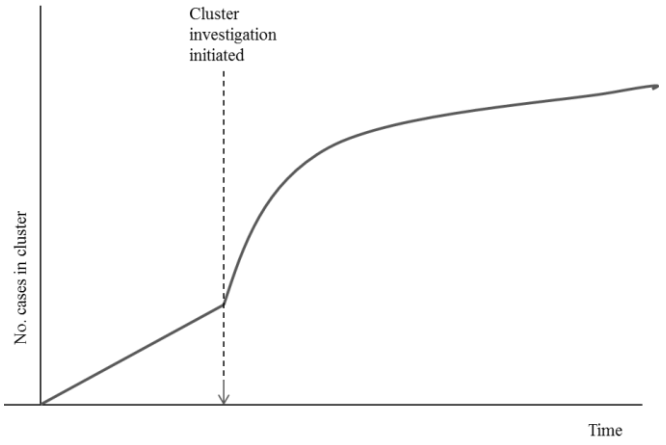
were excluded to take into account possible household transmission (as the diagnostic delay for these two cases will not be affected by the presence or absence of a cluster investigation).

The analysis included 121 pulmonary cases diagnosed after the cluster was investigated, 117 diagnosed before the cluster was investigated and 139 cases that were part of clusters that were not investigated at all. The analysis was stratified by cases that were UK born and non UK born, and was re-run excluding children under the age of 16. Medians and inter-quartile ranges (IQR) were presented and compared using the Wilcoxon rank sum test.

3.7.3 Rate of cluster growth

The rate at which new cases of TB are added to a cluster indicates the rate of ongoing transmission in the community. The relative changes in the size of clusters were explored. It was hypothesised that the rate of cluster growth would differ before and after a cluster investigation was initiated (Figure 23). Cluster investigation may increase apparent rate of cluster growth (as a result of additional cases being identified more quickly) or may decrease the rate of cluster growth (if the earlier identification of cases limits transmission). In ideal circumstances a cluster investigation may transiently increase the rate of cluster growth, later decrease the rate and then plateau as transmission is interrupted (Figure 23).

Figure 23 – Hypothesised rate of cluster growth prior to, and following a cluster investigation



It was hypothesised that the rate at which new cases were added to a cluster would change following the initiation of a cluster investigation. To visualise the data, using a selection of the national TB notification data from 2010 and 2011, the number of cases in a cluster was plotted by the number of days before or after the cluster investigation was started.

2010 and 2011 data from the ETS were merged with the cluster monitoring database. After cleaning the data and excluding the first case in each of the 113 clusters, there remained a total of 949 cases. A univariate linear regression was initially conducted and the final set of variables used in the multivariate regression were: sex, UK born, age group (treated as an ordinal variable), site of disease, case order (the order in which cases were added to the cluster), cases discovered after a cluster investigation was started, and cluster level (local, regional and national) and lineage, the latter being factor variables. See Appendix 4 for the model equations.

3.8 Transmission model

The hypothesised Outcomes of the TB-STS were:

- The TB transmission prevented through the earlier identification and treatment of cases with TB disease and LTBI;
- The costs saved through interrupting transmission and preventing cases; and
- The costs saved by identifying and not treating false positive TB patients (see Figure 11, page 31).

A deterministic age-structured model was used to explore the possible reductions in TB incidence as a result of the TB-STS over a 20-year period (Figure 24). Details of the model are provided in Appendix 5. This extends previous models considering the transmission dynamics of *M. tuberculosis* in England and Wales[112] and recent work on preventive therapy.[208] The model incorporates contact between individuals and rates of immigration and emigration based on Office for National Statistics data.[209] For simplicity, the model considers only pulmonary TB and considers three different epidemiological scenarios: low, medium and high incidence (notification rates of ~7,

~20 and ~120 per 100,000) – comparable to that in the white UK born population (with decreasing annual risk of infection (ARI)), non-white UK born (0.1% ARI), and the non-UK born (1% ARI).

In the absence of the TB-STS, the average diagnostic delay is assumed to be 10 weeks (as estimated for cases that were not in clusters; Table 32, page 155). The effect of a one-week reduction in diagnostic delay due to the TB-STS was explored and it was assumed that patients start TB treatment (on average) two weeks after diagnosis.

In the absence of the TB-STS, 3% of all (recently or latently) infected individuals are assumed to have been detected each year. This proportion is unknown, but was probably low given the low number of latently-infected contacts per index case (see Table 28 page 150 and Appendix 5). The effect of assuming that it increased to 13% after the introduction of the TB-STS was explored. It was assumed that uptake of preventive treatment (PT) among those eligible is 95% and that 85% complete the course of treatment.[48] The model parameters can be found in Appendix 5. In sensitivity analyses, the effect of pessimistic and optimistic assumptions about uptake (30% and 100%, respectively) and completion (50% and 100%, respectively) of PT were explored.

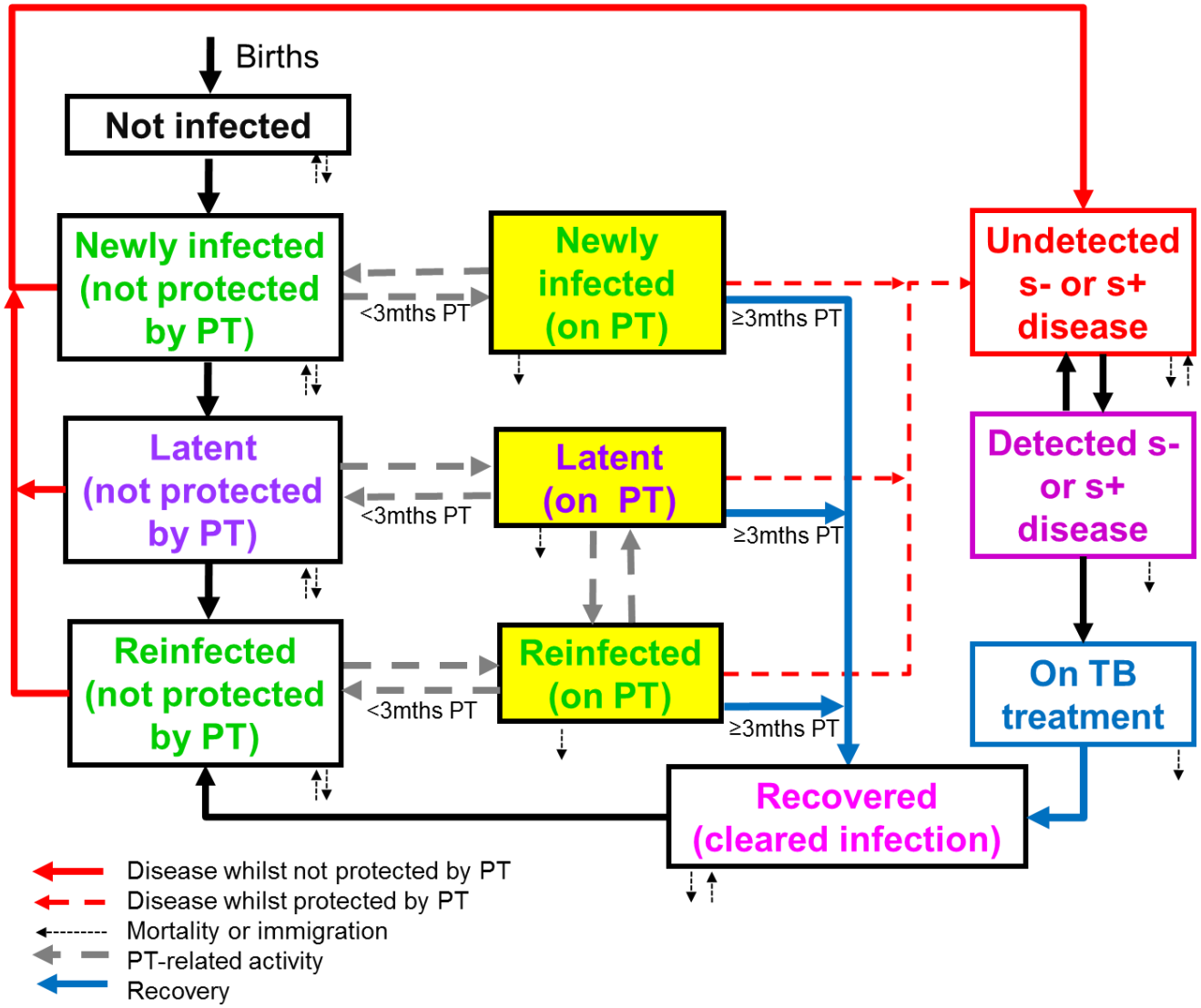


Figure 24 – General structure of the transmission model

PT, preventive treatment. Coloured text and shading is used to reflect similar categories of people: yellow shading is used for people on PT, purple text is used for compartments for people with latent infection, green text is used for newly infected or reinfected people, red is used for diseased people and orange, blue and pink text is used for detected cases, people on treatment and the recovered respectively.

3.9 The cost-effectiveness model

The cost-effectiveness model builds on the transmission model by using the estimates of effectiveness. It incorporated estimates of contact tracing yield, false positive TB isolation and costing collected or estimated as part of this thesis. The aim was to estimate the cost-effectiveness of the TB-STS as an addition to the current system for TB control in England. The cost-effectiveness analysis is illustrated in Figure 25. Further details of the model can be found in Appendix 6.

The analysis adopted a public sector perspective. Estimates of the cost of setting up and operating the national TB-STS were made based on information from Public Health England and the TB Reference Laboratories (see Table 9, page 100). Capital costs were annuitized over an assumed ten-year lifetime for equipment. Costs and health effects (in QALYs) were estimated over a 20-year time horizon, and applying a 3.5% annual discount rate to both costs and QALYs (as recommended by NICE).[210]

The estimated impact of the TB-STS on false positive isolation is presented in Table 20, page 138. For the economic analysis, it was assumed that five cases of unnecessary treatment would be avoided per year due to the TB-STS (ten cases per annum were tested in the sensitivity analysis).

The estimated impact of the TB-STS on contact tracing workload was estimated from the cross-sectional surveys and is presented in Table 16, page 110. In the economic analysis, an opportunity cost for additional time spent by HPU staff on cluster investigations was assumed: 4.4% working time equivalent (WTE) for 26 CCDCs at £99,000pa, costing a total of £113,256 per year (total annual cost of £50,000 per year and £500,000 per year tested in the sensitivity analysis).

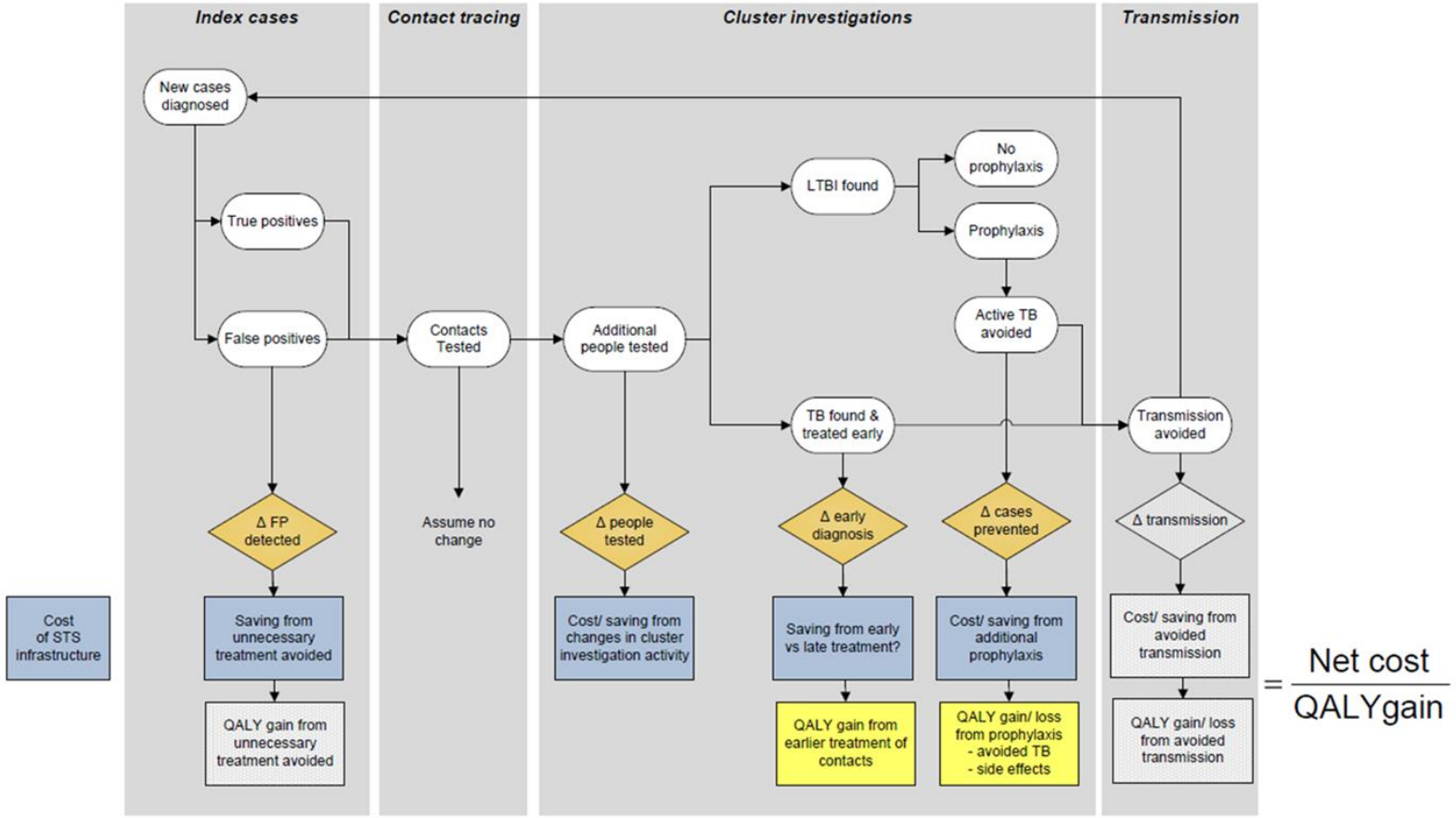


Figure 25 – Structure of the cost-effectiveness analysis

STS TB Strain typing service; FP false positive TB isolation; QALY quality adjusted life year; Δ change

The estimated impact of the TB-STS on case finding and contact tracing yield was estimated from the contact tracing yield analysis and is presented in Table 29, page 151. These results were used to calculate an estimate of the number needed to screen to diagnose LTBI (number screened/number LTBI identified). The estimate was weighted to take into account the number of unique cases, clustered and investigated cases, and clustered but not investigated cases. The number of active TB cases per index case was estimated, based on expert judgement and the results of the contact tracing yield analysis.

The results of the transmission model provided estimates over the 20-year period for each modelled scenario: the number of contacts with LTBI identified, the number of contacts starting preventative treatment, the number of people with active TB diagnosed and starting treatment, as well as the impact on the number of incident TB cases.

The costs associated with diagnosis and the treatment of latent and active TB were estimated based on recommended practice and expert opinion. The unit costs of tests, medications, outpatient contacts and inpatient stays were obtained from national sources which estimate the national average units costs to the NHS: Department of Health Reference Costs 2010-11 for Tuberculosis Specialist Nurse contacts, outpatient consultations (respiratory clinic), and inpatient admissions; British National Formulary, Sept 2012 for medications; and personal communication from PHE Finance Division for theTB-STS.[211,212] Patients with TB who drop out were likely to be identified and offered treatment again at a later time. Such repeat cases were included in the transmission model estimates of the number of people diagnosed per year, and incurred additional costs for diagnosis and treatment in the cost-effectiveness analysis. For simplicity, it was assumed that the cost of diagnosis and treatment was the same for new and repeat cases.

QALYs lost due to TB-related mortality are estimated based on: TB incidence by age;[123] case-fatality rates by age group;[213] life expectancy (ONS); and mean quality of life by age (EQ5D scores) in the general population (Health Survey for England).

The case fatality rates were taken from an analysis of national surveillance data linked to mortality data, with capture-recapture methods used to estimate the number of unascertained deaths.[213] In this analysis, case fatality was defined as a death within 12 months of the start (or notification) of TB treatment, and where TB was mentioned on the death certificate or if treatment outcome monitoring had stated that the death was caused by or contributed by TB. This includes deaths in which TB was reported as a contributory factor, as well as deaths directly caused by TB.

The results are presented in the form of an Incremental Cost-effectiveness Ratio (ICER), which is the additional cost per additional QALY gained with the TB-STS. The difference in costs and the difference in health effects with/without the TB-STS are estimated (Box 3). Any costs or health impacts incurred under both systems were ignored. The ICER was compared with the NICE threshold of £20,000 to £30,000 per QALY gained to determine whether the TB-STS was cost-effective or not.[210]

Box 3 – The Incremental Cost Effectiveness Ratio (ICER) for the TB-STS

$$\text{ICER} = \frac{\text{Incremental cost of the TB-STS (£)}}{\text{Incremental effect of the TB-STS (QALY gained)}}$$

3.10 Ethical Considerations

The research conducted for this thesis was classified as a service evaluation by University College London Hospital Foundation Trust and therefore specific ethical approval was not required. Personal identifiers were removed and data were anonymised prior to analyses. Data were stored on the Public Health England secure server, in adherence to the requirements for national surveillance data and the Caldicott principle.

3.11 Summary of evaluation methods

This prospective evaluation was a mixed-methods approach using primary data collection and routine data sources. Primary data collection included cross-sectional surveys of health protection and clinic staff, semi-structured interviews with health

protection staff, surveys of reference and hospital laboratories and the cluster monitoring database. A summary of the evaluation methods and how each component was used to inform or parameterise other parts of the evaluation is illustrated in Figure 26.

All data were cleaned and analysed using Stata version 12 (*StataCorp. 2011*. Stata Statistical Software: Release 12. *College Station, TX: StataCorp LP*.) and Microsoft Excel (2010), unless stated otherwise.

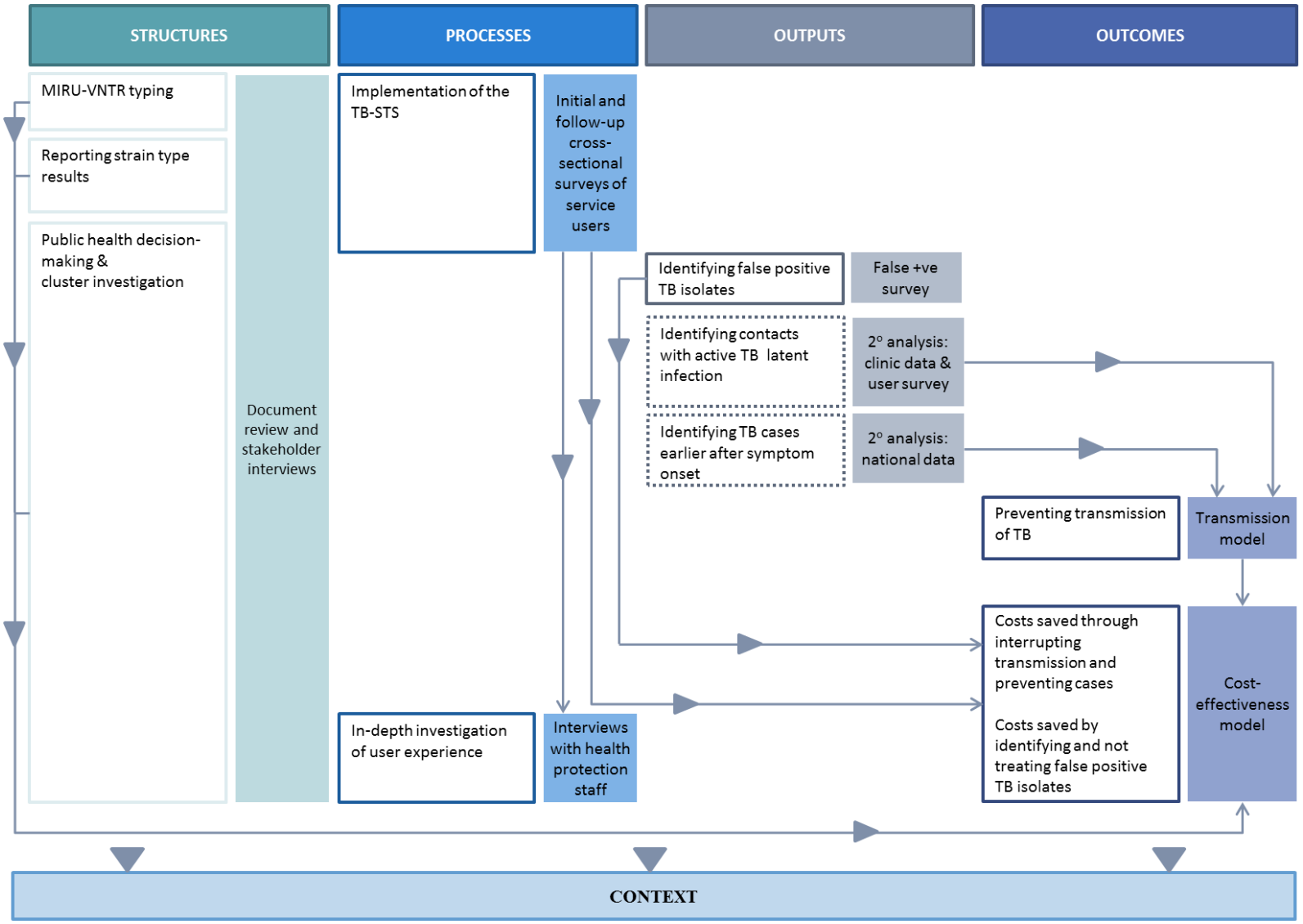
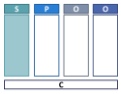


Figure 26 – Interaction of the evaluation components

Arrows represent the flow of data or results from one component to another.
All components are considered in Context of the TB-STS.

Chapter 4. Structures of the TB-STS



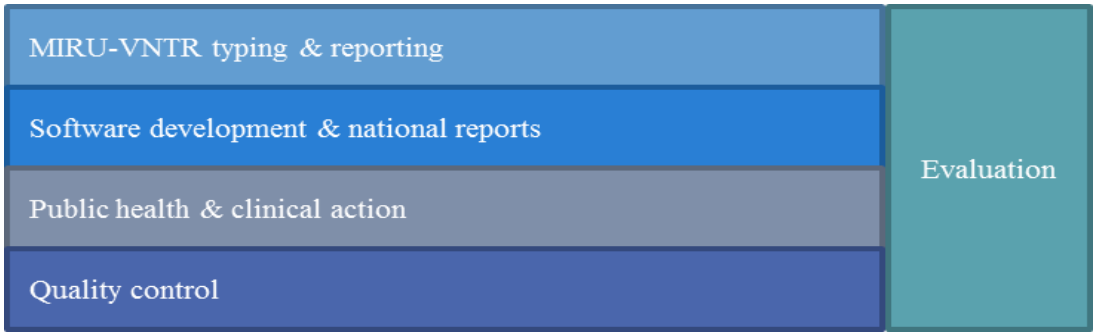
In this chapter I give an overview of the TB-STS. I describe the components of the TB-STS, how it relates to the existing TB control service, and illustrate its implementation on a timeline. Finally, the fixed costs of establishing and running the TB-STS are presented.

—

The TB-STS was based on prospective typing of *Mycobacterium tuberculosis* using 24 loci MIRU-VNTR. It was designed with the intention that strain typing results would be linked to existing national epidemiological and laboratory datasets, and reported in real time to front line teams. It was hypothesised that prospective universal strain typing and analysis of clusters could be used in real time to inform public health action (see Figure 11, page 31).

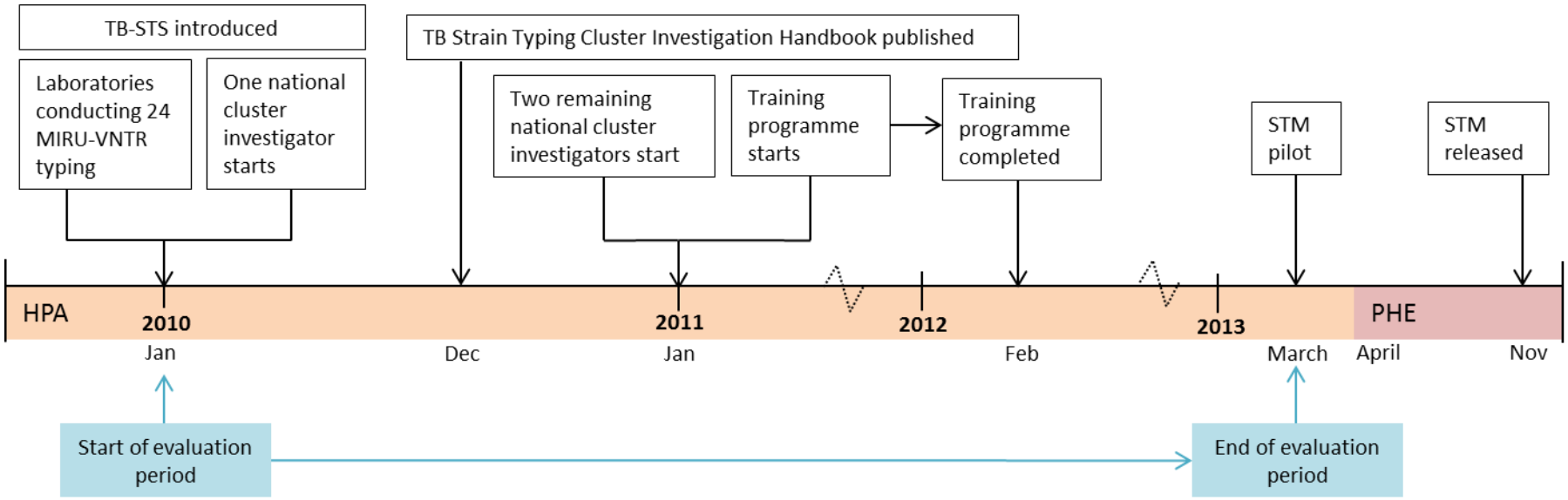
The TB-STS is composed of five operational workstreams: 1) 24 MIRU-VNTR typing and reporting, 2) software development and national reports, 3) public health and clinical action, 4) quality assurance, and 5) the evaluation (Figure 27). Each of these workstreams is described below. Figure 28 shows the implementation of the TB-STS within the first three years, the change from the Health Protection Agency (HPA) to PHE, and the evaluation period. A simplified diagram of the existing TB service and the addition TB-STS is illustrated in Figure 29.

Figure 27 – The five workstreams of the TB-STS



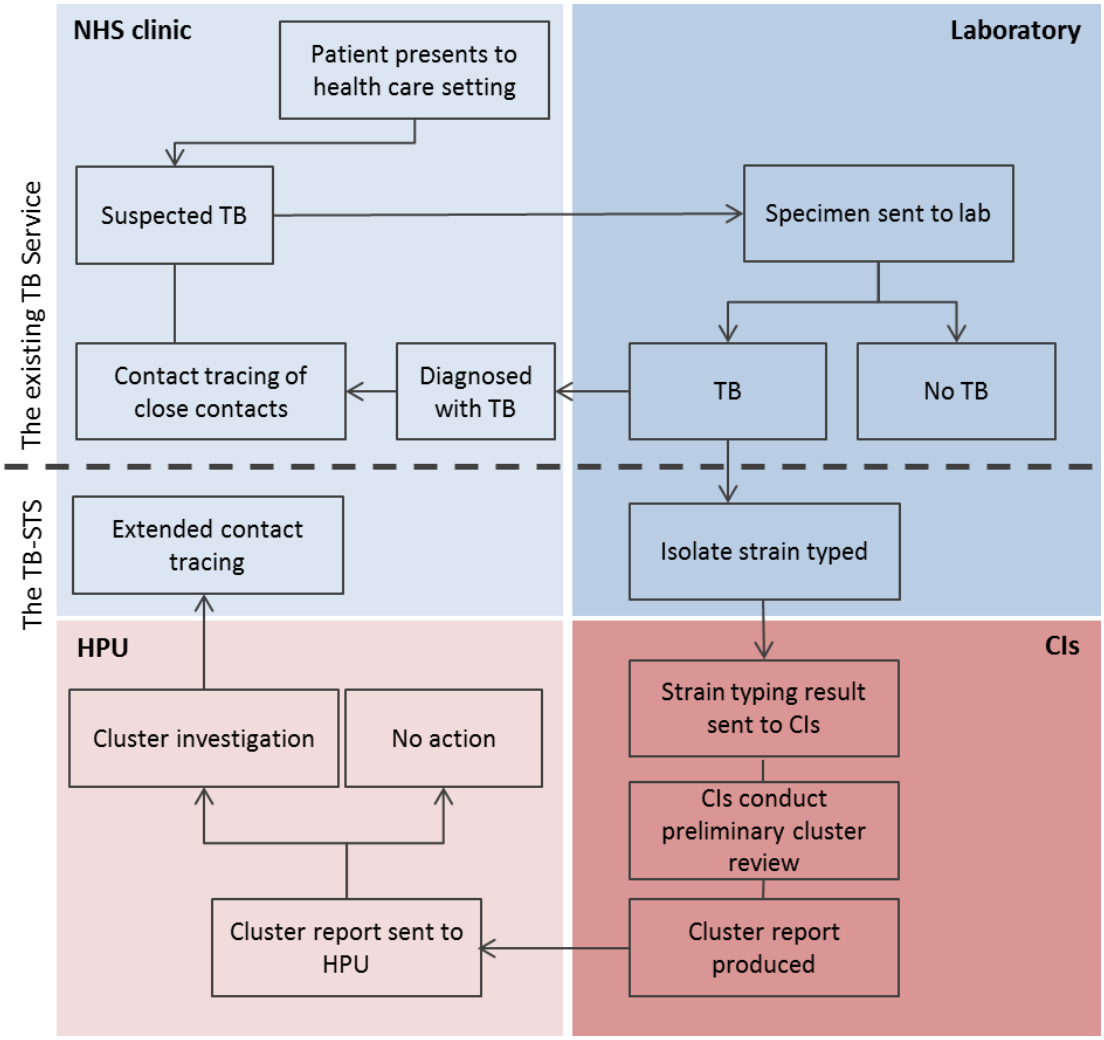
The Evaluation incorporates the remaining four workstreams of the TB-STS

Figure 28 – Timeline showing the change from the HPA (orange) to PHE (red), the implementation of the TB-STS (black) and the evaluation period (turquoise)



STM strain typing module

Figure 29 – Simplified illustration of the existing TB service and the addition of the TB-STS



The components above the dotted line were conducted as part of the existing TB service, and everything below the dotted line is part of the TB-STS. The TB-STS involves the laboratories, PHE national staff (CIs), HPUs, and those working in the clinic setting. These different settings are represented in the different coloured boxes.

4.1 24 MIRU-VNTR typing and reporting

This workstream is initiated when a sample is taken from a suspected TB patient and sent to a local laboratory for culture. Once growth of acid fast bacteria is detected the culture is sent to one of the three Mycobacterium reference laboratories (in London, Birmingham or Newcastle) for identification. If identified as *M. tuberculosis*, drug susceptibility testing will be set up and the first culture per patient will be selected for 24 MIRU-VNTR typing. Subsequent cultures are typed if they are collected more than two months after the first. The document *HPA Mycobacterium tuberculosis*

Strain Typing: A guide to data production and distribution provides more detail on the typing methods and reporting standards.[214]

The 24 MIRU-VNTR strain type for each isolate is reported back to the local laboratory that submitted the positive culture on the standard report containing identification and drug susceptibility results. It is also reported to the local HPU where the patient is being treated and loaded onto MycobNET (the laboratory database held at PHE) so that it can be matched to the Enhanced Tuberculosis Surveillance (ETS) system (the national TB notification system) for surveillance purposes.

Clustering and cluster names are generated using the following algorithm:[214]

i. A cluster comprises two or more isolates with indistinguishable 24 MIRU-VNTR strain types.

ii. For an isolate to be included in a cluster at least 23 loci must have been typed successfully, i.e. the strain type must contain at least 23 digits.

iii. An unequivocal cluster number is given when at least one member of the cluster has a full 24 loci strain type.

iv. Isolates are not included in a cluster if the permitted ≤ 2 missing loci do not at least correspond to all those missing in the most complete member of a cluster.

v. Where missing loci fail to place an isolate unequivocally within a single cluster, typing is repeated.

vi. Prefix letters indicating the phylogenetic lineage are determined.

vii. The cluster designation is completed by adding a unique four digit number following the letter that indicates the lineage.

Cluster reports are sent to HPUs on a monthly basis by the London and Birmingham laboratories. The reports contain new clusters within the HPU, previously reported clusters within the HPU where new cases have been added, and new cases from within the HPU that are part of a regional or national cluster. Cluster reports from all three laboratories are sent to the Cluster Investigators (CIs).

4.2 Software development and national reports

Cluster nomenclature

An online cluster naming tool was developed by PHE using the cluster name designation algorithm (described on the previous page). This tool acts as a repository for cluster names enabling all the laboratories to use the same cluster name for every cluster. When a *complete* 24 loci MIRU-VNTR strain typing profile is submitted, the database of strain types with an assigned cluster name is searched. If the submitted profile matches an existing cluster, this isolate will be added to that cluster. If the submitted profile does not match an existing cluster, a search is conducted of unique profiles stored in the database. Where there is a successful match, a new cluster is created and a name is assigned to that complete MIRU-VNTR profile.

When an *incomplete* 24 loci MIRU-VNTR profile is submitted, the cluster naming tool will attempt to provisionally match the submitted profile to any existing clusters with identical incomplete profiles, and provisionally assign the associated cluster name. If the incomplete profile submitted does not match an existing cluster, unique profiles in the database that contain no missing loci will be searched. Where a match can be found a new cluster will be created and a name assigned.

The tool can be found at:
http://www.hpa-bioinformatics.org.uk/TBCluster/tbhome.php

Strain typing module (STM)

The STM is a software module that is integrated into the ETS. The module is designed to identify and report strain typing clusters. It enables the collation and reporting of the strain typing data with the epidemiological and clinical data stored in ETS, thereby identifying strain typing clusters and presenting them with the epidemiological information to assist with cluster investigation. Clinical and risk factor information relevant to cases in a cluster can then be reviewed. The STM contains questionnaires designed to collect additional information on cases that are

part of a strain typing cluster, and record information on cluster investigation. In addition, the STM can be used to search for cases by cluster name or strain type.

The development of the STM began at the start of the National Strain Typing Project in October 2009 and was due to be released at the end of 2010. However, there were severe delays in the development of the software. In September 2012 the STM was piloted with the Newcastle laboratory and the areas it serves, and was released in November 2013 (Figure 28). This was after the completion of the TB-STS evaluation. Prior to this, the essential functions of the STM were carried out by interim systems (Table 8). As a result, the impact of the STM was not captured by this evaluation.

Table 8 – Functions of the STM that were carried out by other means

STM function	Function re-created in absence of the STM
Identify and report strain typing clusters	Laboratories do this using Bionumerics software and send reports by email and letter to CIs and HPUs
Collating and reporting of strain typing data with epidemiological and clinical data from ETS	CIs do this by hand and create cluster reports by hand, which are emailed to HPUs
Clinical and risk factor information relevant to the cases in the cluster can be reviewed	The CIs collate the information together and include it in the cluster report they produce
Questionnaires designed to collate additional information on cases in a cluster	The CIs post or email questionnaires to the relevant HPUs or TB nurses.
Questionnaires to record information on cluster investigation	Information on cluster investigation activities and outcomes are collected using cluster outcome reporting forms which are sent from the CIs by email or post to HPUs.

4.3 Public health and clinical action

Strain typing information can be used to guide public health and clinical action.[197] Public health action refers to any additional action that is taken in an attempt to interrupt TB transmission. It includes outbreak investigation and extended contact

screening. The public health action that is generated by the TB-STS is initiated through cluster investigations. A cluster investigation is the additional investigation carried out around a cluster of TB cases with indistinguishable strain types. Cluster investigations are initiated based on criteria set out in the TB Strain Typing Cluster Investigation Handbook.[197] The criteria, or thresholds, for initiating a cluster investigation relate to the number of cases in the cluster and the characteristics of the cases within the cluster (Box 4). The Handbook was developed as part of the public health component of the TB-STS to guide the public health application of strain typing in England.

Box 4 – Thresholds for the initiation of a cluster investigation

Thresholds/criteria for the initiation of a cluster investigation, as defined in the TB Strain Typing Cluster Investigation Handbook:[190]

1. Where all the cases in the cluster reside within a single HPU, 5 or more persons within 24 months (2 years), of which 2 occurred in the last 6 months; with TB caused by indistinguishable strains;
2. Where the cases in the cluster reside across more than one HPU within a single region, 10 or more persons within 24 months (2 years), of which 2 occurred in the last 6 months; with TB caused by indistinguishable strains; and
3. Where the cases in the cluster reside across more than one region, 10 or more persons within 24 months (2 years), of which 2 occurred in the last 6 months; with TB caused by indistinguishable strains.

If the cluster contains any of the risk factors identified as increasing the likelihood of recent transmission, a cluster investigation should be considered when the cluster contains fewer cases than stated in the above thresholds (see the Legend to Figure 30 and Figure 31, page 95, for a list of the risk factors for recent transmission).

Figure 30 and Figure 31 are from the Handbook[197] and show the flow of information and processes that lead to, and constitute a local and national cluster investigation. The criteria and thresholds for investigation are shown in the legend for Figure 30 and Figure 31. Cluster investigations are coordinated by CIs, with local HPUs responsible for local cluster investigations and the national cluster team responsible for national cluster investigations.

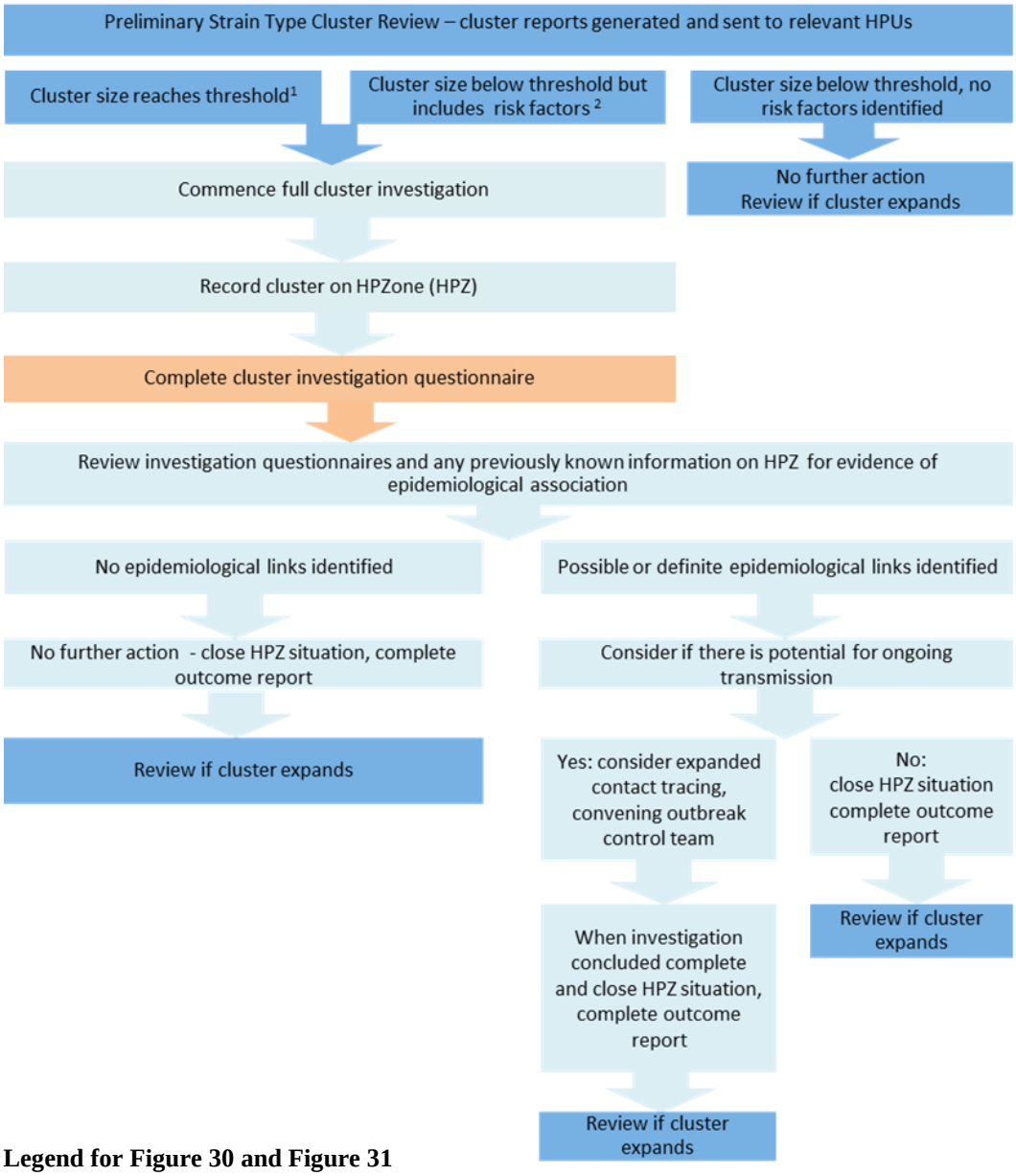
The reference laboratories send cluster reports to the CIs on a monthly basis. The London and Birmingham laboratories also send the cluster reports to the relevant

HPUs. Some HPUs will use the cluster reports sent from the laboratories to inform their decision making. For example, HPUs may use the data to confirm or refute suspected transmission, but many HPUs will wait to receive cluster reports from the CIs as these include recommended actions.

When the CIs receive the cluster reports they carry out a preliminary strain type review. Relevant information about each case is extracted from ETS and, based on the size of the cluster and the characteristics of the cases within the cluster, a judgement is made about whether a cluster should be investigated further and a cluster investigation launched. The summary of the preliminary cluster review and any recommended actions are contained in the cluster report that is then distributed to the relevant HPUs.

From this point, HPUs are responsible for investigating the local clusters identified for further investigation by collecting more information from the clustered patients to try to establish whether there is any evidence for on-going transmission. For national clusters it is the national team that directly contacts TB nurses about their patients. Where there is evidence of on-going transmission, public health interventions may be implemented. Any findings or actions should be reported to the CIs.

Figure 30 – A flow chart showing the flow of information and decision making algorithms involved in a local cluster investigation

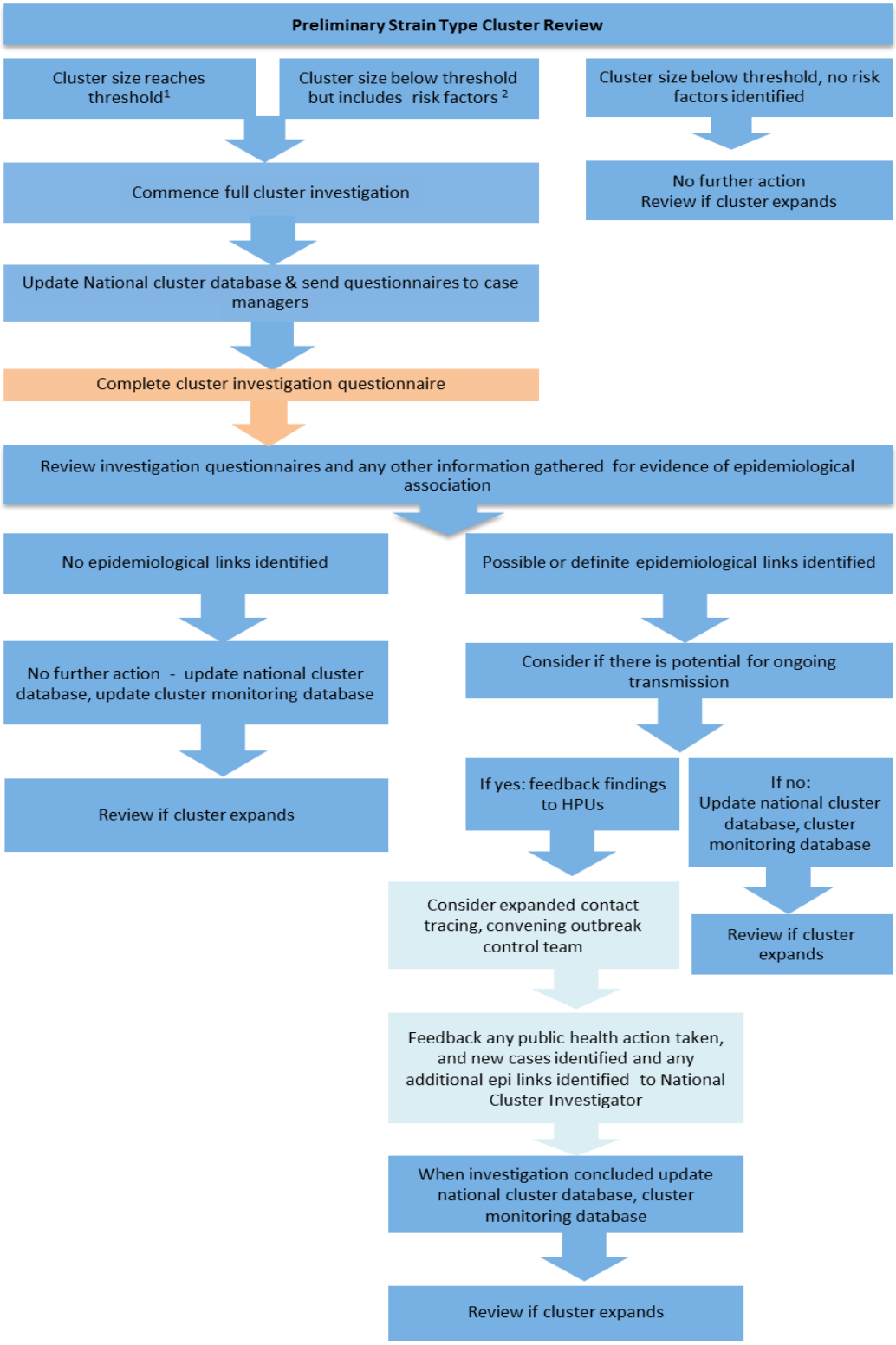


Legend for Figure 30 and Figure 31

Legend/Explanatory Notes	
NHS Lab/NHS TB Team	2 Risk factors: The cluster includes cases who belong to the following groups or gave the following characteristics: • Child (under 16 years) • Health care workers as recorded on ETS • The cluster includes a drug-resistant case • Known HIV infection • Currently homeless/history of being homeless or residence in a hostel or similar temporary accommodation • Recent incarceration in a prison/offender institute • Known previous TB treatment failure • Problem drug or alcohol use • History of severe mental health problems
HPS Cluster Investigators Role	
HPU Role	
1 Threshold ≥ 5 cases in 2 years + 2 in last 6 months	

Source: The TB Strain Typing Cluster Investigation Handbook[197]

Figure 31 – A flow chart showing the flow of information and decision making algorithms involved in a national cluster investigation



Source: The TB Strain Typing Cluster Investigation Handbook[197]

Human resources

The TB-STS created seven new full-time permanent public health positions and one short-term position: one national Cluster Investigator (CI); four CIs responsible for local and regional clusters; one national strain typing scientist; one national strain typing administrator); and one short-term strain typing administrator for 'the rest of England'. To cut costs, two of the local/regional CIs were not appointed, leaving the remaining two responsible for clusters in London and the South East, and the rest of England. This did not include the resources allocated for the evaluation. No additional public health positions were created within HPUs or in the NHS.

Cluster investigations

Three different systems have been developed by the CIs to organise and coordinate cluster investigations, based on their access to administrative and technical support and in lieu of the STM.

1. The national CI leads on national cluster investigations from the PHE Centre for Infectious Disease Surveillance and Control in London, collecting information from nurses and HPUs and advising HPUs on further public health and clinical action that may be required. The national CI is supported by a scientist, an administrator and is line managed by the Head of the TB Section.

2. The CI for London and the South East (SE) sits in the regional office in London and sends out cluster reports advising HPUs which clusters should be investigated. The CI provides support to the HPUs, who lead the local investigations and any further public health and clinical action. Regional clusters in London and the SE and some national clusters are coordinated by the CI. The CI for London and the SE is supported by the London Regional Office and line managed by the London Regional Epidemiologist.

3. The CI for the rest of England sits in the North East region and collates the information for cluster reports before sending them to the relevant HPU with

recommended actions. Local cluster investigations are led by the local HPU, regional cluster investigations are expected to be led by the HPU with the most cases in that cluster, and any national clusters that are being coordinated by this CI are left to the HPUs to investigate locally with some demographic information about the cases from different HPUs included in the cluster report. Neither local nor regional clusters are coordinated by the CI across the rest of England. The CI for the rest of England had short-term administrative support and is line managed by the TB Lead for the North East.

Training

The aim of the training programme was to enable the implementation of the public health component of the TB-STS by equipping public health and clinical teams to use the strain typing information to inform public health decisions. It involved face-to-face and online resources aimed at those working in HPUs and relevant NHS staff, and was carried out between January 2011 and February 2012. The publication of the Handbook[197] and a Q&A sheet[215] in December 2010 was the first stage of the training strategy for the TB-STS and provided health protection staff with guidance on how to understand, interpret and use strain typing for TB control. The rest of the training consisted of a workshop at each HPU between January 2011 and February 2012 facilitated by a CI, and a seminar and webcast at the national Health Protection Conference in September 2012.

4.4 Quality assurance

Laboratory quality control

There are three levels of laboratory quality control that have been built into the service: internal quality control (IQC), PHE quality assurance (PHEQA), and external quality assurance (EQA). The IQC is on-going throughout the year and the PHEQA and the EQA are spread out across the year so that a panel is analysed at each laboratory every two months.

- IQC. Each reference laboratory re-types and analyses a minimum of 1% of all cultures processed at the laboratory.

- PHEQA. Three times a year the National Mycobacterium Reference Laboratory in London produces and distributes a panel of 8 DNA samples to the reference laboratories. Once returned, the results are collated and shared with the participating laboratories. A teleconference is then set up to resolve any inconsistencies between the data produced across the three laboratories and to suggest remedial actions.

- EQA. The reference laboratories participate in the National External Quality Assessment Service (NEQAS) EQA. Four quality assurance samples are sent three times a year. Where the samples contain MTBC, typing data are submitted and the results are scored.

Public health quality control

The PHE STPB one agreed marker of quality for public health quality control.

- For at least 80% of clusters that cross the threshold for a cluster investigation, a decision about further public health action should be made using strain typing information. This should be limited by geographical boundary. Clusters are categorised as:
    - Probable or definite chain of transmission identified and actions taken to interrupt further transmission;
    - No apparent chain of transmission identified, no further action deemed necessary;
    - Investigation inconclusive;
    - To maintain a watching brief;
    - Investigation continuing; or
    - Other.

4.5 Cost of the TB-STS

Using information from the PHE Finance Division and the RMN, it was possible to estimate the cost of the TB-STS based on annual activity. The fixed costs of establishing and running the TB-STS amount to approximately £1 million per year (Table 9).

Table 9 – Fixed costs of establishing and running the TB-STS

		Per year	
PHE set up			
Evaluation		£15,753	[a]
Software development		£30,739	[a]
HPS consultant project work		£8,050	[a]
London laboratory set-up costs		£63,896	[a]
Total		£118,438	[a]
HPU			
Cluster Investigators (3 posts)		£184,905	
CCDC time (+4.4% x 26 WTE CCDC)		£113,256	[b]
Total		£298,161	
London laboratory	4,297 tests per year		
Consumables	£20.02 per test	£86,026	
Overheads	-	-	
Staff **costs**	£51.42 per test	£220,952	
Total		£306,978	
Birmingham laboratory	1,075 tests per year		
Consumables	£65.46 per test	£70,370	
Overheads	£9.21 per test	£9,901	
Staff **costs**	£33.50 per test	£36,013	
Total		£116,283	
Newcastle laboratory	908 tests per year		
Consumables	£70.61 per test	£64,114	
Overheads	£19.03 per test	£17,279	
Staff **costs**	£55.95 per test	£50,803	
Total		£132,196	
Overall total		**£972,056**	

[a] Cost annuitised over 10 years using discount rate of 3.5% per year.
[b] CCDC workload estimated from the initial and follow-up surveys (Chapter 5)
Data source (unless otherwise stated): communication with PHE Finance Division and the RMN CCDC Consultant in Communicable Diseases

4.6 Discussion of the strengths and limitations of the methods used to evaluate the Structures of the TB-STS

Describing the Structures of the TB-STS helped to identify gaps in the system. For example, the delay in the development and release of the STM had a negative impact on the performance of the TB-STS. The importance of this impact on the evaluation findings are discussed on page 200.

The methods used to describe the Structure of the service meant that changes to the implementation of the service could be identified. For example, the Project Initiation Document indicated that the service had funding for four regional CIs, however two posts were not filled in order to provide a saving to the HPUs. The process of information gathering was important in itself, as it engaged all of the service stakeholders, from whom further data were required for the remaining components of the evaluation.

The cost of the TB-STS cannot be compared to other national services as the costing of other national services has not been published. However, the set-up costs are likely to be determined by the number of reference laboratories being equipped to type, and the number of isolates typed per year.

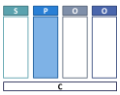
As with any complex intervention, the TB-STS is not a stable system, so presenting an accurate description of the service is difficult. The Structures of the service described in this thesis were correct at the end of the evaluation period, in April 2013, and are representative of the service evaluated.

Summary of findings

The TB-STS is based on prospective typing of *Mycobacterium tuberculosis* using 24 loci MIRU-VNTR. The results are linked to existing national epidemiological and laboratory datasets, and reported in real time to front line teams to inform local TB control. The prospective strain typing and real time reporting of results to inform public health action is what makes the TB-STS a novel intervention. The service is comprised of five components: 24 MIRU-VNTR loci typing and reporting of all first

isolates from TB patients, cluster analysis and reporting from laboratories to public health staff, public health assessment and decision-making around clusters, subsequent cluster investigations, quality assurance (including identifying laboratory contamination), and this prospective evaluation. The overall cost of running the TB-STS is approximately £1 million per year.

Chapter 5. Processes in the TB-STS



In this chapter, I show results from initial and follow-up cross-sectional surveys that I undertook to evaluate the Processes involved in the TB-STS. Specifically, the surveys assessed the implementation of the TB-STS and the knowledge, perceived usefulness and practices of service users. These surveys were designed to capture changes over time between the first and third years of the TB-STS. The survey methods and response are described on page 53.

The findings of the surveys identified areas where more detail needed to be captured on the user experience of the TB-STS. Semi-structured interviews were used to investigate this further and the findings are presented in latter part of this chapter. The interview methods are shown on page 66.

—

5.1 Implementation and perception of the TB-STS: Results of the initial and follow-up cross-sectional surveys

The online initial and follow-up questionnaire survey assessed the knowledge, perceived usefulness and practices of public health staff, physicians and nurses working in TB control in November 2010 and March 2012. The survey methods and response are detailed in the Methods chapter (page 53).

5.1.1 Knowledge

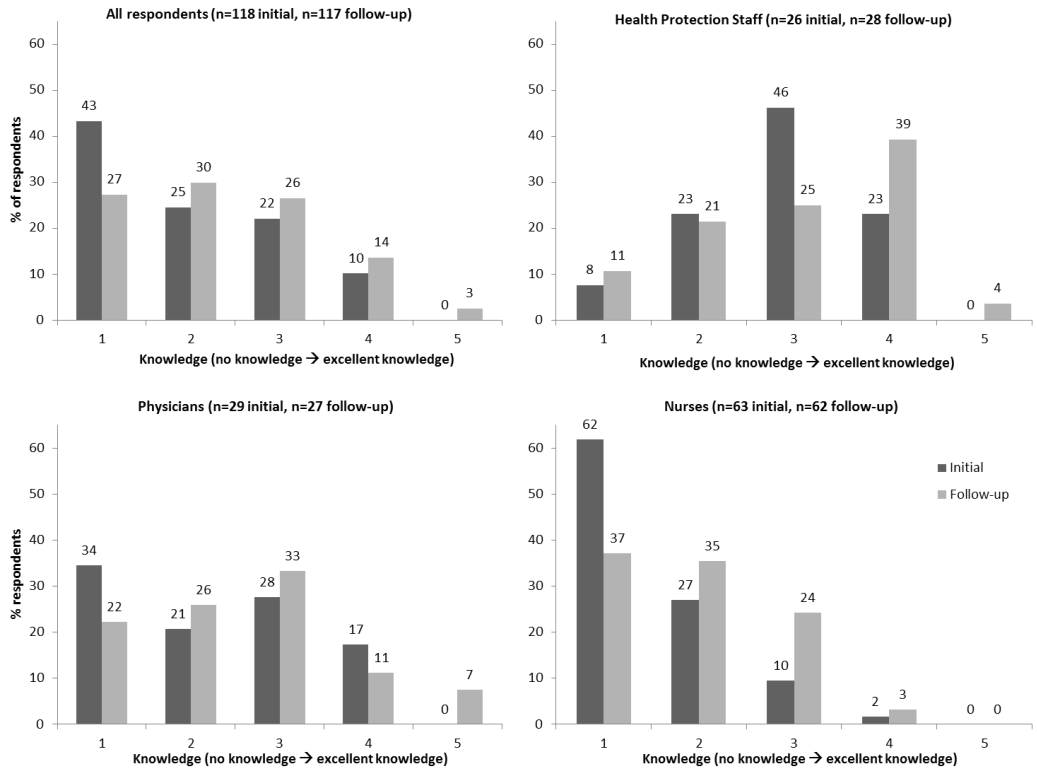
Between the initial and follow-up surveys there were increases in the proportion of respondents who had heard of the TB-STS, had access to strain typing results, had received training, and had access to training resources ($p<0.001$, $p=0.004$, $p=0.003$ and $p=0.032$, respectively; Table 10). The self-rated knowledge of how to use strain typing also increased over time (Figure 32). Nurses reported lower average knowledge in both surveys compared to physicians and health protection staff.

Table 10 – Knowledge: Awareness to the TB-STS and access to strain typing data and resources

		Initial survey		Follow-up survey		
		n	%	n	%	p-value[d]
Heard of the TB-STS[a]	Total	105	85.4	123	99.2	<0.001
Profession	Health protection	28	100	28	100	.
	Physician	20	66.7	30	100	0.001
	Nurse	57	86.4	65	98.5	0.015
TB incidence	Low incidence	49	87.5	56	100	0.006
	Medium incidence	24	72.7	32	97.0	0.010
	High incidence	32	91.4	35	100	0.077
Access to strain typing data[b]	Total	90	72.6	108	87.1	0.004
Profession	Health protection	26	92.9	27	96.4	0.553
	Physician	21	70.0	23	76.7	0.559
	Nurse	43	65.2	58	87.9	0.002
TB incidence[c]	Low incidence	38	67.9	47	83.9	0.047
	Medium incidence	24	72.7	28	84.9	0.228
	High incidence	28	80.0	33	94.3	0.074
Access to training	(health protection staff)	8	28.6	19	67.9	0.003
Access to resources	(health protection staff)	16	57.1	23	82.1	0.042

[a] Have you heard of the TB-STS (apart from in this survey)? *(Yes / No);* [b] Do you have access to strain typing data? *(Yes / No);* [c] HPU-defined area where respondents worked is defined as low, medium and high TB incidence defined as <10/100,000, 10-19/100,000, ≥20/100,000 population, respectively; [d] Chi^2 test of significance comparing responses from the initial and follow-up surveys

Figure 32 – Knowledge: Self-reported knowledge of strain typing



Self-reported knowledge about how to use strain typing was scored on a scale of 1 to 5, where 1 represented 'no knowledge' and 5 represented 'excellent knowledge'. Dark bars represent responses to the initial survey and light bars represent responses to the follow-up survey.

There were significant increases in access to all resources listed: the Handbook, training workshop, colleagues, and webcast/online training (Table 11). The Handbook, written to guide health protection staff (see page 92), was accessed by 75% of health protection staff, and as one would expect given its target audience, by fewer physicians and nurses (7% and 17%, respectively). Similarly, the training workshop was attended by 39% of health protection staff, but fewer clinical staff. The most widely used resource was speaking to colleagues, which was reported by over half of all respondents, and increased significantly between the two surveys in health protection staff and nurses. The webcast/online training session was the least used resource, with fewer than 10% having taken part in one. The 'other' responses remained broadly the same across the surveys and included speaking to microbiologists (initial n=6, follow-up n=2), speaking to colleagues from PHE (n=1, n=3), referring to CDC guidance (n=1, n=1), searching the literature (n=1, n=1), and attending a presentation (n=0, n=1).

Table 11 – Knowledge: Access to resources

	Initial survey		Follow-up survey		
	n	%	n	%	p-value[a]
Handbook	10	8.1	34	27.4	p<0.0001
Health protection	10	35.7	21	75.0	0.003
Physician	0	0.0	2	6.7	0.150
Nurse	0	0.0	11	16.7	0.001
Training workshop	3	2.4	19	15.3	p<0.0001
Health protection	3	10.7	11	39.3	0.014
Physician	0	0.0	1	3.3	0.313
Nurse	0	0.0	7	10.6	0.007
Colleagues	43	34.7	66	53.2	0.003
Health protection	12	42.9	21	75.0	0.014
Physician	15	50.0	18	60.0	0.436
Nurse	16	24.2	27	40.9	0.041
Webcast/online training	0	0.0	11	8.9	0.001
Health protection	0	0.0	5	17.9	0.019
Physician	0	0.0	3	10.0	0.076
Nurse	0	0.0	3	4.5	0.080
Other[b]	10	8.1	8	6.5	

[a] Chi^2 test for significance comparing responses from the initial and follow-up surveys

[b]'Other' included speaking to microbiologists (initial n=6, follow-up n=2), speaking to colleagues from the Health Protection Agency (n=1, n=3), referring to CDC guidance (n=1, n=1), searching the literature (n=1, n=1), attending a presentation (n=0, n=1), and don't know (n=1, n=0).

5.1.2 Perceived usefulness

Respondents' opinion of the usefulness of TB strain typing was high amongst all respondents and did not change significantly between the surveys (95.7% to 94.7%, p=0.667; Table 12). A greater proportion of respondents from low TB incidence areas found strain typing useful at the follow-up survey, compared to those working in high TB incidence areas (97% compared to 89%, respectively), though this result was not statistically significant at the 0.05 level (OR=0.13, 95% CI 0.01-1.13, p=0.075).

Table 12 – Perceived usefulness: Number and proportion of respondents that reported strain typing to be useful[a]

		Initial survey				Follow-up survey				
		Useful		Not useful		Useful		Not useful		
		n	%	n	%	n	%	n	%	p[b]
Total respondents that reported using strain typing		66	95.7	3	4.3	89	94.7	5	5.3	0.667
Profession	Health protection	22	95.7	1	4.3	24	96.0	1	4.0	
	Physician	16	100	0	0.0	20	95.2	1	4.8	
	Nurse	28	93.3	2	6.7	45	93.8	3	6.3	
TB incidence	Low	31	100	0	0.0	38	97.4	1	2.6	
	Medium	16	94.1	1	5.9	26	96.3	1	3.7	
	High	19	90.5	2	9.5	25	89.3	3	10.7	

[a] The following question was asked to respondents who reported that they used strain typing data for TB control (figure 3): Do you find the strain typing information useful? *(Very useful / Quite useful / Not very useful / Useless)* 'Very useful' and 'Quite useful' are grouped into 'useful', and 'Not very useful' is presented as 'Not useful'. No one reported finding the strain typing 'useless' in either survey.
[b] Chi^2 test for significance comparing responses from the initial and follow-up surveys, missing items were excluded.

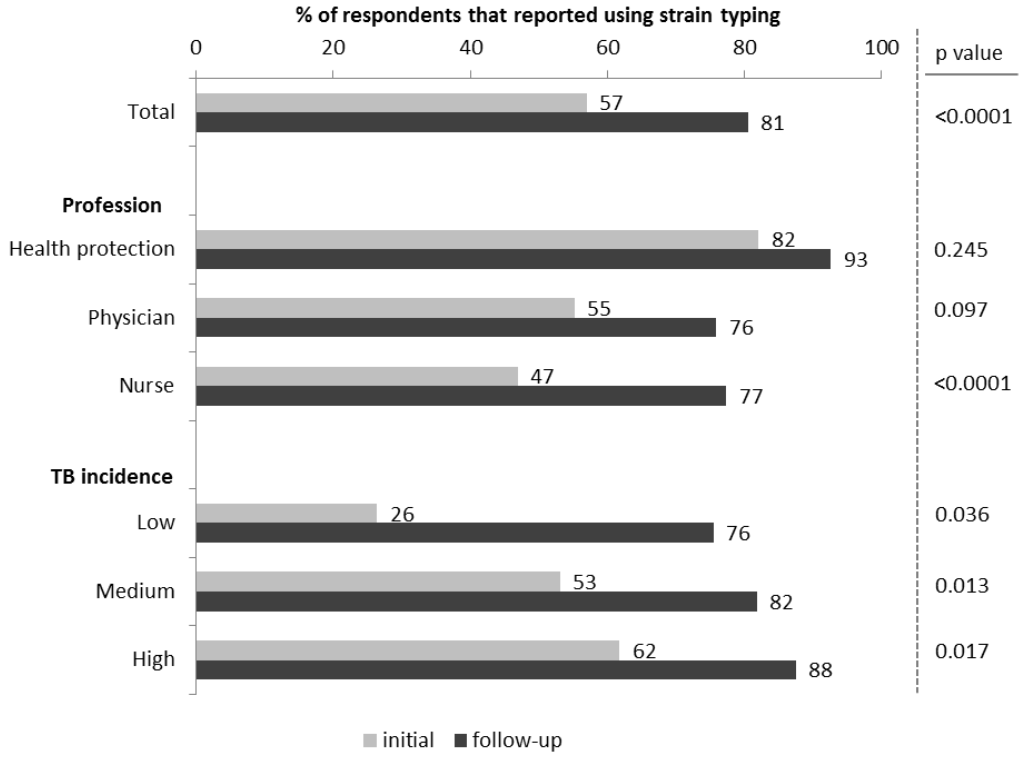
5.1.3 Practices

Nurses and physicians in the follow-up survey reported a mean/median of 7 (range = 2-12) weeks from sample collection to the return of the strain typing result from the laboratory (Table 13). Twenty-four per cent of health protection staff (6/24; missing=3) reported receiving the strain typing data before they completed contact tracing.

There was a significant increase in the number of respondents that reported using strain typing between the initial and follow-up surveys (Figure 33). At the follow-up

survey, the most common way health protection staff accessed strain typing was through CIs and clinic staff through HPU reports (Table 14). There was an increase in the number of respondents who reported using strain typing to: identify links between cases (65.3% to 78.2%, p=0.02; the most common use); disprove links between cases (46.8% to 58.9%, p=0.06); and justify stopping contact tracing (20.2% to 30.6%, p=0.06) (the latter two were not significant at the 0.05 level; Table 15).

Figure 33 – Practices: Respondents that use strain typing for TB control[a]



Dark bars represent responses to the initial survey and light bars represent responses to the follow-up survey.
P-values from Chi^2 tests for significance comparing initial and follow-up proportions are shown.
[a] How often do you use strain typing data in your case management of outbreak investigation? *Never / For few cases / For about half of cases / For many cases / For every case*. 'For few cases', 'for about half of cases', 'for many cases' and 'for every case' were grouped to show the proportion of respondents that use strain typing.

Table 13 – Practices: The average number of weeks it takes for the strain typing result to be returned from the laboratory, as estimated by nurses and physicians at follow-up[a]

	Responses to the question[b]	Mean (SD)	Median (IQR)	Range
Nurse	21	7.2 (2.2)	7 (6-8)	4-12
Physician	14	5.7 (2.5)	6.5 (4-8)	2-10
Total	35	6.6 (2.4)	7 (4-8)	2-12

[a] Nurses and physicians who do not have access to strain typing data excluded from this analysis
Missing items, nurses (n=15) physicians (n=5); [b] 25 responses were not quantified: "many weeks" (n=1 nurse), "variable" (n=3 – 1 physician, 2 nurses), "don't routinely receive" (n=7 – 1 physician, 6 nurses), "have never received" (n=1 nurse), "don't know" (n=13 – 11 nurses, 2 physicians).

Table 14 – Practices: How strain typing data are accessed

	Initial survey		**Follow-up survey**		
	n	%	n	%	p[a]
Request for specific cases	43	34.7	34	27.4	0.217
Health protection	16	57.1	13	46.4	
Physician	14	46.7	12	40.0	
Nurse	13	19.7	9	13.6	
Receive automatically	23	18.5	30	24.2	0.278
Health protection	8	28.6	9	32.1	
Physician	7	23.3	7	23.3	
Nurse	8	12.1	14	21.2	
Contact labs directly	32	25.8	26	21.0	0.368
Health protection	10	35.7	8	28.6	
Physician	7	23.3	12	40.0	
Nurse	15	22.7	6	9.1	
Cluster report from lab	30	24.2	31	25.0	0.883
Health protection	12	42.7	15	53.6	
Physician	5	16.7	5	16.7	
Nurse	13	19.7	11	16.7	
Report from HPU	43	34.7	51	41.1	0.295
Health protection	8	28.6	9	32.1	
Physician	6	20.0	10	33.3	
Nurse	29	43.9	32	48.5	
TB notification system	17	13.7	14	11.3	0.565
Health protection	5	17.9	7	25.0	
Physician	0	0.0	0	0.0	
Nurse	12	18.2	7	10.6	
Cluster investigator[b]	0	0.0	51	41.1	.
Health protection	.	.	18	64.3	
Physician	.	.	8	26.7	
Nurse	.	.	25	37.9	

[a] Chi^2 test for significance comparing proportions between the initial and follow-up survey
[b] The cluster investigators did not start their roles until January 2011, after the initial survey, so this possible answer was only included in the follow-up survey

Table 15 – Practice: How strain typing data are used[a]

Profession	Initial survey		Follow-up survey		
	n	%	n	%	p
Identify clusters and links between cases	81	65.3	97	78.2	0.024
Health protection	22	78.6	25	89.3	
Physician	18	60.0	21	70.0	
Nurse	41	62.1	51	77.3	
Disprove clusters and links between cases	58	46.8	73	58.9	0.056
Health protection	21	75.0	24	85.7	
Physician	13	43.3	15	50.0	
Nurse	24	36.4	34	51.5	
Justify extended contact tracing	51	41.1	60	48.4	0.250
Health protection	16	57.1	19	67.9	
Physician	11	36.7	10	33.3	
Nurse	24	36.4	31	47.0	
Justify stopping contact tracing	25	20.2	38	30.6	0.058
Health protection	13	46.4	13	46.4	
Physician	3	10.0	5	16.7	
Nurse	9	13.6	20	30.3	
To provide more information	34	27.4	44	35.5	0.171
Health protection	13	46.4	10	35.7	
Physician	5	16.7	6	20.0	
Nurse	16	24.2	28	42.4	

[a] What do you use strain typing for? (multiple selections possible) (Don't know / Identify clusters and links between cases / Disprove clusters and links between cases / Justify extended contact tracing / Justify stopping contact tracing / To provide more information / Other (please specify))

Table 16 shows workload reported by nurses and health protection staff. For the nurses, no significant changes in contact tracing workload were reported. Health protection staff reported a significant increase in the mean number of epidemiologically-initiated investigations over a three month period (mean 0.5 to 2.8, p=0.04) and the mean number of these investigations for which strain typing was used to provide more information (0.6 to 1.8, p=0.03), but there was no change in the number that were influenced by the strain typing (1.2 to 0.4, p=0.17). There was no reported difference in the number of clusters investigated because of strain type (though in high incidence areas a large, but non-significant, decrease was reported) and the number of strain typing investigations that identified epidemiological links between cases remained low (Table 16). Overall, the proportion of time health

protection staff spent on cluster investigations increased significantly (from 2.7% to 7.2%, p=0.04).

There was no reported change over time in the frequency at which physicians were called to incident meetings (a meeting, often multi-disciplinary, held to discuss a TB patient, group or cluster of cases that are of particular concern) (p=0.503; most reported once every three months or less (65.5% at in the initial survey and 67.9% at follow-up) and there was no change in the number of physicians who reported strain typing as being relevant to an incident meeting (57.8% to 55.6%, p=0.875)

Table 16 – Practices: The workload associated with the TB-STS for nurses and health protection staff

	Survey[a]	n[b]	median	(IQR)	mean	(SD)	p-value[c]
Nurses							
No. contacts screened in the last month	I	57	21	(11-36)	37.1	(53.5)	
	F	55	20	(8-40)	33.9	(45.1)	0.37
No. hours spent on contact tracing in the last month	I	55	8	(4-16)	12.0	(10.8)	
	F	52	7.5	(3.5-15.5)	16.1	(41.7)	0.24
% time spent on contact tracing	I	57	20	(10-30)	24.2	(16.5)	
	F	54	20	(10-25)	21.7	(17.6)	0.22
Health protection staff							
Investigations initiated because of epidemiological links	I	23	0	(0-1)	0.5	(0.8)	
	F	21	1	(0-2)	2.8	(6.1)	0.04
Strain typing used to provide more information in epidemiological investigation	I	22	0	(0-1)	0.6	(1)	
	F	22	1	(0-2)	1.8	(2.5)	0.03
Strain typing influences an epidemiological investigation	I	23	0	(0-1)	0.8	(1.1)	
	F	14	0.5	(0-2)	1.2	(1.6)	0.17
Investigation initiated because of strain typing	I	23	0	(0-2)	2.2	(6.3)	
	F	22	0	(0-1)	1.1	(2.3)	0.79
Epidemiological links identified in strain typing cluster	I	22	0	(0-0)	0.4	(0.8)	
	F	20	0	(0-0)	0.4	(0.8)	0.52
% time spent on investigations	I	23	1	(0-5)	2.7	(3.2)	
	F	25	5	(0-5)	7.2	(11.1)	0.04

[a] Initial (I) and follow-up (F) surveys
[b] n is numbers of people who answered the question
[c] Paired sample t-test comparing initial and follow-up responses

5.2 User experience: Results of the semi-structured interviews

Free-text comments from the follow-up survey showed that whilst the survey was useful in assessing the implementation of the TB-STS, it did not give much insight into user experiences of the TB-STS. Semi-structured interviews were carried out with the main TB strain typing contact at each HPU. The aim of the interviews was to explore the different processes that are employed around the country to process, disseminate and act upon strain typing information. The methods for these interviews can be found on page 66.

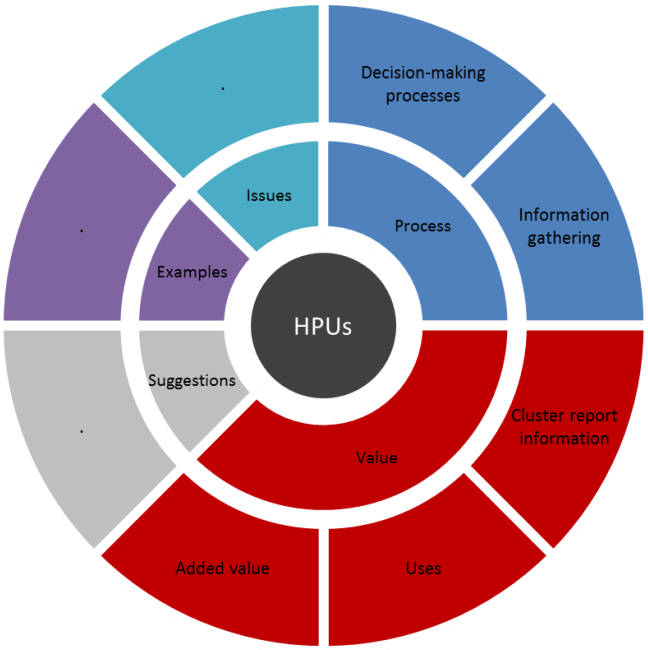
All 26 HPUs were approached for this study and all of them accepted the interview request (two of which were used to pilot the interviews). Five were high TB incidence HPUs, seven were medium incidence and the remaining 14 were low incidence HPUs.

The macro themes that were drawn out of the interviews were:

- The decision-making and information gathering processes involved in the TB-STS (process);
- The value of the TB-STS (value);
- Examples of how the TB-STS had been used (examples);
- Issues relating to the TB-STS; and
- Suggestions for how the TB-STS could be better implemented or used (suggestions).

Figure 34 shows the breakdown of the interviews into themes and sub-themes. These themes are now presented in turn.

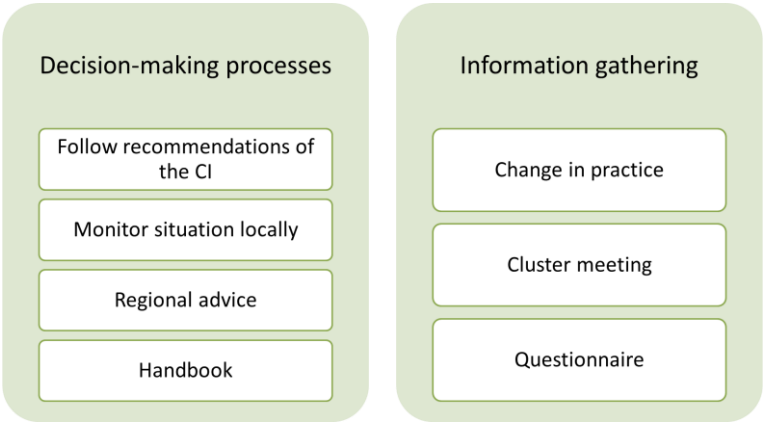
Figure 34 – The themes and sub-themes identified in the semi-structured interviews with health protection staff



5.2.1 Variation in the decision-making and information gathering processes across HPUs

The sub-themes of the processes in the TB-STS that emerged from the interviews are summarised in Figure 35.

Figure 35 – Sub-themes of 'Processes'



Decision-making processes

The interviews identified variation between the HPUs in the way the strain typing information is reviewed and stored, and the gathering of information that constitutes a cluster investigation.

In some areas, the strain typing information was reviewed thoroughly before decisions were made about cluster investigations. This process involved the CCDC lead, the Handbook, advice from colleagues (often including a TB nurse) and the advice of the cluster investigator (CI), and a handful of HPUs had monthly multi-disciplinary meetings where they discussed clusters. In contrast to this holistic approach, there were many HPUs where they relied more thoroughly on the advice of the CI. One HPU reported the involvement of the regional lead, who would review the clusters and give them guidance on the next steps.

> "Use professional judgement, Handbook and the advice of the Cluster Investigator and other colleagues." (CCDC, Low TB incidence)

> "We follow the recommendation on the report." (CCDC, Low TB incidence)

> "We rely on a monthly meeting…we look at cases and clusters and see what we should prioritise." (CCDC, Low TB incidence)

Because many HPUs followed the advice of the CIs, the thresholds for investigation were less relevant to them as they were not taking decisions about whether to investigate themselves. One respondent felt the thresholds, in terms of the number of cases, were appropriate but that the risk factors for early investigation were not. Others found that the thresholds were sometimes appropriate and sometimes inappropriate. Specifically, many low TB incidence areas reported investigating clusters under the threshold, whereas the high TB incidence areas followed the thresholds more closely. Similarly, some people found the Handbook a useful reference tool, whereas others did not use it at all as they would just follow the advice of the CI.

> "I don't know if we've paid much attention to the thresholds. They haven't necessarily impacted on what we've done because two [cases] in certain circumstances might be enough to trigger some action whereas if a cluster of four turns to five it might not be a difference." (CCDC, Low TB incidence)

> "We have the luxury to spend more time to have a quick look if they are potentially linked. We have a lower margin to look at it. We work it [the threshold] as appropriate in our area." (Senior Nurse Specialist and Regional TB Lead, Low TB incidence)

> "Strict about thresholds and liaise with [the CI]…unless we've already picked up some links ourselves through incident management." (Nurse Specialist, High TB incidence)

> "Actually we wait for [the CI]. I don't think we have used the guidelines as such." (CCDC, Medium TB incidence)

Some of the HPUs that were more involved in the decision-making process had also developed a local cluster monitoring database. In these areas, the HPUs mirrored the role of the cluster investigator and carried out a preliminary cluster review by linking the cluster information to information held locally and on the ETS. This was not a widespread activity and tended to be only in low TB incidence areas.

> "Even before we receive the report…we keep an eye on the VNTR and within the unit we create a spreadsheet…and then on an ad hoc basis we can look at it." (CCDC, Low TB incidence HPU)

> "[The London laboratory] sends out monthly spreadsheets. I look to see what is going on in the area. If there is a new cluster I will look at it, even if it is less than the Handbook threshold. I investigate connections between people using ETS, HPZone, ask TB nurses. I organise a meeting to see if we need to take public health action. The cluster report is generated…and I go through the same process." (Health Protection Nurse and TB Lead, Low TB incidence HPU)

Information gathering

There were four main routes of information gathering and TB nurses were critical to them all:

1. Direct contact between the HPU and patients (sometimes through TB nurses);

2. Delegating the cluster investigation to TB nurses and asking them to liaise between themselves;

3. Through multidisciplinary cluster meetings; and

4. By changing contact tracing practices.

The most significant of these was the changing contact tracing practices so that the information that may eventually be required for a cluster investigation was collected when the patient was first diagnosed. Not many HPUs have attempted this change, but many see the value in gathering the information at the earliest possible moment. This change in practice came about in response to nurses being asked to complete cluster questionnaires after their patients had finished treatment, making it difficult to contact them and/or the contact being traumatic for the patient.

> "Some of the TB nurse teams have already incorporated the content of the questionnaire into their normal screening process." (CCDC and Regional TB Lead, Medium TB incidence)

The HPUs that took decisions at monthly meetings would gather some of the information at the meeting because the TB nurses, physicians, microbiologist and HPU are all present to fill in any information gaps.

> "I circulate it to all my TB nurses so they all have a copy and I forward on to others as necessary. I ask them for feedback if there is anything pertinent to tell me, otherwise every other month we go through each report cluster by

> cluster and if there is anything significant we will report it back." (CCDC, Medium TB incidence)

Where monthly meetings did not take place, communication with the TB nurses appeared to be either active or passive. Some HPUs spoke with nurses directly on the phone or arranged to go and see them to gather the information, whereas others "just pass it on to the TB nurses…and ask the TB nurse to liaise with the other areas". There was a distinction between the willingness of HPUs and nurses to collect the information based on whether the TB patient was still on the books or whether they had finished their treatment and had been discharged. A strategy adopted by one HPU was to "email the nurse if the case if from a while ago and ask them to look at their notes to see if anything relevant sticks out…if the case is current I'll ring straight up".

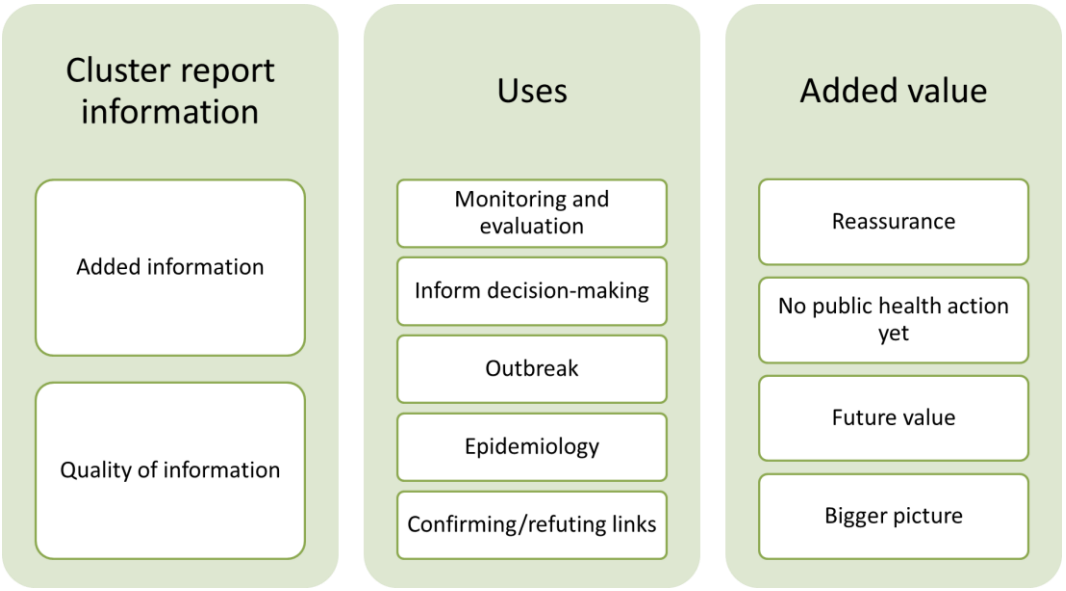
> "We have done quite a bit of work getting a standardised form for TB nurses to use so that they are collecting person, place and time information up front. They've adapted the questionnaire but I cannot give you assurance that it is being implemented properly. In my area the service doesn't see it as a priority. If the patients are still on the books and we think there is a cluster growing we'll ask them for details". (CCDC, Low TB incidence)

The cluster questionnaire (used as a template to gather more information as part of a cluster investigation) was used in some areas and not others. It was deemed useful because it provided a framework for gathering more information and some areas had developed their own local questionnaire based on the cluster questionnaire provided. However, some people talked about the questionnaires being "sent out of the blue", without any further communication or explanation so some nurses "didn't know what to do with them".

5.2.2 The value of the TB-STS

The sub-themes of the value of the TB-STS that emerged from the interviews are shown in Figure 36.

Figure 36 – Sub-themes of 'Value'



The value of the cluster report and cluster investigator

The cluster reports were valued by some HPUs because they lightened their workload by providing them with the summary information that they needed to know and recommending actions for certain clusters. This means that the HPUs did not have to look too closely at the strain typing reports themselves or monitor the strain typing locally. This was valued by HPUs, especially in low TB incidence areas where TB was not a high priority.

> "[Cluster reports] recommended local cluster investigations which is very helpful because when you're faced with growing numbers of cases and clusters it is useful to know where to focus your attention." (CCDC, Low TB incidence)

> "[The CI] sends cluster reports…saves us the effort of going through the [local] database." (CCDC, Medium TB incidence)

The additional information held on the cluster reports can lead to more targeted investigations by the HPU. It was reported particularly helpful in situations where it was difficult to get information from patients as it provided nurses with specific lines of inquiry, rather than more general questioning that elicits an equally vague

response. On the other hand, one HPU argued that the information stored on the ETS was not detailed enough to make a recommendation about whether the cluster should be investigated, and that the decisions about clusters should be led locally where the local intelligence is held.

> "We found it very useful because I didn't know the other cases associated with [the cluster]…so when we get these additional pieces of information it helps the TB nurses go back and ask specific questions - especially where there are language barriers and problems engaging with the patients." (CCDC, Low TB incidence)

> "Sometimes the summary is useful as we are armed with a few extra questions or directions." (CCDC, High TB incidence)

Some HPUs were already aware of the cases in a cluster and their epidemiological links so the reports did not add any additional information and, therefore, did not lead to any public health action. For some it was seen as reassuring when the strain typing information confirmed what they already knew, whereas others felt that the cluster reports were not valuable. This may be explained by the timeliness that cases appear in cluster reports.

One of the main criticisms of the cluster reports was that they were not timely. Many HPUs found that by the time the cluster report arrived they had already completed their local investigation and contact screening activities, leaving no opportunity for the cluster information to influence public health activity. As a result the cluster report did not add anything to local TB control. In fact, some HPUs said that they referred to the laboratories directly for individual strain typing information where they suspected transmission because it was much quicker than waiting for the cluster reports.

> "We do more action even before we receive the cluster report. It's not too late, but definitely not timely to take public health action. We could do with

> a bit more real-time reporting…then we could do away with our local register." (CCDC, Low TB incidence)

> "By the time we get the report we already know about it and have carried out actions. It's always too late… [the service] is not generating new information apart from the typing." (Senior Health Protection Specialist, Medium TB incidence)

> "If we have a feeling about cases then that would be very much in real time so you'd then request it and it would be on diagnosis, rather than a year or two down the line. We would do that anyway. We quite often request the VNTR for a patient and wouldn't necessarily check the cluster information because it is easier and quicker to ask for the VNTR [from the laboratories]." (Nurse Specialist, High TB incidence)

Another criticism of the cluster reports was that they often contained erroneous information or cases that were not from that HPU. This created additional, often unnecessary, work for the HPU and the TB nurses which wasted time, caused confusion, and people lost their confidence in the data. This problem was mainly associated with clusters that were both local and national as the information sharing systems between the local and national clusters was not very robust.

> "It is a muddle when it is a national cluster - sometimes we are asked about patients that aren't even in our area. That has happened a couple of times. Not sure how robust the system is that someone in another part of the country will be identified as part of a cluster - don't feel very confident about that." (Lead Practitioner Health Protection Nurse, Low TB incidence)

In general, however, HPUs enjoyed being able to see how the TB in their area fits into the wider picture and appreciate learning about regional and national clusters. The cluster investigators provided an important advisory role to the HPUs who struggled with the increasing workload and helped them to prioritise the clusters where an investigation may have led to public health action.

> "I value the opinions and guidance and expertise and knowledge about - they know what is going on and what people are doing in other areas. It is useful to have their advice…draw from other experiences." (Health Protection Nurse and TB Lead, Low TB incidence)

How the strain typing is used

Strain typing was mainly used to confirm or refute suspected transmission. Where suspected transmission was confirmed with strain typing it could be used to justify the time and resources then spent on managing complex TB cases, as well as justifying the actions taken. The ability to rule out suspected transmission changed the way that some HPUs managed their incidents, and the strain typing had become part of their decision-making process. However, where transmission was suspected in a particularly high risk setting, HPUs would request the strain typing directly from the laboratory so that transmission could be confirmed or ruled out quickly. In these circumstances the role of the cluster reports was diminished.

> "Useful both in highlighting potentially linked cases we may not have thought of and disproving cases that might have been linked. There is a definite value in that." (Senior Nurse Specialist and Regional TB Lead, Low TB incidence)

> "Most of the benefits from strain typing has been where we started off before all of this - to confirm or refute epidemiological links." (CCDC, Medium TB incidence)

Only a few HPUs reported using strain typing to identify potential new links between cases. This was the more conventional activity around a cluster investigation, requiring more time and resources and often resulted in a lot of effort but no newly identified links.

> "There is implications for it because it tells us to throw the net wider or to stop. Sometimes it makes the work more, and no real outcome, but it does help – another reassurance mechanism. Gives us a justification for what we're doing." (CCDC, High TB incidence)

The strain typing was reported to be particularly useful for on-going outbreaks as it helped to define a case in an outbreak (where the case is culture positive) so one could learn early on about new cases and this could influence the contact tracing strategy and allow for targeted contact tracing as early as possible. Where cases were reluctant to disclose their contacts, strain typing could help to identify relationships with other cases and helped the nurses tailor their interviews with patients.

Some HPUs found the strain typing useful for monitoring and evaluation. Where additional cases were identified with the same strain type and they were linked epidemiologically to another case in the cluster it suggested that the contact tracing was not as thorough as it should have been, or where no more cases were added to the strain typing cluster one could be certain that the contact tracing was carried out well. One HPU has even used the strain typing to argue for more resources for TB so that contact tracing could be carried out more thoroughly.

Added value

The ability to see the bigger picture of TB was identified as an added value of the TB-STS. HPUs were focussed on their own geographical area and the strain typing provided information about how they fit into the national context. However, according to the HPU staff interviewed from areas of low, medium and high TB incidence, in terms of public health output the TB-STS had not yet delivered. Many people, especially in high TB incidence areas, felt that the TB-STS generated more work that led to little benefit.

> "Hasn't added benefits so far, but hasn't highlighted anything we weren't already dealing with...Very happy to receive the strain typing - wouldn't want to not receive it." (CCDC, Medium TB incidence)

> "My gut feeling is that they are perhaps not the most efficient use of scarce resources. I haven't seen anything arising out of these investigations that's actually made a difference in terms of public health management, and yet

there is a lot of traffic of questionnaires being distributed from Colindale to TB nurses." (CCDC, Medium TB incidence)

"Haven't done anything so far that has led to us being able to target wider screening or to feel like we've taken a public health action that wouldn't already have been taken. We haven't delivered anything on the back of it." (CCDC, High TB incidence)

The TB-STS added value in other more general ways that strengthened the TB service as a whole. It helped to engage TB nurses with the public health aspects of TB and developed their relationship with HPUs. On the other hand, the increased demand placed on TB nurses to gather additional information and complete more forms put pressure on this relationship in some areas. This pressure, however, resulted in TB nurses in some areas collecting the data in the cluster questionnaires prospectively for each case, rather than retrospectively and this was recognised as a main benefit of the service so far.

"It's a good thing. It's been helpful for us. It's a good way of getting people more interested in more work around TB. Doing cluster investigations has been a good way also of building up relations with the nurses. They now ring us to ask about strain typing. It has improved dialogue." (CCDC and Regional TB Lead, Medium TB incidence)

Many HPUs saw the potential future value of the TB-STS and acknowledged that the service was still new and that, as more data are collected and analysed, more could be learned from it.

Despite the lack of outcomes and drain on resources, people felt that it would be regressive to stop strain typing.

"From a clinical viewpoint the effects are marginal. Would I keep the cluster work that's being done? Yes, I probably would. Would I fund it out of my budget? Probably not." (CCDC, Medium TB incidence)

> "The general strain typing is like a fishing expedition." (CCDC, Low TB incidence)

5.2.3 Examples of cluster investigations

Respondents from the HPUs provided examples of many types of cluster investigations: where no new epidemiological links were found, where strain typing confirmed or refuted suspected transmission and where strain typing identified new cases and new links between cases or new settings for transmission and where public health actions were required (Table 17).

Table 17 – Examples of cluster investigations and the influence of strain typing

Use of strain typing	Example
No public health outcome	
Low incidence	"Last week a hospital microbiologist rang and said there were two cases of TB with the same VNTR number. Two Philippinos…Do they know each other from another setting? You dig for more information and find out they are absolutely unconnected to each other. Someone should tell us that it is the predominant strain in the Philippines. Later I find out that one in five people in the Philippines have the same VNTR. If I know that this is a common strain then it will help us to say there is no point in throwing in more resources to investigate further. Otherwise we'll be doing more and more and we'll end up doing nothing!"
Medium incidence	"We had a cluster about a year ago of 8 patients in predominantly white UK born with socially complex or chaotic lifestyles. That's as specific as we got I terms of epidemiological links and we couldn't establish, despite all the work, teleconferences, etc....most of these people are chaotic and that's it!"
High incidence	"Regional cluster going on now. Some of the cases were recently discussed at the cohort review so have had local discussion and identified some links between the cases in SW London. No questionnaires were returned. Why? Cases were from one or two years ago - clinics probably ignore they are not currently seeing the person and night hot have time to pull out notes. Once you've gone back beyond a year it is hard to ask people to do it."
Prove/refute suspected transmission	
Low incidence	"Case of TB in a homeless person who was squatting in a garage. Then another case of someone who squatted in the same garage within six months, but they had totally different strains. We were about to chase lots of homeless people around!"
Medium incidence	"One particular family in the last 6 years they've had 11 cases of TB, 7 or which were culture confirmed. I didn't know about them because they had different surnames and different addresses and didn't have the same strain type. The TB nurse asked me about them and she was worried that they weren't managing the cases properly. But the different strain types showed that there was no transmission they could have interrupted."
High incidence	"There were a few cases in prison. Once we started talking to the case they said they knew where they got it from, but we checked the typing and it was different. We heard of another prison in the country who were doing screening and found that they [our case and their case] had links...all had different VNTRs."
New link / case / setting / identified outbreak / public health action	
Low incidence	"We've just had a second case with strain typing in a factory that we screened two years ago because there was a guy who was very infectious and was undetected for six months. We did intensive screening and didn't find anyone else [at the time]. Now we have someone with the same strain type so obviously there was transmission [in the factory] so we'll get the TB service to interview a bit more and raise awareness and keep an eye on the place. That's useful. We couldn't get that information from contact screening at the time."
Medium incidence	"We had a cluster in the Polish homeless people. There were 13 contacts and three of them were cases within six months and a fourth is in hospital. With each case we identify the contacts and try to get in touch with all of them. But the second case comes up and it wasn't identified in the initial contact tracing. [The strain typing] tells us are we finding it out all the contacts."
High incidence	"Put quite a bit of work into a national cluster – alcoholics. Had incident meetings, went carefully through the questionnaires: was there a common context where we could focus some action? We couldn't identify one thing that was shared between them, but there were individual links between some cases. In common was unemployed and alcohol. We did take public health interventions by writing to all GPs in [the area] to make them aware of the cluster, raising awareness with local GPs about the on-going outbreak in that group…and alcohol awareness stuff, awareness raising too. They were general initiatives based on what we knew about the cluster."

5.2.4 Issues

HPUs identified many issues and problems with the TB-STS. The sub-themes of the issues raised during the interviews were:

- Communication
- Process
- Regional and national clusters (compared to local clusters)
- Resources

The time between the date the case was diagnosed and treated and when they appeared on the cluster report was identified as an issue. It was noted that you could be waiting up to two months before a case was included on a cluster report as the reports were sent out every month. This was a problem because if the information was not reported in real-time it could not elicit a real-time response. In addition, due to the time it may take for the investigation threshold to be reached, cases may be included in a cluster report one or two years after they had been discharged. This made the additional information gathering difficult. This was recognised as an important barrier to data collection.

> "No questionnaires were returned. Why? Cases were from one or two years ago - clinics probably ignore they are not currently seeing the person and night hot have tie to pull out notes. Once you've gone back beyond a year it is hard to ask people to do it. If cases are in the last 12 months then it should be easy to pull out notes or remember - if it's real to them, then it's more likely." (CCDC, High TB incidence)

Communication problems within the TB-STS were identified. One HPU claimed that "until a couple of weeks ago I did not know that there was a person who had 'cluster investigator' in their job title!" and others were struggling to fully understand their role and what was expected of them. People had different experiences when investigating usual TB incidents compared to cluster investigations.

> "The incidents that we deal with, we get to know them quite well because we work closely with the clinics, speak with the patients and liaise with

> where it is situated. Whereas the clusters feel much further removed. There is no reason why the HPU would be involved. It feels more of an academic exercise in some ways." (Nurse Specialist, High TB incidence)

The main issue identified by all the HPUs was a lack of resources. HPUs described the lack of resources allocated to TB services in general and questioned whether the TB-STS was a good way to spend money. There was a strong awareness that the HPUs were not given any additional resources to investigate clusters and that TB nurses and physicians working in the NHS were not allocated any additional resources to collect the extra information. This opinion was held across low, medium and high TB incidence areas and some HPUs were concerned that the increasing data requests would put pressure on their relationships with the TB services.

> "I have some concern on the questionnaire side and the full cluster investigation - in the good will of our NHS partners and how much they contribute to that. Most of them have been quite enthusiastic because it is new, sexy and something to learn, but can see them becoming less wilful. There are mutterings about not being commissioned to do them [the questionnaires]". (Senior Nurse Specialist and Regional TB Lead, Low TB incidence)

Some questions and queries were raised by the HPUs. There were many questions surrounding the interpretation of strain typing. People weren't clear on what it meant if someone in their HPU was clustered with someone from a different geographical area, how missing loci should be interpreted, and when a cluster did not represent recent transmission but represented a strain endemic to a certain population. There were also questions about the cluster investigation activities, such as how many times patients should be approached for information and how often you should contact them if you have new information, how the results of a cluster investigation will affect case management, how to investigate without breaking patient confidentiality, how to find links, what should lead to public health action and how to get feedback on national clusters.

5.2.5 Suggestions

The categories of suggestions that were discussed in the interviews were:

- Efficiency and best use of data
- Feedback, communication and learning
- WGS

Efficiency and best use of the data

HPUs suggested that reducing the number of clusters requiring additional work would help to improve people's experience of the service. They suggested that automating the systems and raising the thresholds would help by allowing people to focus on the important clusters, rather than all clusters that appear in their area. In addition, some HPUs suggested that a dedicated team to investigate strain typing clusters could lead to better public health outcomes. One HPU suggested that local engagement could be improved by focussing on the local benefits to clinical and public health services.

> "I think the cluster typing has to inform and strengthen the clinical services because then once supporting the delivery of the clinical and public health programme locally, informing it, you get people engaged, understanding clusters, using them, it becomes one of the daily activities. Real time, as they're developing and using the information directly with services then it becomes relevant to them, rather than just flood people with masses of retrospective clusters and risk factors and patient risk factors." (CCDC, Medium TB incidence)

There was a call for more epidemiological support. Many HPUs had developed a local database of strain types but did not have the skills to analyse it and extract useful information from the data.

Feedback, learning and communication

There was a strong sense that communication, feedback and the ability of the TB-STS to learn from its outputs was not adequate. Concerning communication, HPUs felt there needed to be better data sharing between the local cluster investigators and

the national cluster team as it slowed the service down when they only had access to part of the cluster information. Some HPUs also felt that it wasn't clear what they are expected to do with the strain typing information.

Whole genome sequencing (WGS)

WGS was suggested as a method that would split MIRU-VNTR clusters up, thereby reducing the number of clusters that would require investigation. Concerns, however, were raised that the MIRU-VNTR TB-STS should be properly implemented and integrated into the TB service before introducing a new system.

> "It is important to get our head around how we integrate epidemiological and genomic data. For most people on the front line most people are not at that point and if we miss this opportunity most people will struggle with whole genome sequencing. I certainly think that it potentially has value but I'm not sure we've done it [the MIRU-VNTR strain typing service] justice." (CCDC, Medium TB incidence)

5.3 Discussion of the Processes in the TB-STS

5.3.1 Strengths and limitations of the methods to evaluate the Processes

An initial and follow-up cross-sectional survey was chosen because it is a study design that could be feasibly conducted, at low cost, in the time-frame available. Furthermore, electronic surveys were easily disseminated as everybody in the target group had PHE or NHS email addresses. Additionally, as no information was available prior to the first survey it was a particularly appealing method to capture data about the initial phase of the TB-STS; it was possible to design, pilot and disseminate the survey quickly. It was also easy to repeat the survey at the second time point to assess any changes that may have taken place between two periods.

However, the study has a number of limitations, which are discussed below.

Firstly, the survey was developed after the initiation of the TB-STS so baseline information could not be collected. As a result, we may have underestimated the

difference between the surveys. However, the initial survey was conducted before the roll-out of any training for the TB-STS and prior to the employment of all national staff to coordinate cluster investigations and therefore it seems unlikely that there would have been any major changes prior to the first survey.

Secondly, the target population for the survey consisted of all public health staff, physicians and nurses working in TB control in England. It was not possible to enumerate the sampling frame because no register of clinical and health protection staff working in TB could be identified making it impossible to calculate a response rate and any associated potential biases. Response bias may have occurred if only those with a vested interest in the TB-STS (either positive or negative) responded to the survey; however, given that there was neither overwhelming support nor objection to the TB-STS, this is unlikely to have occurred or these responses may have cancelled each other out. In addition, the survey may have over-represented those with a smaller workload (as people with a greater workload may not have had time to respond to the survey); there were more responses from low TB incidence areas (45%, 27% and 28% from low, medium and high TB incidence areas) but this reflects the distribution of TB across England.

Thirdly, the 50% retention rate between the surveys is quite low which could bias the results if there was a systematic difference between responders and non-responders. It is possible that the opinions and experiences of a particular group of people were omitted. However, the validity of our results is supported by the fact that non-responders to the follow-up survey did not differ significantly compared to those that responded to both surveys based on profession or burden of TB in their geographical area.

Finally, the survey was not powered to look for differences between professions or geographical areas and when testing multiple hypotheses one might find a significant result due to chance. Therefore, it is especially important to consider the findings of the surveys in the wider context of the evaluation and the other methods used, and interpret them accordingly.

Interviews were used to explore the TB-STS user experience. This was appropriate given that the findings were used in a health services research context to supplement the quantitative work conducted around the implementation and perception of the TB-STS. They provide useful insights into the use and perceptions of the TB-STS from the perspective of the main service user (health protection staff). This was useful for understanding the broader findings of the evaluation. Semi-structured interviews were chosen as the method of data collection because they provided an opportunity for more in-depth information on issues identified through describing the service and the initial and follow-up survey, to be gathered and further explored. Data from these interviews also inform the interpretation of the thesis findings (see discussion, page 191).

All the people who were approached by the interviewer accepted to participate; the interviews were representative of all HPUs. Interviews potentially result in a number of biases associated with the interviewee, in this case the strain typing leads in the HPUs and their perception of the interviewer. The strain typing leads had very different roles depending on the burden of TB in their HPU which may have influenced the way they approached the interview: for some, strain typing was a large part of their job so they may have been keen to highlight its importance and utility to justify their role; for others, strain typing was a small part of their overall job so they may have understated its usefulness in order to justify not using it, creating more time for other parts of their job.

The role of the interviewer may also be important. The interviewer presented herself as an external evaluation scientist, who had no conflict of interest. However, people may have assumed that the interviewer worked for PHE. This may have led to a social desirability bias whereby the interviewees say what they think the interviewer wants to hear – if the interviewer was perceived to be from PHE, this could have led to exaggerated support for the TB-STS. Finally, the interviews were all conducted over the telephone, increasing the distance between the interviewer and interviewees, and excluding non-verbal cues. Not being able to see who was conducting the interview may have helped interviewees to speak openly, or it could have made them

feel more guarded as there was less opportunity to build a rapport. Given the service delivery nature of the interviews, rather than those covering personal issues, these biases are likely to have minimal impact.

5.3.2 How this relates to other work

National TB strain typing services have been introduced in other countries,[100,133,136,152] but the knowledge, perceived usefulness and practices of users have not been evaluated. However, the impact of strain typing on contact tracing activities in the Netherlands has been assessed.[127] Consistent with that study, no change was observed in the workload associated with the TB-STS for nurses and physicians, even though strain typing was used by more people at the follow-up survey (indicating the roll-out of the service). This may be because it is difficult to measure marginal changes in workload associated with a particular service where the workforce is already working at full capacity. Health protection staff, however, spent a greater proportion of time on cluster investigations. Given that the Handbook had not been published and all the cluster investigation coordinators were not in position at the time of the initial survey, this is not surprising and suggests that the TB-STS had been integrated into the TB control activities of the HPUs. It is not known what the opportunity cost of this was; what were health protection staff spending *less* time on in order to spend more time on cluster investigations?

The median time it takes for the strain type to be received by TB nurses and physicians following a specimen collection was seven (IQR 4-8) weeks. This is shorter compared to that reported for the USA (8.9 (6.1-12.3) weeks from specimen collection to receipt at typing laboratory and 1.4 (0.6-1.3) weeks from receipt at laboratory to reporting the result),[216] but is still too long for real time reporting. The time delay experienced in England may explain why 76% of health protection staff reported that the strain type is received after contact tracing has already finished.

Based on evidence from the USA one would expect more possible transmission links to be identified when strain typing informs contact tracing.[88,97] This was not detected in the survey despite an increase in the time health protection staff spent on cluster

investigations and the number of investigations that used strain typing. This discordance between the findings on subjective report of utility and the public health outcomes reported could be because the current methods used by public health staff to identify epidemiological links may be inappropriate or ineffective, or there may have been an increase in suspected (but not established) transmission because of the strain typing information.

The TB-STS relies on HPUs and TB nurses to work together to implement any public health action in response to on-going transmission in the community. This raises the important issue of whether the TB service has the capacity to act upon the information collected through cluster investigations. For the TB-STS to have a public health impact and reduce TB transmission, cluster investigations would have to lead to the detection of previously unidentified latently infected and active TB.

5.3.3 Summary of findings

Initial and follow-up cross-sectional surveys of TB-STS users were conducted to assess the implementation of the TB-STS. There was an increase in the number of people that used strain typing and knowledge of the TB-STS between the two surveys. A change in perceived usefulness was not evident as the majority of respondents found strain typing useful to them at both time points. Nurses and physicians waited to receive the strain typing for an average of seven weeks after sending the sample to the laboratory and the majority of health protection staff received the results after completing contact tracing. With respect to workload associated with the TB-STS, there was no change over time in the contact tracing activities of nurses or the frequency of incident meetings attended by physicians; however the proportion of time health protection staff spent on investigating TB transmission increased significantly. Despite strain typing being used to provide more information to public health staff at follow-up, there was no increase in epidemiological links identified.

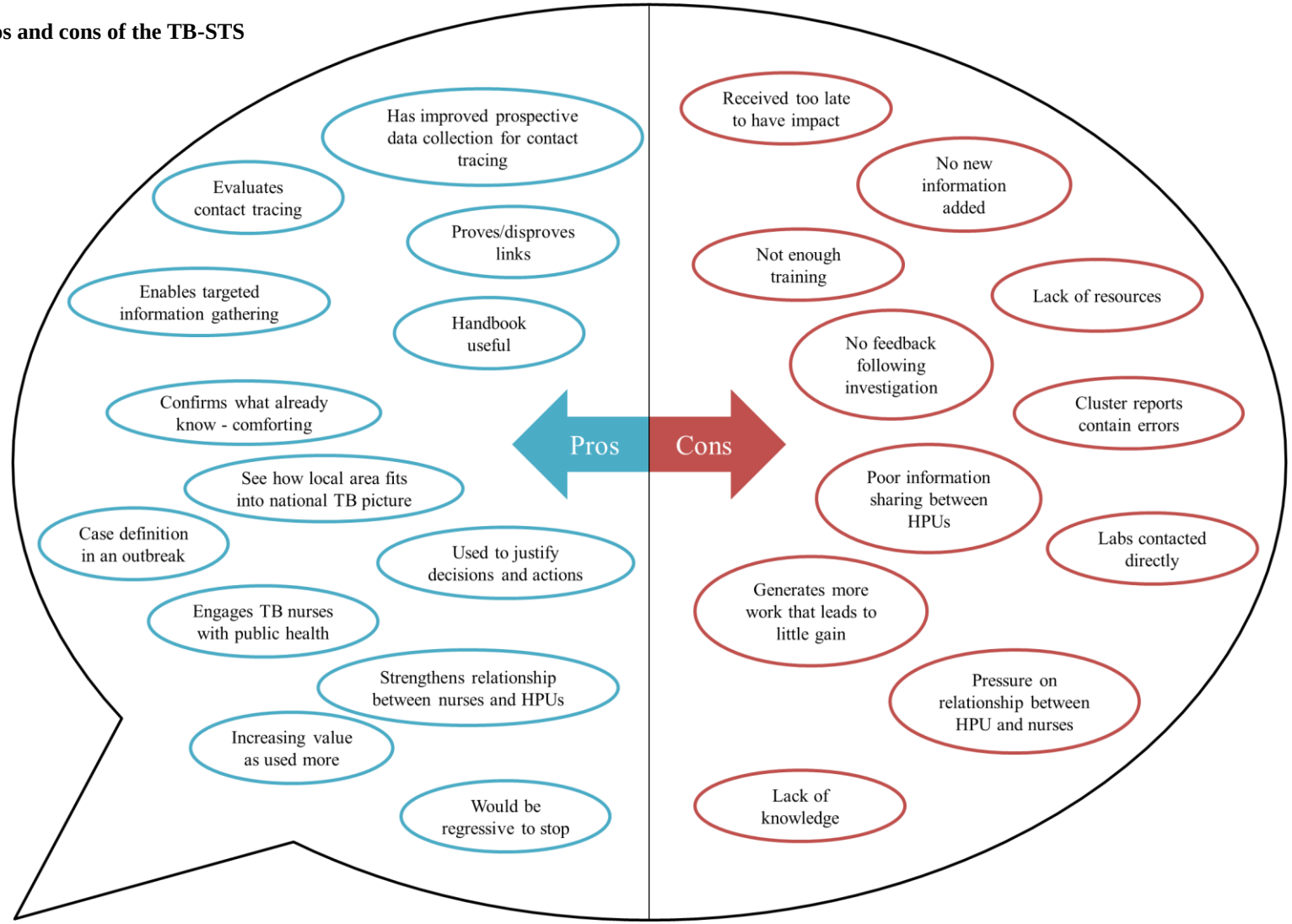
Semi-structured interviews with the strain typing leads in each HPU explore the user experience of the TB-STS. They found that some HPUs reviewed the strain typing

information thoroughly and others just followed the recommendations of the CI. Many low TB incidence areas reported investigating clusters under the threshold, whereas many high TB incidence areas followed the thresholds more closely. The cluster reports were liked by some HPUs because they summarised the strain typing information well; however, some HPUs found that the reports did not add any additional information and therefore did not lead to public health action. Many HPUs felt that, although there were many uses for the strain typing (such as confirming or refuting transmission), the service generated a lot of work but had not yet had a public health impact. Despite this, people felt that it would be regressive to stop the TB-STS.

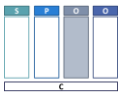
The advantages and disadvantages of the TB-STS identified by HPU strain typing leads are presented in Figure 37 and exemplified by the following quote:

> "Would I keep the cluster work that's being done? Yes, I probably would. Would I fund it out of my budget? Probably not."

Figure 37 –The pros and cons of the TB-STS



Chapter 6. Outputs



In this chapter I present the results of a broad investigation into the Outputs of the TB-STS by presenting the laboratory Outputs, including the detection of false positive TB isolates, followed by the public health Outputs. The false positive TB isolation survey methods can be found on page 69.

Multiple approaches were taken in an attempt to measure the public health Outputs of the service. Firstly, I present the available data on cluster investigations and their outcomes (methods on page 70). Secondly, the results of alternative investigations into the contact tracing yield, diagnostic delay and rate of cluster growth associated with cluster investigations (methods from page 73). The laboratory Outputs and the indirect public health Outputs were fundamental in providing parameters for the transmission and cost-effectiveness models.

—

6.1 Laboratory Outputs

The laboratory Outputs of the service include the number of *M. tuberculosis* isolates typed with 24 MIRU-VNTR and the false positive TB isolates identified by the TB-STS. Table 18 shows the number of isolates typed by the reference laboratories between 2010 and 2012 and the number with at least 22 loci. It shows the variation in workload across the laboratories. The largest proportion (n=11,069, 64.5%) of isolates were typed in the London laboratory, with 3,159 (18.4%) in Birmingham and 2,940 (17.1%) in Newcastle. The total number of isolates typed between 2010 and 2012 was 17,168. Of these, 58.7% had complete 24 loci strain types. Between 2010 and 2012, the proportion of isolates with strain types at least 23 digits long (the minimum requirement suggested in the TB-STS data production guide),[214] was 84.6%.

The total number and proportion of clustered cases, the number of clusters, and estimates of recent transmission in England between 2010 and 2012 are shown in Table 19.

Table 18 – The number and proportion of isolates typed in each of the reference laboratories

Laboratory	Year	Total isolates typed	Complete 24 loci		1 missing locus		At least 23 loci	
			n	%	n	%	n	%
London	2010	3429	1689	49.3	1036	30.2	2725	79.5
	2011	3921	2319	59.1	1117	28.5	3436	87.6
	2012	3719	2239	60.2	1098	29.5	3337	89.7
Birmingham	2010	979	647	66.1	172	17.6	819	83.7
	2011	1092	753	69.0	190	17.4	943	86.4
	2012	1088	796	73.2	165	15.2	961	88.3
Newcastle	2010	995	557	56.0	224	22.5	781	78.5
	2011	969	557	57.5	224	23.1	781	80.6
	2012	976	527	54.0	215	22.0	742	76.0
Total	2010	5403	2893	53.5	1432	26.5	4325	80.0
	2011	5982	3629	60.7	1531	25.6	5160	86.3
	2012	5783	3562	61.6	1478	25.6	5040	87.2
	2010-2012	**17168**	**10084**	**58.7**	**4441**	**25.9**	**14525**	**84.6**

Communication from laboratories January/February 2013

Table 19 – Number of clustered cases and clusters in England and by HPU[a], 2010-2012

Geographical area (country, region and HPU)	Culture confirmed cases	Cases with strain type[b]		No. cases clustered[c]		No. clusters	Estimated proportion of cases due to recent transmission[d]	No. clusters by size		
		n	%	n	%			2-4	5-9	≥10
England[e]	**14248**	**11491**	**80.6**	**5973**	**52.0**	**1371**	**40**	**1084**	**198**	**89**
North of England										
North East	310	223	71.9	46	20.6	17	13	16	1	0
Cumbria and Lancashire	347	238	68.6	79	33.2	22	24	17	4	1
Yorkshire and Humber	1080	819	75.8	228	27.8	65	20	53	9	3
Greater Manchester	890	639	71.8	154	24.1	47	17	41	3	3
Cheshire and Merseyside	218	171	78.4	26	15.2	11	9	11	0	0
Midlands and East of England										
East Midlands	747	620	83.0	197	31.8	63	22	53	9	1
West Midlands	1706	1401	82.1	599	42.7	154	32	129	13	12
Anglia and Essex	466	376	80.7	81	21.5	33	13	32	1	0
South Midlands and Hertfordshire	610	486	79.7	122	25.1	41	17	37	3	1
London	5997	4999	83.4	2286	45.7	593	34	487	69	37
South of England										
Sussex, Surrey and Kent	613	506	82.5	138	27.3	52	17	47	5	0
Thames Valley	516	424	82.2	87	20.5	36	12	35	1	0
Wessex	291	241	82.8	62	25.7	21	17	20	0	1
Devon, Cornwall and Somerset	152	122	80.3	30	24.6	8	18	7	0	1
Avon, Gloucester and Wiltshire	305	226	74.1	55	24.3	20	15	19	1	0

[a]These geographical areas are the recently re-defined HPUs, now known as Health Protection Centres; [b]Culture confirmed cases with a MIRU-VNTR profile with at least 23 complete loci
[c]Clustered cases are clustered within geographical area; [d]Calculated using the n-1 method (number of cases-number of clusters/number of clustered+unique cases); [e]The number of clusters in England is higher than the sum of all HPUs because it includes cluster that span more than one HPU
Source: Tuberculosis in the UK: Annual Report on Tuberculosis Surveillance in the UK, 2013[123]

6.1.1 The TB-STS and false positive TB isolation

One potential benefit of the TB-STS is earlier identification of the false positive TB cases that can result from laboratory contamination. In addition to the avoidance of anxiety for patients and their families, earlier identification of such cases has health and financial implications if treatment is avoided or reduced and the true diagnosis is missed or delayed. The false positive TB isolation survey methods are on page 69.

Between June 2010 and June 2012 11,059 TB isolates were typed at the reference laboratories (Table 20). There were 70 suspected incidents of false positive TB isolation (0·6%), of which 30 (42·9%) were confirmed as false positive, giving a rate of false positive TB isolation in England of 0·3% (30/11,059). Seventeen (56·7%) of the suspected incidents were not known to the source laboratories, and 8 patients were started on unnecessary treatment (Table 20).

These results were used as a parameter in the cost-effectiveness model (page 81).

Table 20 – The number of false positive incidents identified by the reference laboratories and confirmed by the source laboratories between June 2010 and June 2012.

	London		Birmingham		Newcastle		Total	
	n	%	n	%	n	%	n	%
Total isolates typed[1]	7,341		2,023		1,695		11,059	
Total false positive incidents suspected	58	0.79	11	0.54	1	0.06	70	0.63[2]
Incidents confirmed as cross-contamination	26	44.8	3	27.3	1	100	30	42.9[3]
Incidents confirmed as true TB (i.e. not false positive)	25	43.1	4	36.4	0	0.0	29	41.4[3]
Information not known	7	12.1	4	36.4	0	0.0	11	15.7[3]
Incidents known about by source laboratories before contact from reference laboratory	10	17.2	3	27.3	.	.	13	43.3[4]
Number of false positive patients started on treatment	8	13.8	0	0.0	.	.	8	26.7[4]

[1] reported by each laboratory on the data collection form
[2] proportion of all isolates typed
[3] proportion of all incidents suspected
[4] proportion of confirmed cross-contamination incidents

6.2 Public health Outputs

The direct public health Outputs of the TB-STS are the cluster investigations and their outcomes. In addition to this, indirect outcomes of the TB-STS were investigated: the contact tracing yield, diagnostic delay and rate of cluster growth (see the illustration of the aims of the TB-STS, Figure 11, page 31). Information on cluster investigations and their outcomes was designed to be collected, collated and stored using the STM. In the absence of the STM, the following data sources were used: a cluster monitoring database, cluster outcome reporting form, and a database of national clusters (see methods on page 70)

6.2.1 Direct measures: cluster investigations and their outcomes

The cluster monitoring database was developed in order that CIs could record the characteristics and activities surrounding each investigated cluster and the investigation outcomes. Between 2010 and 2012, 188 clusters were recorded in the database and 160 of these included a start date for the investigation (Table 21).

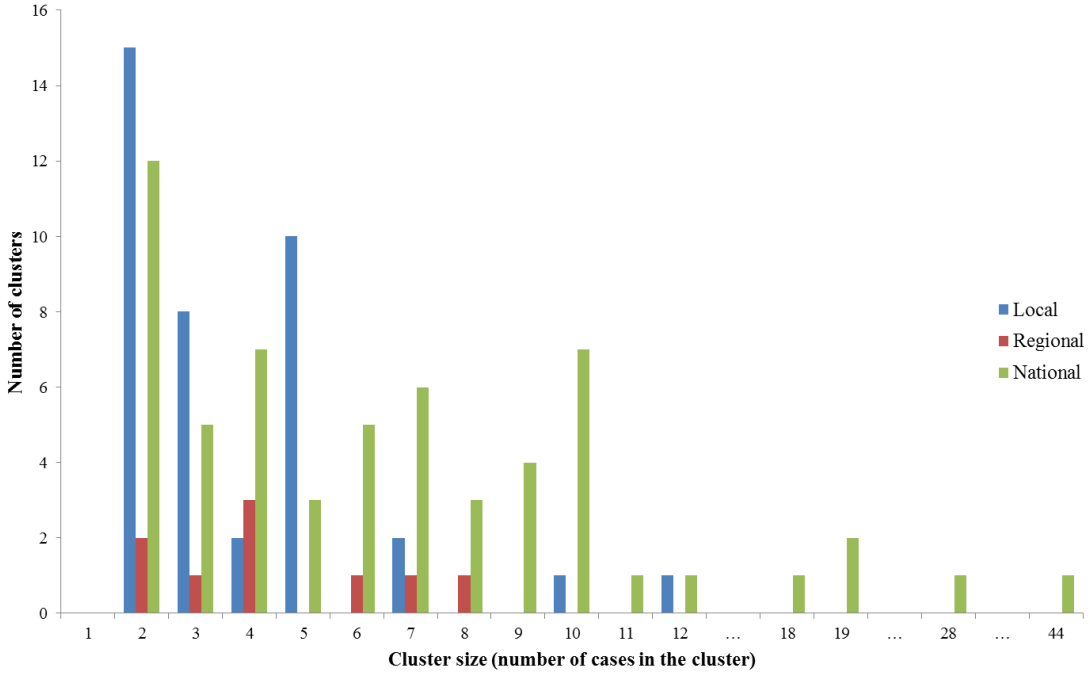
Table 21 – The number of local, regional and national clusters by the year in which a cluster investigation was first launched, as reported in the cluster monitoring database

Cluster type	Year in which cluster first became active				Total in the cluster database
	2010	2011	2012[a]	2010-2012[a]	2010-2012[ab]
Local	2	32	30	64	69
Regional	0	7	6	13	15
National	15	37	31	83	104
Total	**17**	**76**	**67**	**160**	**188**

[a]Jan 2010-Oct 2012
[b]This column is greater than the previous column due to missing information about the date a cluster is first investigated

Cluster size at the point at which an investigation became active ranged from two to 12 for local clusters and two to 44 for national clusters (Figure 38). Figure 38 shows that the threshold of five cases for a local investigation and ten cases for a national investigation is often disregarded, either because the cluster contains cases with risk factors for transmission, and/or HPUs and national teams are choosing to investigate clusters earlier. There are peaks at two cases, three cases and five cases for local investigations and two cases and ten cases for national investigations.

Figure 38 – The distribution of cluster size on the date a cluster investigation is initiated, 2010-2011



The median number of days a cluster investigation remained active (of those investigations that had already been closed at the time of the evaluation, n=86) was 129 days (IQR 50 to 188). The median length of a cluster investigation was longer for national clusters compared to local clusters (171 (IQR 60 to 336) and 97 (43 to 157), respectively).

Outcomes of cluster investigations

Cluster outcome reporting forms collected information on the outcomes of local and regional cluster investigations carried out by HPUs (i.e. the number of additional epidemiological links identified between clustered patients and the number of secondary cases identified as a result of the strain typing information). These forms were not received by the CIs for all clusters that were investigated. Between the publication of the form in the second version of the Handbook in September 2011, and February 2013, and 92 cluster investigations were launched (according to the cluster monitoring database). The content of the forms are summarised in Table 22 and Table 23.

Table 22 – The cluster outcome reporting form: Information about the actions and outcomes of cluster investigations and usefulness of the TB-STS

	Local	Regional	National	Total (%)	Not answered
Number of clusters	26	6	10	42	.
Reason for investigation					
Threshold reached	18	2	6	26 (62)	1
Below threshold, but with risk factors	3	1	0	4 (10)	1
Actions					.
Extended contact tracing	7	0	2	9 (21)	14
Additional contacts found/screened in cluster	5	0	2	7 (17)	35
New epidemiological links found in cluster	15	2	3	20 (48)	3
Public health outcomes [a]					8
Probable/definite transmission	4	1	2	7 (17)	.
Investigation inconclusive	9	2	2	13 (31)	.
No apparent transmission	5	3	4	12 (29)	.
Investigation on-going	2	0	0	2 (5)	.
Was the strain typing useful? [b]					15
Strain typing was useful	11	3	2	16 (38)	.
Strain typing was **not** useful	6	2	3	11 (26)	.

Numbers presented are numbers of clusters. The denominator for the proportions presented is 42.
[a] Only possible to select one option
b "Was the strain typing information useful to you?" *Yes / No*

The total number of cluster outcome reporting forms returned to the local/regional CIs since 2011 was 42 (9 national clusters, 6 regional clusters and 27 local clusters). 26 investigations were launched because the threshold was reached and four were launched below the threshold because of patient risk factors. Nine investigations lead to extended contact investigations and seven investigations screened more contacts, identifying LTBI and active TB cases (Table 23). New epidemiological links were found in 20 clusters. Strain typing was reported to be useful for 16 clusters (and not useful for 11 clusters).

The contact tracing activity and yield from screening additional contacts as part of a cluster investigation was only reported for seven clusters (labelled A to G) (Table 23). Across all seven cluster investigations, 1275 further contacts were screened, 23 (1.8%) were diagnosed with active TB, 125 (9.9%) had LTBI of which 100 (80%) were treated with prophylaxis. It is not known how this differs to the contacts initially screened in these clusters. The final public health decision for clusters is reported in Table 22.

Table 23 – Information on the contact screening activity and outcomes reported for the additional contacts screened

Cluster	Additional contacts screened	Active cases (%)[a]	LTBI (%)[a]	Prophylaxis given (%)[b]
A	118	12 (10.2)	21 (17.8)	21 (100)
B	628	0 (0)	37 (5.9)	37 (100)
C	157	2 (1.3)	7 (4.5)	5 (71.4)
D	81	1 (1.2)	23 (28.4)	.
E	6	0 (0)	0 (0)	0 (0)
F	102	8 (7.8)	24 (23.5)	24 (100)
G	165	0 (0)	13 (7.9)	13 (100)
Total	**1257**	**23 (1.8)**	**125 (9.9)**	**100 (80.0)**

Information presented in this table was only available for seven clusters (labelled A-G)
[a]As a proportion of the additional contacts screened
[b]As a proportion of LTBI
[a and b] These proportions do not include the contacts screened as part of traditional contact tracing that would have been conducted prior to the cluster investigation

More details on a greater number of national clusters were input onto the cluster monitoring database. Table 24 shows the total number (and proportion) of national clusters reported in 2010 and 2011, the number that were recorded as having no epidemiological links between the cases, the number that reported some epidemiological links between cases and the number where the links were found following a cluster investigation. This is broken down by the number of cases in the clusters. 369 national clusters were reported onto the cluster monitoring database in 2010 and 2011. 33 (10.6%) clusters reported to have epidemiological links known between some cases and nine (2.4%) clusters reported finding epidemiological links following an investigation.

Table 24 – Number (%) of national cluster investigations where epidemiological links have been identified, by the size of the clusters

Number (%) of national clusters	Number of cases in cluster					Total
	<5	5 to 9	10 to 19	20 to 49	50+	
National clusters	260	63	31	11	4	369
Clusters **without** reported epidemiological links	254 (97.7)	54 (85.7)	16 (51.6)	4 (36.4)	1 (25)	329 (89.2)
Clusters with reported epidemiological links	6 (2.3)	9 (14.3)	13 (41.9)	8 (72.7)	3 (75)	39 (10.6)
Epidemiological links established following a cluster investigation	1 (0.4)	2 (3.2)	2 (6.5)	3 (27.3)	1 (25)	9 (2.4)

6.2.2 Indirect measures of the impact of the TB-STS on TB control: Contact tracing yield

It was hypothesised that the TB-STS would help to identify more cases of active disease and LTBI through better targeted or extended contact tracing. Contact tracing activity was examined using two data sources: information on index cases reported by nurses in the initial and follow-up survey, and clinic data from NCL sector and Leicester. The methods can be found on page 73.

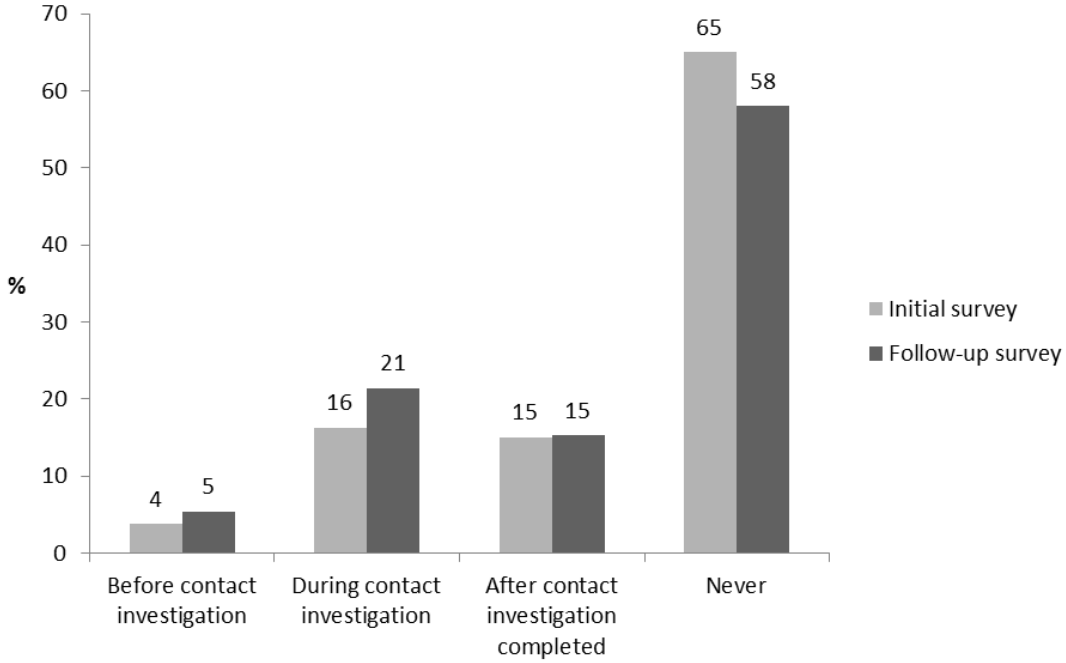
6.2.2.1 Overall estimates of contact tracing activity and yield

Of the 66 nurse respondents to the survey, 30% and 49% completed the questions about contact tracing yield in the initial and follow-up surveys respectively, with the majority (75%) reporting on five recent index cases. In total, the details of 86 index cases were recorded in the initial survey and 141 in the follow-up survey. The majority of index cases were from low TB incidence areas (51%), 28% were from medium incidence areas and 18% were from high incidence areas, which is proportional to the TB incidence across the country.

For the majority of cases, nurses reported that the strain type for the index cases never became available (65% and 58% from the initial and follow-up surveys; Figure 39). Only 4% of strain typing results in the initial survey and 5% in the follow-up survey were available before contact tracing activities had started. When the strain typing data were made available to nurses did not change between the two surveys (chi^2=1.34, p=0.72).

561 contacts were identified by 85 index cases in the initial survey and 1172 contacts were identified by 138 index cases in the follow-up survey (missing = 1 from initial and 3 from follow-up surveys; Table 25). The median number of contacts identified and screened was between 3 and 4 per index case, and the median number of contacts found to have active disease and LTBI was zero across the two surveys. The median number of contacts identified by the index case, screened, and found to have active TB disease or LTBI did not change significantly over time (p>0.01 for all).

Figure 39 – Contact tracing yield: The point at which the strain typing data for individual cases was made available to nurses



The proportion of contacts screened and the proportion of contacts with LTBI decreased significantly between the two surveys (contacts screened decreased from 91.4% to 86.9%, p=0.005; contacts with LTBI decreased from 12.1% to 8.4%, p=0.023; Table 26). The proportion of contacts with active disease did not change significantly between the surveys (2.1% and 1.8%, p=0.610).

Table 25 –The mean (standard deviation), median (inter-quartile range) and range of the number of contacts identified, screened, with active TB disease and with LTBI: responses from the initial and follow-up cross-sectional surveys.

		Initial survey (n=86)					Follow-up survey (n=141)					
	Index case site of disease	Number of index cases	Total contacts	Mean (SD)	Median (IQR)	Range	Number of index cases	Total contacts	Mean (SD)	Median (IQR)	Range	p-value
Contacts identified by the index case												
Total		85	561	6.6 (10.4)	3 (2-7)	0-64	138	1172	8.5 (22.2)	4 (2-7)	0-200	0.300
	Pulmonary	55	471	8.6 (12.4)	4 (2-10)	0-64	88	1027	11.7 (27.3)	5 (3-9)	0-200	0.288
	Extra-pulmonary	30	90	3 (2.7)	2 (1-5)	0-10	50	145	2.9 (2.0)	3 (2-4)	0-10	0.662
Contacts screened												
Total		83	513	6.2 (10.2)	3 (1-6)	0-64	134	1018	7.6 (18.2)	3.5 (2-7)	0-150	0.609
	Pulmonary	55	432	7.9 (12.1)	3 (2-9)	0-64	87	904	10.4 (22.1)	5 (3-9)	0-150	0.315
	Extra-pulmonary	28	81	2.9 (2.6)	2 (1-4)	0-10	47	114	2.4 (1.9)	3 (1-4)	0-7	0.665
Contacts with active disease												
Total		83	11	0.1 (0.7)	0 (0-0)	0-5	131	18	0.1 (0.5)	0 (0-0)	0-4	0.333
	Pulmonary	55	11	0.2 (0.9)	0 (0-0)	0-5	85	16	0.2 (0.6)	0 (0-0)	0-4	0.531
	Extra-pulmonary	28	0	0 (0)	0 (0-0)	0-0	46	2	0.0 (0.2)	0 (0-0)	0-1	0.267
Contacts with LTBI												
Total		83	62	0.7 (1.9)	0 (0-1)	0-13	130	86	0.6 (1.5)	0 (0-1)	0-10	0.977
	Pulmonary	55	59	1.1 (2.2)	0 (0-1)	0-13	84	78	0.9 (1.8)	0 (0-1)	0-10	0.933
	Extra-pulmonary	28	3	0.1 (0.3)	0 (0-0)	0-1	46	8	0.2 (0.5)	0 (0-0)	0-2	0.909

[a]p-value from the Wilcoxon rank sum test
Results are reported by site of disease of the index case.

Table 26 –The proportion (CI)[a] of contacts screened, with active disease and LTBI: responses from the initial and follow-up cross-sectional surveys.

	Initial survey % (CI)	Follow-up survey % (CI)	p-value[b]
% (CI) of identified contacts screened			
All index cases	91.4 (88.8-93.6)	86.9 (84.8-88.7)	0.005
Index case pulmonary TB	91.7 (88.9-94.0)	88.0 (85.9-89.9)	0.032
Index case extra-pulmonary TB	90.0 (81.9-95.3)	78.6 (71.0-85.0)	0.024
% (CI) of contacts screened with active TB			
All index cases	2.1 (1.1-3.8)	1.8 (1.1-2.8)	0.610
Index case pulmonary TB	2.5 (1.3-4.5)	1.8 (1.0-2.9)	0.346
Index case extra-pulmonary TB	0.0 (0.0-4.5)	1.8 (0.2-6.2)	0.512[c]
% (CI) of contacts screened with LTBI			
All index cases	12.1 (9.4-15.2)	8.4 (6.8-10.3)	0.023
Index case pulmonary TB	13.7 (10.6-17.3)	8.6 (6.9-10.7)	0.005
Index case extra-pulmonary TB	3.7 (0.8-10.4)	7.0 (3.1-13.4)	0.367[c]

[a]Exact confidence intervals calculated based on the binomial distribution
[b]Chi^2 test for significance
[c]Fisher's exact test for significance

6.2.2.2 The impact of clustering and cluster investigations on contact tracing yield

In 2011, for NCL sector and Leicester TB clinics, there were 220 pulmonary TB cases with unique strain types and 97 clustered pulmonary TB cases who had information on their traced contacts. Of the clustered cases, 29 were in a cluster that had been investigated and 68 were in a cluster that was not investigated.

The characteristics of cases in clusters that were investigated versus clusters that were not investigated were broadly similar (Table 27). There were some exceptions: the proportion of cases with any drug resistance was higher in clusters that were **not** investigated, though non-significant (3.0% versus 14.5%, p=0.081), and recent incarceration in prison was significantly higher in clusters that were not investigated (6.1% versus 13.0%, p=0.018).

There was a greater proportion of non-UK born cases (compared to UK born) in clustered cases (investigated or not), consistent with national surveillance data.[3] A significantly larger proportion of cases in a cluster that was investigated were UK born compared to cases in a cluster that was not investigated (39.4% versus 23.3%, p=0.020).

The median number of contacts screened, with active TB and LTBI was significantly greater in clustered compared to unique cases (Table 29). No difference was observed between cases that were part of a cluster that was investigated compared to not investigated (p>0·1 for all).

A sensitivity analysis was conducted to assess the impact of missing data. The analysis assumed that index cases with missing contact tracing information yielded no (zero) cases of active disease or LTBI. The results showed that the comparison of median yields in unique and clustered cases remained the same, but were no longer

significant (p=0.06 for active cases and p>0.1 for LTBI) (Table 30). Median yields were broadly similar for cases in clusters investigated or not (p>0.1 for both).

The estimated proportions of contacts screened in clustered cases that were investigated or not investigated that had active disease and LTBI were high as all missing data were excluded from the analysis (2.4% and 5.1% of contacts in clustered cases that were investigated or not investigated, respectively, were diagnosed with active TB, and 22.1% and 27.8% with LTBI) (Table 29). Under these assumptions, the contact tracing yield was greater in cases that were in clusters that were not investigated, compared to those investigated, but these findings were not significant (p>0.1).

A more realistic estimation of the proportion of contacts screened with active disease and LTBI is shown in the sensitivity analysis, in which 1.6% and 3.0% of contacts had active TB and 16.8% and 20.5% of contacts found had LTBI in clustered cases that were investigated or not investigated, respectively. Under these assumptions, there was no significant difference between the contact tracing yield in clusters that were investigated versus not investigated (p>0.1) (Table 30). The assumptions in the sensitivity analysis mean that the denominator was total contacts screened in the dataset, rather than contacts of index cases for which information was available.

The mean number of contacts screened per LTBI found is shown in Table 28, taking into account the proportion of cases that had a unique strain type, in a cluster that was investigated or in a cluster that was not investigated. These estimates were used to parameterise the cost-effectiveness model (page 175).

Table 27 – Characteristics of cases in clusters that were investigated versus cases in clusters that were not investigated.

Characteristics of cases	Clustered investigated (%)	Clustered not investigated (%)
Total	33	69
Smear positive	14 (42.4)	28 (40.6)
Drug resistant	1 (3.0)	10 (14.5)
<16 years old	2 (6.1)	1 (1.4)
Homeless	1 (3.0)	1 (1.4)
Prison	2 (6.1)	9 (13.0) *
Drug/alcohol abuse	4 (12.1)	7 (10.1)
Previous TB treatment	0 (0)	2 (2.9)
UK born	13 (39.4)	16 (23.2) *

HIV status and history of severe mental health were not available. Characteristics of clusters such as size when cluster investigated, presence of a child, and cluster level (local/regional/national) were not available.
* Chi2 test using 'diagnosed after cluster investigated' as the reference category, p-value<0.05

Table 28 – The mean number of contacts screened per LTBI found, weighted by whether the strain type was unique, clustered and investigated, or clustered and not investigated

	n (%)	Mean number of contacts per index case screened			Contacts screened per LTBI found
		Screened	Active TB	LTBI	
Not clustered	220 (38)	4.36	0.11	0.93	4.68
Clustered investigated	29 (5)	6.59	0.21	1.68	3.91
Clustered not investigated	68 (12)	5.44	0.33	1.90	2.86
Total	**317**	**2.6**	**0.1**	**0.7**	**3.97**
Sensitivity analysis: missing=0					
Not clustered	227 (39)	4.22	0.06	0.65	6.48
Clustered investigated	33 (6)	5.79	0.09	0.97	5.97
Clustered not investigated	69 (12)	5.36	0.16	1.10	4.87
Total	**329**	**2.6**	**0.0**	**0.4**	**5.94**

Table 29 – Contact tracing yield for index cases by clustering and whether the cluster was investigated: NCL and Leicester clinic data.

	Unique cases	Total clustered cases	p-value[a]	Cases in a cluster that was investigated	Cases that were in a cluster that was not investigated	p-value[b]
Contacts screened						
Index cases (n)	220	97		29	68	
Contacts screened (n)	959	561		191	370	
Median screened (IQR)	3 (1-5)	4 (2-7)	0·008[c]	4 (2-9)	4 (2-6)	0.474[c]
Contacts with active disease						
Index cases with information available (n)	131	47		14	33	
Contacts screened (n)	593	341		125	216	
Contacts with active disease (n)	14	14		3	11	
% (CI) contacts screened with active TB[de]	2.4 (1.3-3.9)	4.1 (2.3-6.8)	0.132[f]	2.4 (0.5-6.9)	5.1 (2.6-8.9)	0.271[g]
Median active (IQR)	0 (0-0)	0 (0-0)	0·011[c]	0 (0-0)	0 (0-0)	0.896[c]
Contacts with LTBI						
Index cases with information available (n)	159	59		19	40	
Contacts screened (n)	761	410		145	265	
Contacts with LTBI (n)	148	108		32	76	
% (CI) contacts screened with LTBI[de]	19.4 (16.7-22.4)	26.3 (22.1-30.9)	0.006[f]	22.1 (15.6-29.7)	27.8 (23.3-34.5)	0.146[f]
Median latent(IQR)	0 (0-1)	1 (0-2)	0·016[c]	1 (0-2)	1 (0-2)	0.330[c]

[a] Comparing unique cases with clustered cases;
[b] Comparing cases in a cluster that was investigated with case that were in a cluster that was not investigated
[c] Wilcoxon rank sum test
[d] Exact confidence intervals calculated based on the binomial distribution
[e] Number with active disease \ number of contacts screened *100
[f] Chi^2 test of significance
[g] Fisher's exact test of significance

Table 30 – Sensitivity analysis. Contact tracing yield for index cases by clustering and whether the cluster was investigated, assuming that index cases with missing contact tracing information yielded no cases of active disease or LTBI: NCL and Leicester clinic data

	Unique cases	Total clustered cases	p-value[a]	Cases in a cluster that was investigated	Cases that were in a cluster that was not investigated	p-value[b]
Contacts screened						
Index cases (n)	227	102		33	69	
Contacts screened (n)	959	561		191	370	
Median screened (IQR)	3 (1-5)	4 (1-7)	0.02 [c]	4 (1-8)	4 (2-6)	0.872 [c]
Contacts with active disease						
Index cases with information available (n)	227	102		33	69	
Contacts screened (n)	959	561		191	370	
Contacts with active disease (n)	14	14		3	11	
% (CI) contacts screened with active TB[de]	1.5 (0.8-2.4)	2.5 (1.4-4.2)	0.147[f]	1.6 (0.3-4.5)	3.0 (1.5-5.3)	0.401[g]
Median active (IQR)	0 (0-0)	0 (0-0)	0.059 [c]	0 (0-0)	0 (0-0)	0.819 [c]
Contacts with LTBI						
Index cases with information available (n)	227	102		33	69	
Contacts screened (n)	959	561		191	370	
Contacts with LTBI (n)	148	108		32	76	
% (CI) contacts screened with LTBI[de]	15.4 (13.2-17.9)	19.3 (16.1-22.8)	0.055 [f]	16.8 (11.8-22.8)	20.5 (16.5, 25.0)	0.288 [f]
Median latent (IQR)	0 (0-1)	0 (0-1)	0.418 [c]	0 (0-2)	0 (0-1)	0.436 [c]

[a] Comparing unique cases with clustered cases
[b] Comparing cases in a cluster that was investigated with case that were in a cluster that was not investigated
[c] Wilcoxon rank sum test
[d] Exact confidence intervals calculated based on the binomial distribution
[e] Number with active disease \ number of contacts screened *100
[f] Chi^2 test of significance
[g] Fisher's exact test of significance

6.2.3 Indirect measures of the impact of the TB-STS on TB control: Diagnostic delay

An additional mechanism that the TB-STS may help to interrupt TB transmission is through the earlier diagnosis of TB disease, preventing further transmission in the community. The earlier diagnosis may be possible if an investigation of a cluster leads to more targeted screening and public health action. This could lead to the identification of contacts or people exposed to TB who have TB disease but may not have sought care until later, thereby reducing diagnostic delay. The methods for this case-control study can be found on page 76.

There were 526 clustered cases of pulmonary TB in England in 2011, after removing the first and second cases within each cluster to take into account possible household transmission. 208 were in a cluster that was not investigated and 318 were in a cluster that was investigated. Information on diagnostic delay was available for 139 and 238 cases, respectively (Table 32).

The characteristics of cases in clusters that were not investigated, cases that were diagnosed before the cluster was investigated and cases diagnosed after the cluster was investigated were broadly similar (Table 31). There were, however, some differences. There was a higher (non-significant) proportion of smear positive cases in clusters that were investigated compared to cases in a cluster that were not investigated (61.5% or 60.3% versus 55.4%, respectively, $p>0.05$) – this is not surprising as smear positivity is a risk factor for transmission. There was a significantly higher proportion of drug resistant cases in clusters that were not investigated (22.3% versus 7.4%, $p>0.05$), similar to the population from NCL and Leicester (Table 27, page 150).

The median diagnostic delay was not significantly different in cases that were in a cluster that was not investigated (n=139; 62 days) or diagnosed before a cluster investigation was launched (n=117; 85 days), compared to those diagnosed after the start of the cluster investigation (n=121; 77 days). The results did not change when stratifying by UK born (n=293), non UK born (n=328) and excluding cases under the

age of 16 (n=12) (Table 33). As previously observed, UK born cases had a longer diagnostic delay compared to the non-UK born cases.[35]

A sensitivity analysis was conducted to see the effect of using date of notification instead of date of treatment to define diagnostic delay, but this did not change the results.

Table 31 – Characteristics of cases in clusters that were not investigated, cases that were diagnosed before the cluster was investigated, cases diagnosed after the cluster was investigated.

		Cluster investigated	
Characteristics of cases	**Cluster not investigated (%)**	**Diagnosed before cluster investigated (%)**	**Diagnosed after cluster investigated (%)**
Total	139	117	121
Smear positive	77 (55.4)	72 (61.5)	73 (60.3)
Drug resistant	31 (22.3) [a]	10 (8.5)	9(7.4)
<16 years old	7 (5.0)	6 (5.1)	4 (3.3)
Homeless	13 (9.4)	12 (10.3)	14 (11.6)
Prison	20 (14.4)	13 (11.1)	15 (12.4)
Drug/alcohol abuse	9 (6.5)	5 (4.3)	6 (5.0)
Previous TB	9 (6.5)	14 (12.0)	6 (5.0)
UK born	85 (61.2)	74 (63.2)	82 (67.8)

HIV status and history of severe mental health were not available. Characteristics of clusters such as size when cluster investigated, presence of a child, and cluster level (local/regional/national) were not available.
[a] Chi2 test using 'diagnosed after cluster investigated' as the reference category, p-value<0.05

Table 32 – The diagnostic delay for clustered pulmonary cases from 2011, based on whether they were in a cluster that was investigated, and whether they were diagnosed before or after the investigation was launched

	Cases that were in a cluster that was not investigated	Cases in a cluster that was investigated	
		Diagnosed before the investigation was launched	Diagnosed after the investigation was launched
Number of clusters	120	58	38
Number of cases	139	117	121
Mean (SD) (days)	113.2 (117.9)	140.8 (271.9)	113.5 (128.7)
Median (IQR) (days)	62 (32-127)	86 (47-155)	77 (41-157)
p value[a]	0.1567	0.4257	ref

[a]p-value for the Wilcoxon rank sum test

Table 33 – The diagnostic delay for UK born, non-UK born, and cases over the age of 16 who are clustered pulmonary TB cases from 2011, by cluster investigation and whether they were diagnosed before or after the investigation was launched.

	Cases that were in a cluster that was <u>not</u> investigated	Cases in a cluster that was investigated	
		Diagnosed <u>before</u> the investigation was launched	Diagnosed <u>after</u> the investigation was launched
UK born			
Number of clusters	38	48	28
Number of cases	54	56	66
Mean (SD)	120.2 (85.4)	135.1 (226.2)	110.3 (93.1)
Median (IQR)	109 (55-168)	76 (42.5-144.5)	81.5 (42-162)
p value[a]	0.3278	0.6962	ref
Non-UK born			
Number of clusters	27	81	24
Number of cases	59	71	51
Mean (SD)	161.6 (374.2)	104.8 (143.9)	124.8 (166.4)
Median (IQR)	69 (43-133)	60 (31-122)	80(45-161)
p value[a]	0.7146	0.1411	ref
≥16 years			
Number of clusters	118	58	38
Number of cases	137	114	117
Mean (SD)	113.2 (179.1)	144.1 (274.8)	116.2 (130.0)
Median (IQR)	62 (32-125)	92 (51-155)	80 (42-161)
p value[a]	0.0868	0.4200	ref

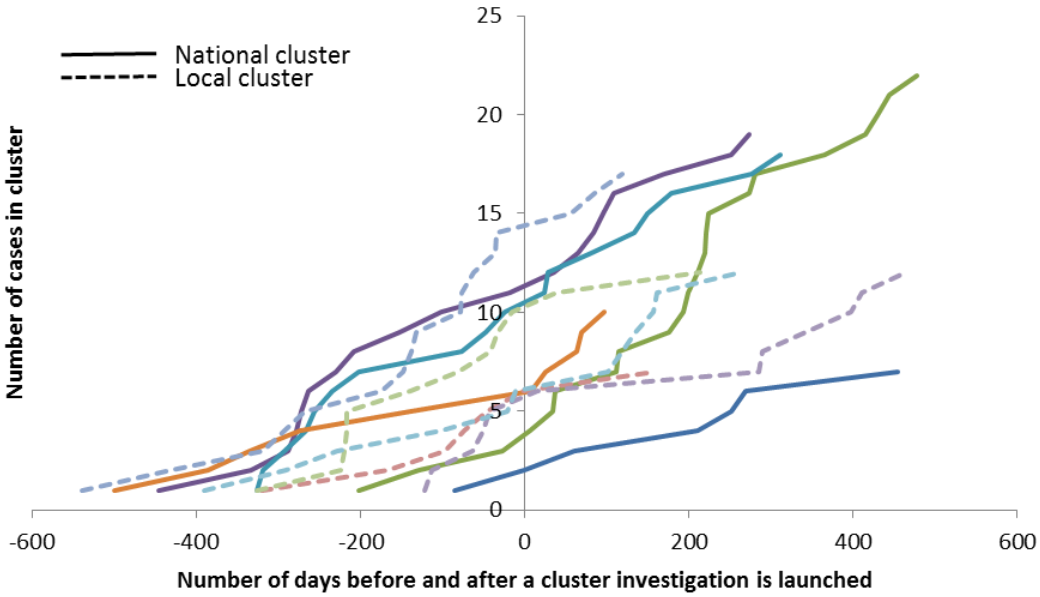
[a]p-value for the Wilcoxon rank sum test

6.2.4 Indirect measures of the impact of the TB-STS on TB control: Rate of cluster growth

The relative change in the size of clusters was explored. It was hypothesised that the rate of cluster growth would differ before and after a cluster investigation was initiated. Cluster investigation may increase apparent rate of cluster growth (as a result of additional cases being identified more quickly) or may decrease the rate of cluster growth (if the earlier identification of cases limits transmission). In an ideal circumstance a cluster investigation may transiently increase the rate of cluster growth and then later decrease the rate. The methods for this study are on page 77.

Figure 40 shows examples of cluster growth over time and the size of clusters at the point they are investigated. The local clusters shown here are not representative as they do not include any clusters that are investigated with just two cases (as this isn't well visualised as a line graph). Figure 38 (page 141) shows the distribution of clusters sizes at the point they are investigated, giving an overview of the size of clusters when they are investigated.

Figure 40 – Examples of the rate at which TB cases were added to national and local clusters during 2010 and 2011. The figure shows the rate of cluster growth in relation to the cluster investigation.



The univariate analysis identified cluster type (local, regional and national), case order and whether diagnosed before or after the cluster investigation as potential factors associated with the rate of cluster growth. The multivariable linear regression had significant ($p<0.01$) coefficients for the intercept, the cluster level being national (i.e. additional cases accruing in 'national' clusters are added more quickly compared with cases occurring in local clusters), and the case order (i.e. each subsequent case in a cluster arrives sooner than the previous ones, possibly reflecting the fact that as clusters spread there are more infectious TB cases so more opportunities for transmission and thus the intervals between subsequent cases are expected to decrease, or that intervals between cases are more likely to represent differences in variability in incubation period rather than intergenerational time as clusters grow). No other variables were significant (Table 34).

Table 34 – Regression analysis investigating the impact of cluster type (local, regional, national), case order and whether diagnosed before or after the cluster investigation on the rate of cluster growth

		Coefficient	95% CI	p-value
Cluster type	Local	ref		
	Regional	-4.9	-29.9, 20.0	0.698
	National	-46.0	-60.7, -31.3	<0.001
Case order		-1.1	-1.5, -0.7	<0.001
Diagnosed after investigation		0.0	-11.3, 11.3	0.997
Intercept		99.4	65.5, 133.2	<0.001

6.3 Discussion

6.3.1 Strengths and limitations of the methods used to evaluate the Outputs of the TB-STS

Laboratory Outputs

The false positive TB isolation survey was designed to observe how frequently the reference laboratories queried isolates for potential cross-contamination based on the strain typing information, and to establish how many of these queries had not been previously identified. It did not attempt to measure the cross-contamination rates within TB laboratories in England. This was important to ensure cooperation from

the laboratories who might not have shared their data so freely if the service they provide was being evaluated. As a result, the findings can only be interpreted in the context of the TB-STS, and cannot be used to assess the quality of TB laboratory services in general.

The two-step process of collecting information from the reference laboratories and then following up with the source laboratories was necessary because although the reference laboratories would write to the source laboratories to alert them of a possible cross-contamination event, they would not follow up with them to find out the outcome of their query. In fact, the reference laboratories had heard back from source laboratories in 19 incidents out of 70 (27%). The follow-up survey to source laboratories was an appropriate method of data collection (compared to a telephone survey) as it gave the recipient time to find out the result of an investigation before responding. However, due to the survey method, very few questions could be posed, so the information collected was limited. Source laboratories may have wished to appear well-informed and therefore responded that they were already aware of the false positive diagnosis when they were not. This would have biased the survey results, underestimating the impact of the TB-STS at detecting these incidents. Future evaluations should conduct an audit to find out the outcomes of each incident, especially the eight patients who were unnecessarily started on treatment (Table 20).

Public health Outputs

Cluster investigation activities and outcomes

The delay in the development of the STM until after the evaluation period severely limited the data collection around cluster investigations. Interim reporting systems were established in order to capture some of this activity. These were designed with stakeholders to ensure their utility (for the service to be monitored as well as for the evaluation) and usability. There were, however, many limitations to these systems so they could not be used to parameterise the transmission and cost-effectiveness models.

The cluster monitoring database was not linked to the routine laboratory, clinical and epidemiological data, and it required CIs to manually input each case that was part of a cluster. As a result, the database had lots of missing fields and was not regularly updated. The cluster monitoring database focussed on whether epidemiological links between clustered cases were identified. However, this was retrospective and was only relevant if it had an impact on subsequent public health action. The database should have collected information on the actions taken because of the strain typing information and the resulting outcomes of these actions.

In addition to justifying further investigation, the TB-STS could be used to disprove transmission between cases where transmission had been suspected. For example, the diagnosis of TB disease in multiple students attending the same college within a few weeks of each other may suggest on-going transmission at the college. If the strain types are distinguishable, there is no evidence for transmission between the students. In this scenario, the strain typing results will have prevented a large investigation and screening programme around the students and the college. Data on the activity avoided as a result of the TB-STS was not collected (Figure 11, page 31). Capturing activities that do not happen is challenging. This problem was addressed by using the cross-sectional survey and semi-structured interviews to gather examples of where strain typing has been used to disprove suspected transmission. Although these insights captured through qualitative methods cannot tell us the frequency of this event, they can be used to demonstrate whether these events occur at all.

Refuting possible transmission was one of the main uses of the strain typing information that was reported in the initial and follow-up survey. In addition, this was one of the sub-themes identified in the semi-structured interviews with HPUs. See Table 17 (page 124) for examples of refuting possible transmission between cases. Recall bias may play an important role here as people may remember incidents where the strain typing results prevented public health action more clearly than where strain typing did not add anything to a particular incident. Alternatively, it may work in the other direction where people are less likely to remember where

they did not take action, compared to where the strain typing led to further public health action.

Contact tracing yield

Strengths and limitations of the initial and follow-up survey design are discussed in Chapter 4 (page 128). The specific questions around contact tracing are discussed here. To minimise the potential recall bias that might occur when asking nurses about the details of patients, they were asked to record the details of five recent cases for whom contact tracing had been completed. However, it is possible that they reported the five most memorable recent cases (e.g. with unusual outcomes), the details of which they could remember most easily. To try to avoid this, they were prompted to get their case notes before completing the contact tracing questions.

The significant decrease in the proportion of contacts screened and with LTBI is surprising given that the TB-STS aims to help target contact tracing more effectively. The decrease in the proportion of contacts screened may be because there was an increase in the number of contacts identified between the two surveys (though this was not significant). Alternatively, the TB-STS may have encouraged people to prioritise clustered cases, so that the contacts of other cases are not screened. However, this would not explain why the yield of LTBI decreased as well – under these circumstances you would expect them to increase, unless patients that had infected fewer of their contacts were being prioritised. Regardless, these results reject the hypothesis that the TB-STS will increase contact tracing yield.

The data used to investigate the impact of cluster investigations on contact tracing yield were from two study sites in England: the North Central London Sector and Leicester. This was a convenience sample based on which TB services routinely collected the relevant data. The outcomes of contact investigations were not collected on the central notification system. Whilst this sample covers two TB services in different regions of England, it is not a representative sample as both settings have a high TB incidence (Leicester City 57/100,000 in 2011 and NCL 34/100,000 in 2011).[4,217] There may be a self-selection bias whereby the sites that collect data on

their contact tracing activities might therefore, due to an increased interest in the metric, provide a more effective contact tracing service compared to sites that do not collect these data. If this were the case, then the impact of the TB-STS may have been harder to observe due to its marginal effect on an effective contact tracing service.

The analysis only included pulmonary TB cases, which accounts for ~50% of TB cases in England, so the impact of the TB-STS on contact tracing for extra-pulmonary TB cases is not known. Because of the smaller chance of transmission from a case of extra-pulmonary TB, contact tracing yield is usually lower in these cases,[24] suggesting that the contact tracing yield estimated in this thesis is overestimated.

The characteristics of cases in clusters that were investigated versus clusters that were not investigated were broadly similar. However, the proportion of cases with any drug resistance was higher in clusters that were **not** investigated and recent incarceration in prison was significantly higher in clusters that were not investigated. This is surprising given that drug resistance or recent incarceration in prison are criteria for cluster investigation, but may also reflect incidents where lots is known about the epidemiology of TB in these cases before the strain typing results are received (for example, once all exposed contacts in a prison are screened following diagnosis of the index case, further screening around a secondary case is unlikely). This might go some way to explain why the contact tracing yield is slightly higher (though not significantly so) in cases that were not investigated compared to those investigated. In addition, greater effort may go into contact tracing around a drug resistant patient or someone with a history of prison (in line with national guidance).[48]

It is surprising that the contact tracing yield was greater (though not significantly) for cases in clusters that were *not* investigated compared to cases that were in clusters that were investigated. This suggests that the screening process was *less* effective with the addition of the TB-STS, possibly due to the selection criteria to actively

investigate clusters (this interpretation influenced the recommendations about the future of cluster investigations (Box 6, page 204)). On the other hand, this could be expected as the greatest contact tracing yield of an infectious index case is likely to be found through the initial household contact tracing and any further screening associated with a cluster investigation is, by its nature of 'casting the net wider', going to have a lower return. If this latter hypothesis is true, the benefit of finding and treating an undiagnosed case of LTBI or active disease must outweigh the additional cost of the *less* efficient contact tracing as well as the costs of the TB-STS for the service to be cost-effective.

Another explanation for this finding is that some cases would have been diagnosed and immediately be part of an ongoing cluster investigation. These cases may have been a contact of another case in the cluster, and therefore have fewer contacts to trace (as they may have shared some contacts with their index case). Bearing this scenario in mind, as more cases are diagnosed through active case finding (e.g. cluster investigation), the contact tracing yield may decrease, rather than increase as more cases have shared contacts.

A strength of this analysis, is the link between the clinic data on contact tracing yield and the clustering data, enabling the investigation into the impact of cluster investigation on contact tracing yield. However, with only one year of data available (2011), it is not surprising that no impact was observed. Data for cases diagnosed in 2011 were collected up to July 2012, allowing for the time it takes to conduct contact tracing and cluster investigations.

The sensitivity analysis conducted (which assumed that missing data on the number of contacts screened with active disease or LTBI were zeros) is more realistic because it is unlikely that active cases and LTBI identified through contact tracing would not be reported; it is more likely that missing values represent zeros.

Diagnostic delay

Diagnostic delay is a key indicator for the success of a TB programme. Despite this, the date of symptom onset necessary to calculate diagnostic delay was only recorded in 75% (238/318) of pulmonary TB cases diagnosed in 2011 that were in an investigated cluster. Thus the sampling fraction was reduced and a bias potentially introduced as those cases with no recorded symptom onset may have had a particularly long delay to diagnosis if, for example, they could not remember when their symptoms had started.

The characteristics of cases in clusters that were not investigated, cases that were diagnosed before the cluster was investigated and cases diagnosed after the cluster was investigated were broadly similar. The higher proportion of drug resistant cases in clusters that were not investigated compared to those that were diagnosed after the investigated was initiated was, on the one hand, unexpected as drug resistance is a criterion for cluster investigation. On the other hand, if a large drug resistance outbreak is known about in detail, the identification of an additional case may not lead to an investigation.[91] This was consistent with the data from NCL and Leicester used for the investigation of contact tracing yield (though numbers were too small in for a significant difference to be detected). One would expect drug resistant cases to be identified earlier and have a shorter diagnostic delay because of the increased contact tracing effort surround a case with drug resistant TB. Therefore diagnostic delay may have been underestimated in the cases that were not in a cluster that was investigated. There were no significant differences between cases that were diagnosed before or after cluster investigations were initiated.

This analysis gives a snapshot of the impact of cluster investigations initiated in 2011 on the diagnostic delay in cases diagnosed during the same year. A strength of this analysis is that it compares the diagnostic delay in cases diagnosed before or after a cluster investigation, something that has not been done before. It is limited by having only one year of data, and may be under powered to detect the observed difference.

In this analysis there were three groups: cases diagnosed before a cluster investigation, after an investigation and cases in a cluster that was not investigated. Ideally, you would want to compare patients with similar characteristics and risk profiles who are clustered and investigated, to patients who are clustered and not investigated and then randomly assign each cluster to be either investigated or not. Or, more specifically, randomly assign each cluster that is recommended for investigation following the preliminary cluster review (though this would not be ethnically sound as there is evidence that certain patient characteristics are risk factors for transmission). The problem with the approach taken here is that those cases diagnosed before an investigation might be different to those diagnosed after an investigation. And cases in a cluster that is not investigated are likely to be different to cases in a cluster that is investigated, as characteristics of cases within a cluster are used to determine whether a cluster should be investigated or not. However, given that the service was implemented nationally, without the opportunity to randomise clusters to be investigated or not, comparing the diagnostic delay before and after the initiation of a cluster investigation is an indication of how close the TB service is to interrupting transmission (on the assumption that if diagnostic delay decreases to zero, transmission has been interrupted).

Rate of cluster growth

The multivariate linear regression examining cluster growth is appropriate because it takes full advantage of all the available data (i.e. data from the cluster monitoring database and the ETS). However, there are some notable limitations, for example it may be biased toward the null by potentially "averaging out" the positive (through more aggressive case finding) and negative (through limiting transmission opportunities from cases detected more rapidly) effects of the cluster investigation on cluster growth, and the model excludes the first case in each cluster.

This analysis was conducted using two years of data, meaning that there was a maximum two-year follow-up per cluster, with most clusters having less than two years of follow-up (assuming the second case in each cluster was not diagnosed on

1st January 2010). This may not have been enough time to observe the change in rate of cluster growth after the initiation of a cluster investigation.

In addition, the date of the start of an investigation may not be very meaningful. The start of an investigation is the date the CI assesses the cluster as having met the threshold for investigation and launches an investigation by sending the cluster report and cluster questionnaires to the relevant HPUs or TB nurses. The date that these reports are then acted upon (which is the date from which any impact could be seen) is unknown. The time between the launch of the investigation by the CIs and the action taken could be very lengthy, rendering the investigation start date arbitrary.

6.3.2 How this relates to previous studies

Laboratory outcomes

A study conducted in London reported a cross-contamination rate of between 0.5% and 0.9%, which is higher than the 0.3% reported here. A slightly lower estimate is to be expected because the figure reported here is based solely on the incidents queried because of the TB-STS, rather than any other information (such as clinical suspicion). One explanation for this is the variation between the laboratories. Although a protocol for identifying false positive isolates using strain typing was shared between the three laboratories, there was a difference in the number of queries raised, with the Newcastle laboratory only querying one incident (0.06%) in a 24-month period. Excluding Newcastle from the analysis, however, did not change the cross-contamination estimate because Newcastle type a small proportion of overall isolates. The discrepancy between the laboratories raises questions about how closely the protocol is followed and whether the threshold for querying isolates was appropriate. In addition to the low rate of cross-contamination, the number of true false positive isolates identified per year because of the TB-STS was lower than expected.[153]

The unexpectedly low numbers of false positive TB isolates identified by the TB-STS may be explained by the timing of the evaluation. Although the TB-STS was not

launched until January 2010, two of the three TB reference laboratories had been conducting TB strain typing for research undertaken during the previous decade (using IS*6110* RFLP, or 12 or 15 MIRU-VNTR). Any major problems with laboratory protocol may have already been picked up by the strain typing data and changes to the protocol already made. As a result, we may be trying to measure something very small. The impact of strain typing on false positives is likely to come just after the introduction of strain typing to a population. However, given that the identification of false positive TB isolation was one of the justifications for the introduction of the TB-STS and was hypothesised to influence the results of the cost-effectiveness analysis (Figure 11, page 31), our finding of a minimal impact is important.

Cluster investigation activity

Electronic systems to link strain typing data to surveillance data exist in other countries with strain typing services (e.g. TB-GIMS in the USA).[218] Nationwide population studies reporting the cluster investigation activities of the Netherlands and the USA suggest that, with the necessary systems in place, these data can be collected systematically.[127] Where comprehensive data collection and reporting systems exist, assessing the cluster investigation activity can give useful insight into the impact of strain typing on TB control.[127]

Contact tracing yield

The estimates of contact tracing yield from the survey population were comparable to published UK estimates of the contact tracing yield for pulmonary TB cases (Table 35).[21,24] The contact tracing yield in clustered and unique cases has not been previously studied. The estimated number of clustered and unique cases from NCL and Leicester were higher than the survey sub-study, those reported in the cluster outcome reporting form, and previous studies from the UK (but more consistent with studies conducted elsewhere in Europe[26] or on smear positive pulmonary TB cases in the USA[31] (Table 1, page 12)), but the sensitivity analysis results were closer to the survey estimates. The differences might be explained by the high TB incidence in the study sites, or the self-selection bias described above (whereby sites that have a

stronger contact tracing service may be more likely to have enhanced data collection). The survey estimates, on the other hand, may be subject to recall bias in addition to being mostly from low TB incidence areas (51% of TB cases from the survey estimate were from a low TB incidence HPU, 28% medium incidence and 18% high incidence). Therefore, the survey estimates are likely to be better representative of the contact tracing yield across England, whereas the clinic estimates more represent the contact tracing yield in areas with a high burden of TB. The cluster outcome reporting form contains important information on the impact of cluster investigations on contact tracing yield and provides (weak) evidence that yield does not change with a cluster investigation.

Table 35 – Estimates of contact tracing yield for pulmonary TB cases from the thesis sub-studies and other published studies from the UK

Data source	Table (page)		% contacts with active TB	% contacts with LTBI
Cross-sectional survey	Table 26 (p147)	Initial	2.5	13.7
		Follow-up	1.8	8.6
Cluster outcome reporting form	Table 22 (p142)	Clustered	1.8	9.9
Clinic data from NCL and Leicester (missing values excluded)	Table 29 (p151)	Unique	2.4	19.4
		Clustered	4.1	26.3
Clinic data from NCL and Leicester (missing values=0)	Table 30 (p152)	Unique	1.5	15.4
		Clustered	2.5	19.3
Saunders, Birmingham 1990-2010[24]			2.3	7
Underwood, Tower Hamlets 1997-1999[21]			3.8	6.9

Unique=unique strain types
Clustered=in a strain typing cluster
All studies include pulmonary TB only

The finding that strain typing is rarely received by nurses is not surprising, given that cluster reports are sent to health protection staff in HPUs. What is important, however, is that it suggests that cluster information is not received when it could influence the contact tracing activities. Rather, it is received after contact tracing has already been completed, if at all. This reduces the potential impact of the strain typing information as additional information gathering for a cluster investigation will always be retrospective. Similarly, reports on the genotyping service in the USA found that the median time from specimen collection to the date the strain typing

result was linked to the surveillance data was 13.6 weeks (IQR 9.2-19.0),[216] which is presumably after contact tracing around the patient has been completed. In the Netherlands, however, contact tracing is standardised and the information required for a cluster investigation is collected at the time of contact tracing, rather than retrospectively once the strain typing result has been received. This process should result in the strain type influencing decisions more rapidly as the necessary information about a clustered TB case is already available and nurses are less likely to have to re-interview the patient to collect more information.

Diagnostic delay

This is the first assessment of the TB-STS and its impact on diagnostic delay. Since the presentation of the TB-STS evaluation to PHE in April 2013, diagnostic delay has been recognised as an important indicator for a TB service and was reported in PHE's Tuberculosis Annual Report.[3] The median diagnostic delay across the UK in 2013 was 72 days, which lies within the range reported here (62 to 85 days in clusters that were not investigated and cases diagnosed before the cluster was investigated, respectively).

It is not surprising that the median diagnostic delay was shorter (62 days) in clusters that were not investigated as these are more likely to be smaller clusters of household transmission. Though not significant, the median diagnostic delay was eight days shorter in cases diagnosed after the initiation of a cluster investigation, compared to before, suggesting that the cluster investigation may be identifying cases earlier.

Rate of cluster growth

The rate at which new cases of TB are added to a cluster indicates the rate of ongoing transmission in the community. The rate of cluster growth has not been explored in the UK, but the number and proportion of clusters of various sizes are reported in the national surveillance report as an indicator for TB control.[3,123] The majority of clusters remain small; in the UK between 2010 and 2013 the proportion of clusters of size two was 46.3% (858/1854) and the median cluster size was three (range 2 to 166).[3]

Previous studies have not evaluated the impact of cluster investigations on the rate of cluster growth. For example, Driver *et al*[103] investigated the influence of the demographic, clinical, and epidemiological characteristics of the first two cases in a cluster in New York City on the rate of cluster growth, and Kik *et al*[102] investigated the characteristics of the first two cases in a cluster to see if they predicted whether the cluster would grow to include five or more cases. The analysis presented here suggests that cluster investigations do not have an impact on the rate TB clusters grow within two years of the initiation of the cluster investigation. This could be because the transmission chain has already been interrupted, suggesting that the cluster investigation was either successful, or unnecessary.

During 2010 and 2011 the average length of a cluster investigation was 129 days (n=86, IQR 50-188; see page 140), excluding any ongoing investigations. These findings raise questions about how long a TB service can afford to conduct a cluster investigation before affecting the rate at which the cluster is growing, and whether cluster investigations are being conducted appropriately and in clusters where they could have the most impact.

6.3.3 Summary of findings

Laboratory Outputs

The laboratories typed 17,168 *M. tuberculosis* isolates between 2010 and 2012, 84.6% of which had at least 23 loci.

Seventy (0.6%) of the total isolates typed between 2010 and 2012 were suspected as false positive TB isolates because of their strain type. Thirty of these (43%) were confirmed as cross contamination incidents, of which 17 had not been previously identified by the source laboratories.

Public health Outputs

Cluster investigation activities and outcomes

In the absence of the STM, interim databases and data collection forms were produced to collect data on cluster investigation activities and their outcomes (Table

8, page 92). Cluster investigation activities are summarised across three databases and a case management tool. 188 clusters were recorded in the cluster monitoring database between 2010 and 2012. Clusters were investigated when they included between 2 and 44 cases, and remained open for a median of 129 days (IQR 50-188). The median length of cluster investigations is an underestimate because investigations that had not yet closed were not included in the estimate. The large number of clusters that are investigated when they contain two cases could be because the cases are assessed as high risk in terms of transmission (based on the criteria in the Handbook), or low incidence areas may investigate clusters earlier due to the rarity of the disease and/or availability of resources. (Cluster outcome reporting forms were completed and returned for 42 local and regional clusters. New epidemiological links were found within 20 clusters, but public health action was taken in only three clusters. Strain typing was reported to be useful for 16 clusters (and not useful for 11 clusters). Epidemiological links were established following a cluster investigation for only 2.4% of national clusters that were reported during 2010 and 2011.

Contact tracing yield

There was no difference in the time point that nurses received strain typing data (before, during or after the contact investigation, or never) between the initial and follow-up surveys, with the majority of nurses never receiving it. The median contact tracing yield did not change between the surveys. The proportion of contacts screened and the proportion of contacts with LTBI decreased between the two surveys; the proportion of contacts with active disease did not change.

In NCL and Leicester, the median contact tracing yield was significantly greater in clustered cases compared to unique cases. There was no significant difference in contact tracing yield between cases that were in clusters that were investigated versus clusters that were not investigated.

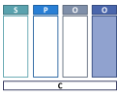
Diagnostic delay

Compared to clustered cases diagnosed after the start of the cluster investigation, the median diagnostic delay was not significantly different in clustered cases that were not investigated or cases diagnosed before a cluster investigation was launched.

Rate of cluster growth

The null hypothesis that cases arriving after initiation of a cluster investigation are associated with a change in the rate new cases are added to a cluster compared to cases arriving before the cluster investigation (adjusting for other variables) cannot be rejected. Therefore, there is no evidence to suggest that the rate at which clusters grow is different before or after a cluster investigation is launched.

Chapter 7. Outcomes



This chapter presents the results of the TB transmission model and the cost-effectiveness model, developed to estimate the effectiveness and cost-effectiveness of the TB-STS. The structure of each model is shown in Figure 24 (page 80) and Figure 25 (page 82), respectively, and the underlying assumptions for the models are described on page 78 and 81, respectively. Further details of the modelling methods and results can be found on pages 78 and 81 and in Appendix 5 and Appendix 6. The model parameters that were estimated from elsewhere in this thesis are shown in Table 36. For my contribution to the transmission model and cost-effectiveness model please refer to 'My role in this thesis' on page xxiv.

Briefly, the baseline scenario (S0) without the TB-STS assumes 3% of LTBI are detected and a diagnostic delay of 12 weeks. Given that there was no measurable impact of the TB-STS on public health Outputs (Chapter 6), all other scenarios represent the potential (not realised) impact of the TB-STS. Therefore, the models investigate what impact the TB-STS would have to have in order to be effective and cost-effective.

Table 36 – Model parameters estimated from sub-studies in this thesis

Parameter	Study source	Page no.
Cost of the TB-STS	Description of Structures	100
Average TB-STS associated workload at an HPU	Initial and follow-up surveys	103
False positive TB isolation	False positive TB isolation survey	138
Mean number of contacts screened per index case	Contact tracing yield	148
Average number of LTBI and active disease identified per index case	Contact tracing yield	148
Average diagnostic delay	Diagnostic delay	153
Impact of the TB-STS on the number of incident TB cases	Transmission model	172

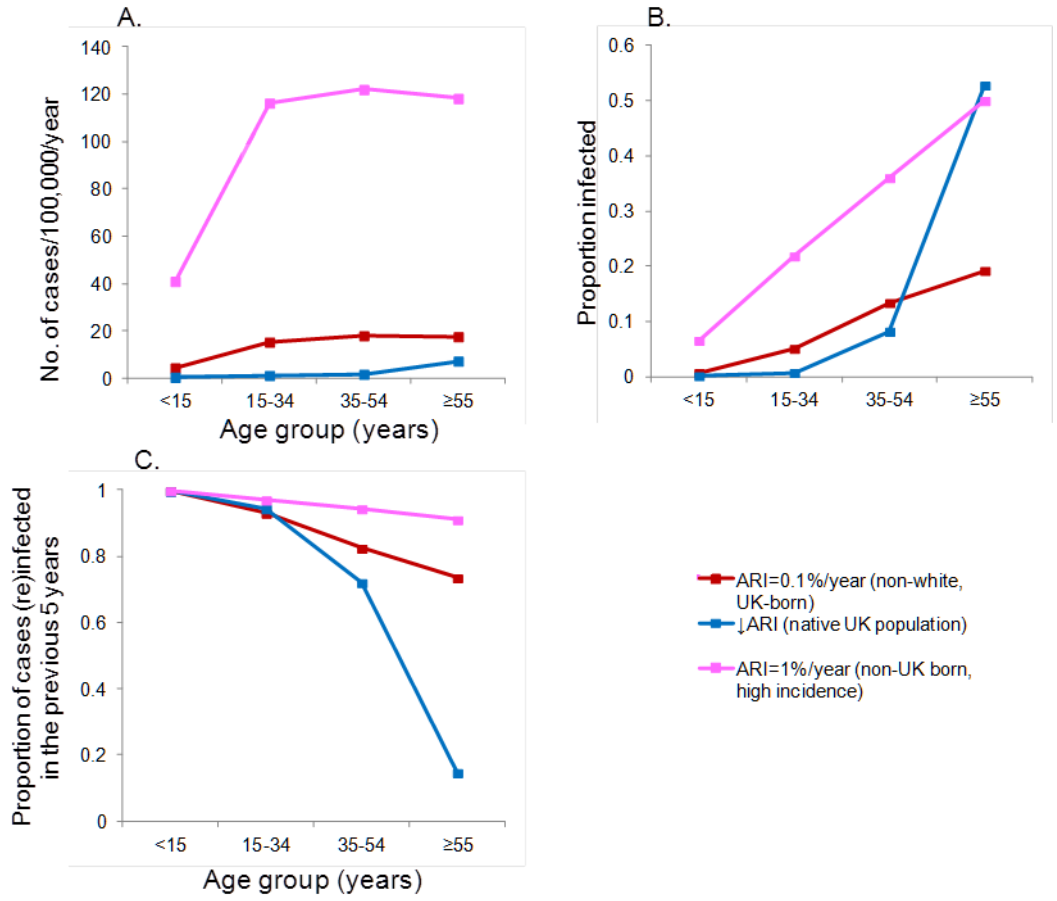
—

7.1 Modelling the effectiveness of the TB-STS

The model structure is shown in Figure 24, page 80.Three epidemiological scenarios were considered in the transmission model, representing the variation in ARI within the UK population. For each scenario, the age-specific annual TB incidence,

proportion of individuals who have ever been infected, and proportion of new cases that have been newly infected or recently re-infected are shown in Figure 41.

Figure 41 – Characteristics of the epidemiological scenarios considered in the transmission model



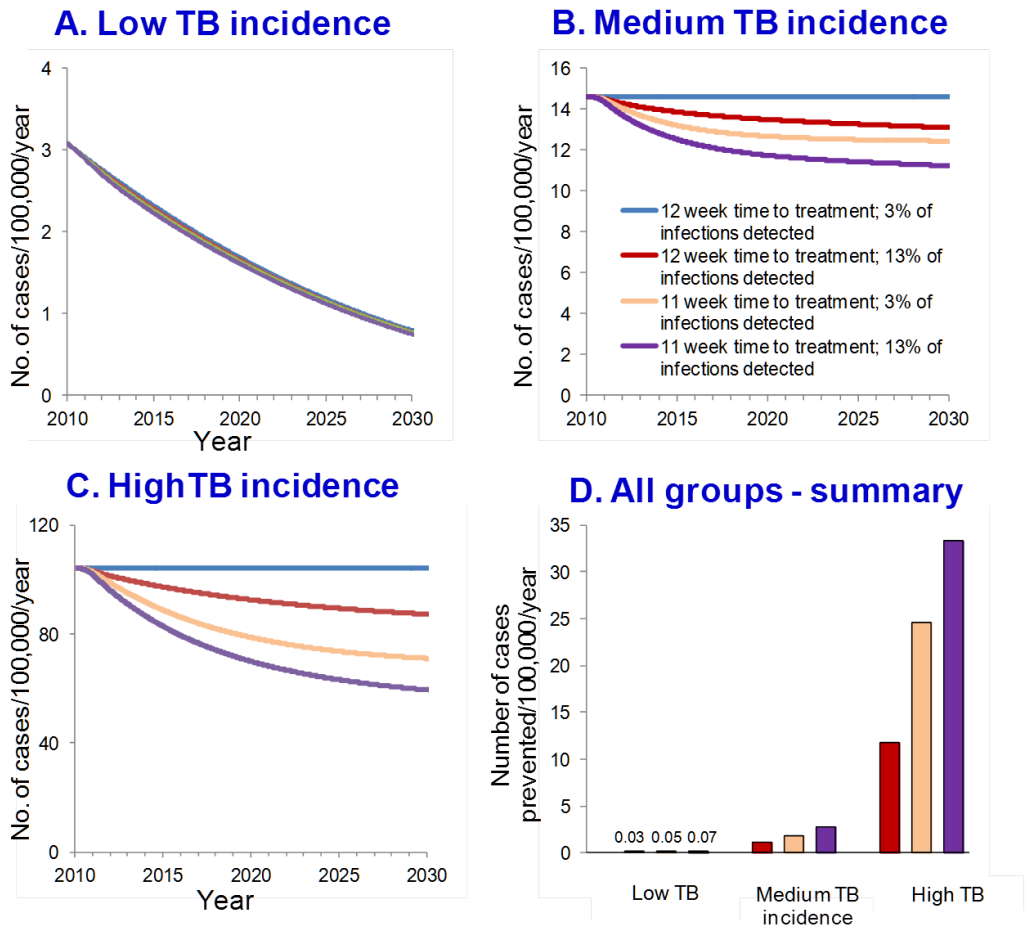
A. The age-specific annual TB incidence per 100,000 population.
B. The age-specific proportion of individuals who have ever been infected.
C. The age-specific proportion of new cases that have been newly infected or reinfected in the previous 5 years
ARI annual risk of infection

Figure 42 summarizes predictions of the impact of the TB-STS on TB incidence for the three epidemiological scenarios considered. For the white UK population, the predicted incidence decreased from four to less than one per 100,000 per year between 2005 and 2030 in the absence of additional interventions. Reducing diagnostic delay and/or increasing the proportion of infections that were detected in this population because of the TB-STS is predicted to have little impact on TB incidence (Figure 42A).

For populations with a similar incidence to that of the non-white UK born population, increasing the proportion of infections detected from 3% per year to 13% is predicted to lead to a small reduction in TB incidence, compared to that in the absence of the TB-STS. For example, for this scenario, 20 years after the introduction of the TB-STS, the predicted incidence is 11% less than that in the absence of TB-STS (Figure 42B) (ranging between 4% and 12% for pessimistic and optimistic assumptions, respectively, relating to uptake and completion of preventive treatment), with about one case prevented per 100,000 per year over this period (Figure 42B). Combining an increase in the proportion of infections detected to 13% per year with a one week reduction in diagnostic delay is predicted to approximately double the reduction in TB incidence, with just over two cases prevented per 100,000 per year over the period 2010-2030 (Figure 42D).

Of the scenarios considered, the TB-STS is predicted to lead to the greatest reductions in TB incidence in high transmission settings, similar to the non-UK born population. For this scenario, 20 years after the introduction of the TB-STS, the predicted incidence is about 15% less than that in the absence of TB-STS (Figure 42C), with about 10 cases prevented per 100,000 per year over this period (Figure 42C). When an increase in the proportion of infections detected to 13% per year is combined with a one week reduction in diagnostic delay, the reduction in TB incidence increases to 40% (Figure 42C), with over 30 cases prevented per 100,000 per year over the period 2010-2030 (Figure 42D). Under this scenario, and broadly accounting for the relative proportions of people living in low, medium and high incidence areas in the population (based on primary care trust-level data),[204] we would estimate that the TB-STS would prevent approximately 1400 cases in England per year (although this does not account for the range in the number of cases prevented in a given setting, nor the age distribution, migration rate or HIV prevalence in the different populations).

Figure 42 – Predicted impact of the TB-STS



Predicted impact of reducing diagnostic delay from 12 to 11 weeks and increasing the proportion of infections that are detected by the TB-STS from 3% to 13%. The number of cases occurring per 100,000 per year in a setting in which the tuberculosis incidence is (A) similar to that in the white UK population (declining ARI)), (B) similar to that in the non-white UK born population group (ARI=0.1%/year), (C) similar to that in a high transmission, non-UK born population group (ARI=1%/year), and (D) shows the average annual number of cases prevented per 100,000 population for these scenarios over 20 years after the introduction of the TB-STS.

7.2 The cost-effectiveness of the TB-STS

The structure of the cost-effectiveness model is shown in Figure 25, page 82. Cost of the TB-STS and QALY estimates associated with the TB-STS are shown in Appendix 6.

Transmission model outputs

The results of the transmission model provided estimates for the population of England over a 20-year period (2010-2030) assuming a constant risk of infection of

0.1% per annum for each modelled scenario of: the number of contacts with LTBI identified, the number of contacts starting preventative treatment, the number of people with active TB diagnosed and starting treatment, as well as the impact on the number of incident TB cases (Table 37). The outputs of the transmission model shown in Table 37 were input into the cost-effectiveness model. The baseline (S0) is intended to reflect the expected costs and Outcomes of the TB control system in the absence of the TB-STS. This is modelled assuming that 3% of previously infected individuals are identified per year and that mean diagnostic delay is 12 weeks. A further 14 scenarios are considered, that vary the assumptions for the proportion of LTBI detected between 3% and 13% and diagnostic delay between 8 and 12 weeks. This exercise is intended to assess what impact the TB-STS would have to have in order to be cost-effective.

The cost-effectiveness scenarios of the TB-STS are estimated under the assumptions that it:

a) Increases the proportion of infected cases detected from 3% to 4%, 5%, …, 13% (S1 to S10) with a constant diagnostic delay of 12 weeks.

or

b) Reduces the diagnostic delay from 12 weeks to 11, 10, 9, 8 weeks (S11 to S14) with a fixed LTBI of 3%.

Table 37 – Summary of transmission model results for baseline (S0) and 14 scenarios[a]

Scenario	% LTBI found	DD[b] (weeks)	LTBI diagnosed	LTBI starting treatment	New TB cases[c]	TB cases diagnosed[c]	TB cases starting treatment
S0	3%	12	8,987	8,538	6,672	7,502	6,640
S1	4%	12	11,298	10,733	6,585	7,408	6,556
S2	5%	12	13,334	12,668	6,502	7,316	6,475
S3	6%	12	15,131	14,375	6,420	7,227	6,397
S4	7%	12	16,718	15,882	6,342	7,140	6,321
S5	8%	12	18,120	17,214	6,265	7,056	6,246
S6	9%	12	19,361	18,393	6,191	6,974	6,174
S7	10%	12	20,459	19,436	6,118	6,895	6,104
S8	11%	12	21,432	20,360	6,048	6,817	6,036
S9	12%	12	22,295	21,180	5,979	6,742	5,969
S10	13%	12	23,060	21,907	5,912	6,668	5,904
S11	3%	11	7,835	7,443	4,964	5,631	4,990
S12	3%	10	7,011	6,661	3,828	4,379	3,884
S13	3%	9	6,412	6,091	3,051	3,520	3,125
S14	3%	8	5,964	5,666	2,504	2,912	2,588

[a]Mean number of cases per year for population of England (53m) over 20 years, assuming constant ARI of 0.1%.
[b]DD Diagnostic delay
[c]The estimated number of cases diagnosed exceeds the number of new cases in each year as there is a pool of cases who have previously not been diagnosed or who have defaulted from treatment.

Scenario analysis

The incremental costs and effects of the TB-STS under each scenario are shown in Table 38. Under our base case assumptions, if the TB-STS had increased the proportion of LTBI detected from 3% to 4% with no impact on the mean time to diagnosis for active cases, it would not be cost-effective. Although the improvement would have prevented an estimated 1,726 cases of TB (over 20 years for the population of 53m), saving approximately £3.8m in diagnosis and treatment costs, this cost was more than offset by the direct cost of the TB-STS (£14.3m), the additional costs of screening contacts (£32.5m) and of prophylactic treatment (£22.2m). The net impact on health expenditure was an estimated increase of £65.2m. This cost increase is associated with a QALY gain of around 682 years of healthy life, giving an estimated ICER of £95,628 per QALY gained, which is well above the range usually considered to be cost-effective in the NHS (a maximum of £30,000 per QALY gained).

Estimated cost-effectiveness did improve under the assumption that the TB-STS achieved a greater increase in the proportion of LTBI detected. However, over the range tested this improvement was still not sufficient to bring the ICER below the £30,000 threshold (scenarios S1 to S10 compared to S0, Table 38). If the introduction of the TB-STS has increased the identification of an additional 10% of prevalent LTBI - an additional 281,461 people diagnosed with LTBI over 20 years - the estimated cost per QALY gained was £54,539 (scenario S10 compared to S0, Table 38).

In contrast, the results were very sensitive to small reductions in the average time from onset of symptoms to the start of treatment for active disease. A reduction from 12 weeks to 11 weeks was estimated to yield a large reduction in the number of incident TB cases, and hence in the numbers of contacts to be screened and in people requiring prophylactic treatment (S11 and S12 compared to S0, Table 38). There was therefore a net saving in healthcare expenditure (almost £85m saved), as well as a large health improvement (16,000 QALYs gained). Bigger reductions in the diagnostic delay, would achieve even larger cost savings and health improvements.

Deterministic sensitivity analysis

Results under a range of other changes to the model parameters are shown in Table 39. Unless stated otherwise, these analyses all relate to the comparison between scenarios S1 and S0 (1% increase in the proportion of prevalent LTBI cases diagnosed with TB-STS; no difference in diagnostic delay), and with all other parameters held constant at the base case values.

Other than reductions in diagnostic delay, the only changes tested that gave an estimated ICER below the usual NICE threshold of £30,000 per QALY related to an increase in the QALY loss from TB. However, in order to achieve this result, quite strong assumptions were required about the TB-related mortality and/or morbidity: equivalent to an overall mean loss of two full years of healthy life per case (see row in Table 39 labelled “TB case fatality”).

Table 38 – The incremental costs and effect and the resulting ICERs for the scenarios explored compared to the baseline (S0)[a]

Scenario	% LTBI detected	DD[b]	Contact screening £	Prophylactic treatment £ (QALYs)	TB diagnosis & treatment £ (QALYs)	Total incremental cost[c] £	Total incremental effect (QALYs)	**ICER[d] (£ per QALY gained)**
S1	4	12	32,539,484	22,240,304 (-27.70)	-3,809,472 (707.27)	65,219,944	682	**95,628**
S2	5	12	61,643,979	42,132,714 (-52.48)	-7,511,038 (1,394.66)	110,515,283	1,345	**82,190**
S3	6	12	87,710,528	59,948,754 (-74.68)	-11,110,571 (2,063.18)	150,798,339	1,991	**75,742**
S4	7	12	111,087,030	75,926,436 (-94.58)	-14,613,304 (2,713.84)	186,649,790	2,622	**71,194**
S5	8	12	132,081,043	90,275,317 (-112.45)	-18,024,077 (3,347.55)	218,581,910	3,238	**67,515**
S6	9	12	150,961,412	103,179,853 (-128.53)	-21,347,548 (3,965.19)	247,043,344	3,839	**64,349**
S7	10	12	167,965,309	114,801,926 (-143.01)	-24,587,951 (4,657.47)	272,428,913	4,427	**61,539**
S8	11	12	183,301,973	125,284,356 (-156.06)	-27,749,123 (5,155.15)	295,086,833	5,002	**59,999**
S9	12	12	197,154,669	134,752,505 (-167.86)	-30,834,794 (5,728.84)	315,322,008	5,563	**56,678**
S10	13	12	209,685,394	143,317,116 (-178.53)	-33,848,402 (6,289.25)	333,403,736	6,113	**54,539**
S11	3	11	-13,640,543	-9323,117 (11.61)	-76,153,948 (16,032.60)	-84,867,979	16,047	*****
S12	3	10	-23,552,487	-16,097,775 (20.05)	-128,625,846 (26,950.25)	-154,026,480	26,973	*****

[a]Baseline scenario (S0) assumes 3% LTBI detected and 12 weeks diagnostic delay
[b]DD Diagnostic delay
[c]Total incremental cost includes the incremental cost of running the TB-STS (£14,298,781) and the costs saved from the identification of false positive TB diagnoses (-£49,153) and the associated QALY gain (2.45)
[d]ICER = Incremental cost of the TB-STS(£) / Incremental effect of the TB-STS (in QALYs). The NICE threshold of cost-effectiveness is £20,000-30,000 per QALY gained
*The costs saved and the QALYs gained outweigh the additional costs associated with the TB-STS in these scenarios

Table 39 – Results of the deterministic sensitivity analysis

Parameter changed (base case value)	Parameter values tested	Incremental cost (£)	Incremental effect (QALYs)	ICER (£ per QALY)
% LTBI detected	+1% with TB-STS (base case)	£65,219,944	682	£95,628
	+10% with TB-STS	£333,403,736	6,113	£54,539
Diagnostic delay	-1 week with TB-STS	-£64,867,979	16,047	Dominant
(no reduction)	-4 weeks with TB-STS	-£239,774,497	40,078	Dominant
Discount rates	1.5% for QALYs	£65,219,944	1,093	£59,682
(3.5% QALYs, 3.5% costs)	0% for QALYs	£65,219,944	1,643	£39,707
Time horizon	15 years	£55,453,941	431	£128,807
(20 years)	10 years	£42,475,021	212	£200,212
Cost of software	£500,000	£65,622,236	682	£96,218
(£264,593)	£1,000,000	£66,476,695	682	£97,471
Cost to HPUs	£50,000 pa	£64,289,459	682	£94,234
(£113,256 pa)	£500,000 pa	£70,908,886	682	£103,970
False positives	10 cases avoided pa	£65,170,792	684	£95,214
(5 avoiding treatment pa)	100 cases avoided pa	£64,286,046	729	£88,233
Utility loss from FP	0.5 for 4 months	£65,219,944	692	£94,273
(basecase 0.1 for 4 months)	0.5 for 12 months	£65,121,639	716	£90,909
Contacts screened (basecase 4.11)	2 contacts per LTBI diagnosed	£49,073,148	682	£71,953
	6 contacts per LTBI diagnosed	£81,858,521	682	£120,025
Adverse effects of CPx	0%	£65,219,944	710	£91,896
(basecase 10%)	100%	£65,219,944	433	£150,737
Yield of TB diagnosis	2 per case	£66,844,254	682	£98,010
(5 investigated per case)	10 per case	£62,512,761	682	£91,659
TB case fatality	1% all ages	£65,219,944	330	£197,537
(0.3% 0-4 to 17.6% 55+)	10% all ages	£65,219,944	2,249	£28,995
TB morbidity	0.05 QALYs lost per case	£65,219,944	599	£108,832
(0.12 QALYs lost per case)	1 QALY lost per case	£65,219,944	1,684	£38,732
Cost of prophylaxis	£300 per case	£51,955,851	682	£77,180
(£743 per case)	£1,000 per case	£72,900,343	682	£106,890
Cost of treatment	£500 per case	£65,854,499	682	£96,559
(£1,114 per case)	£5,000 per case	£61,202,523	682	£89,738

7.3 Discussion

7.3.1 Strengths and limitations of the methods used to evaluate the Outcomes of the TB-STS

The transmission model extends previous models considering the transmission dynamics of *M. tuberculosis*.[112,208] The model takes into account the variation in transmission dynamics across age, migration patterns, and the use and effectiveness of preventative therapy. It is integrated with the rest of this evaluation, using data from the investigations into contact tracing yield and diagnostic delay (Figure 26, page 86). In addition, it inputs into the cost-effectiveness model, allowing for the effectiveness and cost-effectiveness of the TB-STS to be estimated over a 20-year time horizon. Likewise, the cost-effectiveness model is integrated into the rest of the evaluation. As well as being directly linked to the transmission model, it derives parameters from the cross-sectional surveys, contact tracing yield and diagnostic delay analyses, and the false positive isolation survey (Table 36, page 172).

The outputs of the transmission model suggest that the impact of the TB-STS is different across the three epidemiological scenarios, most notably having very little impact in the white UK-born population. These differences result from the model assumptions; specifically those around the assumptions underlying the scenarios considered and PT. Firstly, the assumptions regarding the proportion of LTBI and TB in the population across different age groups, and the proportions due to reactivation versus recent transmission are important in understanding the transmission model outputs (Figure 41). Secondly, no one aged ≥35 years and 95% of those aged <35 years who are identified as having been infected are assumed to start PT,[219] with 85% of these completing the full course of treatment. PT is assumed to provide 65% protection against disease whilst individuals are taking the treatment[220,221] and the full course of PT is assumed to fully cure the infection so that individuals can only develop disease subsequently if they are reinfected.

Considering these assumptions, the model estimated that the TB-STS would have minimal impact in the white UK-born population. This is because the proportion of

individuals who have been infected is low (<10%) for those aged <55 years, but increases to 50% for those aged ≥55 years (Figure 41B). Therefore the impact of identifying an increased proportion of LTBI will not have an impact on TB incidence because those being identified would not be eligible for PT. In addition, because the predicted TB incidence is highest amongst those aged ≥55 years (Figure 41A), and the proportion of disease that is attributable to recent transmission decreases steadily with increasing age, reaching <10% for those aged ≥55 years, the impact of reducing diagnostic delay in the white UK-born population is smaller compared with that in the alternative epidemiological scenarios where the proportion of TB attributable to recent transmission is correspondingly greater. Conversely, the non-white UK-born and non-UK-born scenarios assume that a greater proportion of LTBI is in those aged <35 years (Figure 41B), compared to that for the white UK-born scenario, so can be given PT. This means that, due to a large proportion of TB in these scenarios being in younger age groups (Figure 41A), the proportion of TB attributable to recent transmission in the overall population is >90%, which, in turn, leads to an increased impact of increasing the proportion of LTBI detected and reducing diagnostic delay on TB transmission in these populations.

There are several limitations to the transmission model. An important limitation is that there are no data on the percentage of all infections in the population that were detected before the introduction of the TB-STS and it is therefore assumed that this equalled 3%. Large screening studies may help to calculate this percentage. A further limitation is that there are no data on the ARI in the population, which is likely to vary between communities and different parts of the country. We have also assumed that preventative treatment cures infections. The extent to which this is realistic is presently unclear[208] and, if preventative treatment does not cure infections, we may have overestimated the number of cases prevented through detecting a given proportion of infections. The model did not include HIV prevalence or drug resistance because the prevalence of HIV and drug resistance is relatively low in England[3] and to limit the complexity of the model.

This analysis failed to demonstrate that the TB-STS is a cost-effective use of

resources. It suggests that it is unlikely that earlier identification of false positive cases related to laboratory contamination, or increases in the identification and prophylactic treatment of contacts with LTBI could, on their own, justify the cost of the system. It was not possible to conduct a probabilistic sensitivity analysis to characterise the overall impact of uncertainty over the parameters and assumptions over the transmission model and cost-effectiveness analysis. This is because of time and resource restraints and because of the high level of uncertainty around the model parameters in the first place. However, simple deterministic sensitivity analysis suggested that the results are, with one major exception, quite robust to plausible changes in most parameters. The key uncertainty relates to the lack of evidence over whether the TB-STS is associated with earlier diagnosis and treatment for active cases.

Although the proportion of LTBI detected is not known, the possibility of the TB-STS increasing this by 1% seems more likely than decreasing the diagnostic delay by one week. This is because diagnostic delay is driven by the passive presentation of people with TB due to recent transmission and reactivation and those identified through active case finding, and the TB-STS can only have an impact on those identified through the latter route. The proportion of prevalent LTBI detected increases as more people are screened. The non-significant increase in the mean number of contacts identified and screened per index case between the initial and follow-up surveys (Table 25, page 146) suggests that the TB-STS might increase the number of contacts screened, so could plausibly increase the proportion of LTBI detected. In addition, if the TB-STS were to improve contact tracing practices (as postulated by HPU strain typing leads in the semi-structured interviews), an increase in the number of contacts screened would be expected.[222]

The finding that the service is not cost-effective is consistent with evidence from the Birmingham and Bradford areas. The two studies demonstrated the high workload associated with cluster investigations.[223,224] One of the investigations presented did not identify any further transmission of TB, but required an estimated 20 hours of

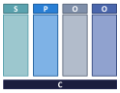
work. If a full cost-effectiveness analysis had been conducted in that setting, it seems unlikely that it would have been found to be cost-effective.

7.3.2 Summary of findings

The mathematical modelling suggested that increasing the proportion of infections detected would have little value in reducing TB incidence in the white UK born population. However, in the non-white UK born and non-UK born populations, over 20 years, if detection of LTBI increases from 3% to 13% per year then TB incidence would decrease by 11%; reducing diagnostic delay by one week could lead to 25% reduction in incidence.

However, even assuming the TB-STS leads to a small (1%) improvement in the contact tracing yield in a medium-incidence population (similar incidence to that of the non-white UK born population), the service was not estimated to be cost-effective over a 20-year period: £95,628 per quality adjusted life year (QALY) if the proportion of LTBI detected were to increase from 3% to 4%. Assuming an increase from 3% to 13%, the estimated incremental cost per QALY gained (£54,539) still did not reach a level considered to be cost-effective in the UK. These results were much more sensitive to reductions in the diagnostic delay: e.g. if the TB-STS reduced diagnostic delay by one week, the system would save £85m and gain over 16,000 QALYs over 20 years.

Chapter 8. Discussion



The discussion aims to contextualise the findings of this evaluation. In this chapter I summarise and evaluate the main findings presented in this thesis. This chapter discusses the strengths and limitations of the research as a whole, outlines recommendations for implementing and evaluating a public health intervention on a national scale, and suggests directions for future research. The strengths and limitations of each sub-study and how the findings relate to previous studies have been discussed at the end of each chapter (Structures: page 101; Processes: page 128; Outputs: page 157; Outcomes: page 181).

—

8.1 Summary of thesis

The TB-STS involves MIRU-VNTR typing of the first isolate from every TB patient for the prospective identification, reporting and investigation of strain typing clusters, to be used in real time to inform public health action.

A mixed-method evaluation of the TB-STS in England was conducted between January 2010 and April 2013. Using an adapted Donabedian evaluation framework, the Structures were described, the Processes investigated, the Outputs measured, and the Outcomes estimated. The Context of the TB-STS is explored here. The methods used and the main findings are summarised in Table 40.

Table 40 – Table of objective, study, main findings, publication

Objective	Methods	Main findings	Page	Pub.
1. To examine the effect of study design and setting on the estimation of the proportion of clustering by MIRU-VNTR strain typing	Systematic review of the literature	The number of MIRU-VNTR loci typed, requiring consent to type patient isolates (as a proxy for sampling fraction), the TB incidence and the maximum cluster size had a significant association with the proportion of clustering.	40	[94]
2. To describe the Structures of the TB-STS	Description of the service based on grey literature, and conversations with stakeholders/key informants	The TB-STS is a complex intervention with multiple interacting components, relying on different elements of the public health and national health services in England to communicate and cooperate. Crucially, the software to link the services together was not completed within the evaluation time period. Interviews found people had mixed experiences, but identified broader benefits, of the TB-STS.	87	
3. To evaluate the Processes of the TB-STS	Initial and follow-up cross-sectional surveys and semi-structured interviews with health protection staff	The TB-STS had been integrated into the TB services across England by March 2012, but perceived usefulness of the service did not change. There was improved reported knowledge of strain typing and increased frequency of use. The majority of health protection staff received the strain typing after they had completed contact tracing activities and the proportion of time spent on investigating TB transmission increased. Interviews explored mixed experiences of the TB-STS, but identified broader benefits of the TB-STS such as the overall strengthening of information collection for contact tracing and its contribution to a greater understanding of TB epidemiology.	103	[202]
4. To evaluate the Outputs of the TB-STS	Analysis of cluster investigations, contact tracing yield, diagnostic delay, rate of cluster growth, and false positive TB isolation.	Between 2010 and 2012, 17,168 isolates were typed, 84.6% of which had at least 23 completed MIRU-VNTR loci. The TB-STS identified 17 additional false positive TB diagnoses. The direct outputs and outcomes of cluster investigations were not well captured so could not be used to inform the evaluation. Indirect outputs of cluster investigations were explored instead. The TB-STS had no significant effect on contact tracing yield, diagnostic delay or the rate of cluster growth.	111	[225]
5. To evaluate the Outcomes of the TB-STS	Deterministic TB transmission model and cost-effectiveness model	The mathematical model estimated that the TB-STS would have little value in reducing TB incidence in low incidence settings and, of the scenarios explored, the greatest potential impact is estimated in high incidence settings if the proportion of LTBI that are detected increases to 13% and diagnostic delay reduces by one week. The TB-STS was not estimated to be cost-effective over 20 years (£95,628/QALY).	172	[225]
6. To contextualise the findings of this evaluation	Discussion	The TB-STS is not effective or cost-effective, but broader benefits that were not accounted for in the modelling justify its continuation.	185	

Pub. Publication resulting from this thesis

8.2 The wider context of the TB-STS

The scientific, economic and technological context of the TB-STS can be considered in terms of:

1. The increase in funding for the application of molecular biology to TB epidemiology and public health;
2. Global economic downturn leading to austerity measures; and
3. The political will to utilise novel molecular techniques to improve public health.

During the evaluation period, there has been increasing interest in the application of molecular biology to epidemiology and public health. In 2009 the UK Clinical Research Consortium awarded £5.1m to the Modernising Medical Microbiology Consortium in Oxford to investigate the application of WGS to infectious disease surveillance and molecular epidemiology – *M. tuberculosis* was one pathogen of interest.[85] The Oxford Consortium includes the Oxford Biomedical Research Centre, Wellcome Trust Sanger Institute in Cambridge and PHE, involving PHE in the forefront of molecular epidemiology and infectious disease surveillance. In addition two Health Innovation Challenge Fund grants have been awarded by the Wellcome Trust to Oxford and Cambridge Universities to translate WGS into routine clinical microbiological practice and to develop a world-class system of active surveillance.[227] In 2013, the UK Prime Minister pledged £100 million to the 100k Genome Project, which includes the whole genome sequencing of *M. tuberculosis*.[228] This increase in funding for the application of molecular biology to TB epidemiology and public health reflects the scientific and funding communities' interest in the area and their incentive to make it work.

Following an economic downturn at the end of the 2000s, the UK entered a period of 'austerity', whereby spending cuts were applied across the Government and quasi autonomous non-governmental organisations (quangos) were dismantled and either shut down or moved into Government. The Health Protection Agency was a quango and following large budget cuts from 2010 onwards, in April 2013 it became Public

Health England, part of the Government's Department of Health. This had implications for the resources made available to the TB-STS and hence impacted on its implementation. While new resources were being invested in developing WGS for TB, substantial reductions were made to the public health infrastructure. This has limited PHE's ability to utilise and evaluate information generated by the TB-STS and more recently WGS, through the investigation of potential outbreaks – one of the main justifications for such a service.

A strong political will to utilise novel molecular techniques to protect and improve public health is evident in England. Despite the findings of this evaluation, PHE are exploring the use of universal prospective WGS for front-line TB control and are piloting such a service in the Midlands (*Prof Ibrahim Abubakar, personal communication*). Although an evaluation of the service is planned, the results will not be made publicly available. It is clear that WGS has some advantages over a MIRU-VNTR typing service,[50,78,82] but it is currently more expensive so its cost-effectiveness is uncertain.

8.3 Evaluation framework and design

As a result of the complexity of national public health interventions or policies, many are not evaluated. Or, if they are, they do not take into account the entirety of the intervention and reasons why or how it might or might not work. Instead, they may focus on the easy (or easier) to measure outcomes. This leaves policy makers vulnerable to the enthusiasm around new technologies and the idea that because this new technology is useful in some settings, it will be useful across the country. It is, therefore, commendable that PHE included a prospective evaluation in the development of the TB-STS.

The framework adopted for the evaluation of the TB-STS was an adapted version of the Donabedian healthcare evaluation framework,[161] designed to describe and evaluate the entirety of the service. The Donabedian framework is a simple evaluation model that divides the service into Structures, Processes, Outputs and

Outcomes. This ensures that all aspects of the service are taken into account, from the service design, to its day-to-day functioning, to its overall impact. The framework prevents the evaluator from focussing on just one element of the service, and makes sure that the parts of the service that can be adjusted and changed are included in the evaluation. As a result, recommendations have been made to PHE about how to improve the service that involve structural and process-driven changes that are easily implemented, rather than introducing only output and outcome-related targets (Box 6, page 204).

An additional advantage of the Donabedian framework is that it helps to identify how to assess each component and how the results of each component might feed into another. In this evaluation, the programme cost of the TB-STS was determined from the description of the Structures, results of the cross-sectional surveys were used to develop the semi-structured interviews and parameterise the cost-effectiveness model, the false positive TB isolation results were input into the cost-effectiveness model, the analysis of contact tracing yield and diagnostic delay were used in the transmission model, which input into the cost-effectiveness model (Figure 26, page 86).

The Donabedian model was adapted by including an overarching component which acknowledges the context of the TB-STS. This discusses the overarching benefits and costs of the TB-STS that may not have been captured by the quantitative evaluation methods.

The evaluation framework applied here also has limitations. For one, some might argue that it is too simplistic. There is a wealth of evaluation literature prescribing different approaches to evaluation and whilst these can be helpful in challenging a researcher to think through their model of evaluation and what questions are being asked, the evaluation theories are often more than approaches, but methodologies, thereby being prescriptive about how they should be done.

This evaluation was commissioned by PHE (then the Health Protection Agency) to inform their TB strain typing strategy. Although PHE can be praised for their inclusion of the evaluation into the TB-STS structure, the timing of, and budget for the evaluation restricted the evaluation design options. In addition, PHE employees were helpful and compliant with the data collection conducted for the evaluation and the gathering of information about the service, but further involvement of PHE staff would not have been appropriate. On the other hand, NHS staff were not required to support the evaluation, so any information from them was voluntary. The Donabedian framework did not limit the evaluation to a particular methodology and, as a result, the design of the evaluation could reflect these constraints.

Ideally, the evaluation would have been either a clustered trial comparing sites that were randomly allocated the TB-STS to those that did not have the TB-STS, or a controlled before and after design where the baseline before the service could be compared to a time after the service had been implemented.[229] However, due to the timing of the evaluation, the national implementation of the service, and the limited availability of baseline data, these evaluation designs could not be used. Therefore, a mixed-methods approach was used to evaluate this service, including qualitative, quantitative, direct and indirect methods. A mixed methods approach, as opposed to randomised trial, can be the most appropriate way of evaluating a national program.[154]

This flexible and responsive approach to evaluation is shared by Habicht and colleagues[230] who argue that the evaluation design will depend on who the evaluation is for (who the decision-maker is) and what types of decisions will be made as a consequence of the evaluation. Their framework is based on two axes that concern the indicators of interest (provision, utilisation, coverage or impact of the intervention) and the type of inference to be made (adequacy, plausibility or probability). In the context of the TB-STS, which was evaluated to inform PHE's strain typing strategy in England, the lack of adequate data have resulted in recommendations based on both probability and plausibility.

A recent evaluation of an intervention to find and treat hard to reach TB patients in London is an example of how a pragmatic evaluation of a complex intervention can influence policy.[11] The 'find and treat' evaluation could not be designed as a randomised trial. Despite this, the findings of the evaluation were used to influence policy on targeted active case finding for TB.[231]

8.4 Evaluating the Processes

Evaluating the Processes can provide the insight and context necessary to interpret the Outputs of a service (e.g. has the service been implemented properly? Has the service been implemented differently in different places?), help to explain differences between the observed and expected Outcomes, and identify ways of intervening to improve the service.[154] The semi-structured interviews helped to better understand how the TB-STS was implemented; how the TB-STS had been implemented in different HPUs and how health protection staff used the service. Although many limitations of the TB-STS were raised, the interviews also elicited support for the service.

The implementation of the service, supporting the interpretation of the outputs, the identification of ways of improving the service, and the broader benefits of the service that were identified through the interviews are now discussed in turn.

Implementation

Evidence for the successful implementation of the laboratory elements of the TB-STS (such as the typing of isolates and reporting of strain typing results) can be found in the laboratory Outputs (page 135). The implementation of the public health component of the TB-STS is more complex, as was identified by evaluating the Processes: findings from the surveys and the interviews suggest that some of the public health elements of the TB-STS (such as the preliminary cluster reviews and the dissemination of cluster reports) were well implemented, but other elements were less well implemented (such as the reach of the training provided about the TB-STS and support for HPUs to conduct cluster investigations). These findings impact on

the way the public health Outputs of the TB-STS were interpreted, and helped to identify areas that could be improved within the service.

The lack of a functioning STM raises concerns of type III error. The failure to develop the STM within the evaluation period resulted in insufficient data collection and data quality, as well as creating the problem of evaluating a service that is not properly implemented. However, whilst there would have been benefits attained with the release of the STM in the standardisation of reporting and data access, the extent to which a functioning STM would impact on the Outputs of the service and the consequent impact on the effectiveness of the service is unknown.

Taking into account the interim systems that were implemented to perform the functions of the STM, such as the cluster reports generated by the laboratories and CIs, and the alternative data collection and storage methods developed (see Table 8, page 92 for more detail), one cannot conclude that the introduction of the STM would have a profound impact on the functioning of the service. In addition, it is not clear that the STM would have a significant impact on the diagnostic delay – a key driver for the effectiveness of the TB-STS – for the following reasons:

1. The analysis and reporting of clusters based on the criteria in the strain typing Handbook for local and national clusters happened promptly and was sent to HPUs and clinics regularly. However, it would have been useful to reconcile this information with other data in the surveillance system using the STM, improving the timeliness of the cluster reports by removing the need for CIs to collate the information.

2. A major issue identified by the semi-structured interviews with HPUs was that by the time isolates were cultured and subsequently typed prospectively, most HPUs already knew about the links between cases that they considered to be important. Therefore, the strain typing arrived too late not because the laboratories were not undertaking the software based analysis promptly due

to the lack of the STM, but because the process of getting to the fingerprint takes too long.

3. One area where the STM would have been particularly useful (even if it did not inform the investigation of clusters) is in providing the data to demonstrate the effect of the TB-STS on the diagnosis of additional cases in clusters. Again, for many clusters there was an attempt to collect the information that would have gone into STM using paper forms but the response rate was very variable. There is no reason why this information would be more complete using the STM.

4. In addition, one can speculate that providing the software is one step, but getting people to use it is another step entirely. PHE's experience of developing and launching the ETS tells us that it can be very difficult to get people working across PHE and the NHS to adopt a new software system – the ETS was introduced in 1999 and had been adopted electronically by the entire country by 2010 (*personal communication with Dr Laura Anderson from PHE, 2010*).

Supporting information to help interpret findings

The interview findings can be integrated to help better understand and interpret the results of the rest of the evaluation. The semi-structured interviews were used to find out more about what a cluster investigation involves in each HPU and how strain typing influenced decisions. It was identified that HPUs used the strain typing differently and received different support from the CIs, thereby helping to explain the variation in the implementation of the TB-STS and helping to elucidate the survey findings around workload and cluster investigation activities. For example, the increasing cluster investigation workload for HPU staff without any increase in the number of epidemiological links found between clustered cases may reflect the lack of resources available for clinic staff to conduct further contact tracing around an historical patient. Alternatively, this may be because the criteria used to initiate a cluster investigation are not appropriate – importantly, the semi-structured interviews

identified that HPUs did not adhere to the thresholds stated in the guidelines, but adapted them to their local situation to make them more feasible given local case load and the resources available.

The cross-sectional surveys suggested that the strain typing was usually received after contact tracing had already been completed. Consistent with this, many interviewees commented that the strain typing was often received too late to have an impact on their decision-making. In some instances, this was reassuring as the strain typing confirmed what was already suspected. However, this was viewed by some as frustrating as the strain typing provided no additional contribution to TB control. In conclusion, the hypothesis that the TB-STS would contribute towards contact investigations in real time was not plausible.

Identifying ways to improve the service

Firstly, the surveys identified gaps in the training strategy. Nurses were not included in the training strategy, but the surveys found that nurses had the lowest self-rated knowledge of strain typing. In addition, interviews identified that some HPUs used the strain typing as a way of engaging nurses in TB control, suggesting that, since the service relied on nurses to collect more information for cluster investigations, including nurses in the training strategy could help improve the functioning of the service.

Secondly, the interviews identified a lack of resource for the public health implementation of the TB-STS. The problem with the design of the TB-STS was that it did not account for the input required from TB nurses in a cluster investigation, so no resources were allocated to support them. Alongside this, the introduction of the TB-STS coincided with the period of austerity in the UK following the global economic downturn. As a result, PHE had to make significant savings, and the public health side of the intervention was altered by removing two permanent CI positions (who had not yet been hired). This resulted in the remaining CIs having too great a workload, therefore not being able to support the HPUs sufficiently.

The perceived lack of resources for the public health component of the TB-STS was a key theme from the interviews that was not captured by the other evaluation methods. This finding goes some way to explain why there was no measurable impact of the cluster investigations – if there were no resources available to investigate clusters then the impact of the investigations would be negligible. As the economic instability and political context dictate that public funds are coming under increasing pressure, additional new resources for the TB-STS are unlikely (unless it is shown to be highly cost-effective, which is not the case). Therefore, cluster investigations are unlikely to become effective (or cost-effective).

Thirdly, the interviews provided a platform to collect suggestions for service improvement from the HPUs, the main users of the TB-STS. The suggestions were helpful in further understanding the current problems with the service, as well as assisting with the development of the recommendations for the future of the service. For example, HPUs suggested that reducing the number of clusters requiring additional investigation might improve their experience of the TB-STS. This suggestion, along with the finding that more investigations were being conducted but no impact on TB transmission could be detected, led to the development of a recommendation that routine cluster investigations should be discontinued (Box 6, page 204).

Broader benefits of the service

The interviews identified broader benefits of the service that had not been captured by the other evaluation studies. For example, the improved data collection by nurses, the move towards standardised data collection for contact tracing in some areas, the increased engagement of nurses with the public health aspects of TB (as opposed to just the clinical aspects), the improved relationships between HPUs and TB nurses, and the improved understanding of TB epidemiology in England. In addition, the value of the strain typing information for TB research and its ability to contribute to our understanding of TB in England and worldwide was acknowledged. Interestingly, even though many people had not experienced any benefit of the TB-STS to public health, or were frustrated by the additional workload associated with it,

some people felt that it would be 'regressive' to stop the service. This finding has important implications for the introduction of other national interventions, especially where the effectiveness and cost-effectiveness is not formally evaluated; it is difficult to justify taking something away if there is any perceived benefit, even if the gain is small. The broader benefits of the TB-STS are discussed further on page 201.

8.5 Evaluating the Outputs

Traditional evaluations measure the outputs of an intervention, as they are usually measurable indicators of a programme's impact. The false positive TB isolation identified by the TB-STS is an example of where this is true – to measure the number of false positive isolates identified, data were collected from the reference laboratories on the number of incidents queried and source laboratories were surveyed to establish the outcome of the query and whether the incident had been previously identified. However, because the platform for cluster reporting and data synthesis was not available during the evaluation, collecting information on cluster investigations and their outcomes proved difficult. One way around this was to develop interim data collection strategies such as cluster outcome forms and a cluster monitoring database. However, there were no additional resources allocated to enter the data onto these platforms, so the data collected were sparse. Therefore, consistent with the MRC framework,[154] indirect Outputs of the TB-STS were investigated based on the hypothesised impact of the service: contact tracing yield, diagnostic delay, and rate of cluster growth were all analysed (Figure 11, page 31). There was no evidence to suggest that the TB-STS had an effect on the indirect Outputs of the TB-STS.

The approach of this evaluation was pragmatic; using the best quality data available to find a different way of measuring the Outputs of the TB-STS. This is an example of how the evaluation design can be responsive and flexible to the system it is evaluating: by working with the intervention stakeholders to design and implement interim solutions, and find alternative indirect ways to measure the impact of the intervention. With this, however, come some considerable limitations.

There is uncertainty surrounding the conclusions drawn from these studies measuring the indirect Outputs of the TB-STS. The studies did not have the most rigorous control or comparison group, and were reliant on the limited amount of data collected since the initiation of the service in 2010 (which was further reduced to data from 2011 to 2012 because 2010 was considered the 'roll out' year and the evaluation period ended at the end of 2012). As a result, it cannot be concluded that the TB-STS did not have an impact on contact tracing yield, diagnostic delay or rate of cluster growth; rather, that there is no evidence to suggest an impact. The potential variability in the estimate of effect on contact tracing yield and diagnostic delay was explored using a deterministic sensitivity analysis in the mathematical and cost-effectiveness models.

The indirect methods adopted are useful indicators for assessing the quality of a TB control programme. Specifically, contact tracing yield is an indication of how effective active case finding is, and diagnostic delay is an indicator of the opportunity for onward transmission in the community. The rate of cluster growth is not widely used as so few countries have a universal typing service, but may be a useful indicator for rate of transmission. Though not direct Outputs of cluster investigations, if the TB-STS is helping to interrupt transmission, an increase in contact tracing yield, a decrease in diagnostic delay and a reduction in the rate of cluster growth should eventually be detectable. The studies presented here could provide a useful comparison for a future study after the TB-STS matures.

As one of the drivers for setting up the TB-STS, it is surprising that the TB-STS did not identify much additional false positive TB. This suggests good quality practices throughout the laboratory service. If the TB-STS had been introduced to a laboratory service with less rigorous quality control procedures, or a service that had not previously conducted much strain typing, it may have had a greater impact. The experience of conducting strain typing in a UK laboratory since 2003 is likely to have resulted in improved systems across all three laboratories. Using strain typing to confirm or refute suspected false positive TB identified through another channel such as clinical suspicion, might be of more value as it would prevent clusters of false

positive patients from being investigated. However, this evaluation did not capture those incidents, as it was focussed on the false positive isolation suspected because of strain typing. A future evaluation might consider including this in their assessment of the TB-STS.

8.6 Evaluating the Outcomes

The inclusion of a transmission model and a cost-effectiveness model to this evaluation makes the results of the evaluation useful to the decision-makers at PHE as well as other countries within the UK and internationally, where the introduction of similar services is under consideration. It is also consistent with the MRC framework for evaluating complex interventions.[154] Using the effect of the TB-STS on TB transmission during the first three years of the service, the models have enabled us to estimate its impact over the next 20 years and to postulate whether it is a cost-effective use of public funds.

A benefit of developing these models is that they have identified elements of the TB control strategy in England – the proportion of infections detected and the diagnostic delay, for example – that, if improved, will have a measurable impact on TB incidence over the next 20 years. This gives policy makers a clear objective for improving TB control in England, irrespective of the TB-STS. An additional strength of the models is their integration with each other and the rest of the evaluation (Figure 26, page 86).

The transmission model and the cost-effectiveness model are limited by a general lack of available data (e.g. the proportion of LTBI currently detected, the annual risk of infection in different population groups, and the effectiveness of preventative therapy are unknown). Simple deterministic sensitivity analyses were used to explore the impact of the TB-STS over a 20 year period. It was not possible to conduct a probabilistic sensitivity analysis to characterise the overall impact of uncertainty over the parameters and assumptions over the transmission model and cost-effectiveness analysis due to the level of uncertainty already in the models.

More data are needed to determine whether an increase in the proportion of LTBI detected is possible. It will likely depend on the quality and completeness of the contact tracing that is conducted locally. However, when considering how to spend limited public funds for TB control, the screening of migrants from high incidence countries has been shown to be cost-effective[232] and is likely to have a greater impact on the proportion of LTBI detected than the TB-STS.

Assuming that about 15% of all culture positive cases are in a cluster that is investigated (~50% clustering over all (Table 19, page 137) and ~30% of clustered cases are in a cluster that is investigated (Table 29, page 151)), in theory, the average diagnostic delay of all TB cases could be reduced by reducing the diagnostic delay among cases that were detected as a result of cluster investigations. However, a large reduction would be necessary to have an impact on the overall average. One could argue that a reduction of up to six weeks may be plausible given that information from the TB-STS would not be available to the local team until about four weeks after diagnosis of the index case (

Table 13, page 108). By that time the initial examination of the domestic contacts is unlikely to have been completed so cluster information could be used to decide where else to look for contacts, which would take at least another two weeks. To date, there are no data on the proportion of cases that are detected through cluster investigations or passive presentation. However, in the Netherlands, less than 10% of cases are detected through cluster investigations,[122] which would suggest that a reduction in the diagnostic delay of more than 10% (i.e. one week) is unlikely to have been achieved in the UK.

No other strain typing services in the world have been assessed for their cost-effectiveness, making this the first of its kind. This might be because such a service is unlikely to be estimated cost-effective (because the impact is likely to be either very small and/or very difficult to measure), but the public health and research communities rely on such services for surveillance and data collection. There is also a competitive prestige associated with being at the forefront of typing across the

globe, and once the infrastructure has been developed in order to implement the service, it is unlikely to be removed.

Based on the findings of the transmission and cost-effectiveness models, the TB-STS was neither effective nor cost-effective and therefore should not be continued. This is especially true where there is an opportunity cost to any public health strategy (money spent on the TB-STS cannot be spent on anything else), and where other TB control interventions have been shown to be cost-effective.[11,232] However, it has been argued that cost-effectiveness alone should not be the determinant of a health investment.[233,234] Consistent with this view and the evaluation framework used in this thesis, these findings are discussed in the following sections of this chapter alongside the broader benefits identified in the semi-structured interviews with service users, and the political and research contexts.

8.7 Lack of impact or lack of evidence?

The lack of observed impact of the TB-STS may be due to:

1. A true lack of impact;
2. The inability to observe the impact within the observation period; or
3. The (limited) implementation of the service.

Firstly, there may have been no impact, even if the TB-STS had been fully implemented. Strain typing information may not reach TB service staff early enough to inform contact screening decisions in a meaningful way. Possible reasons for this include the time needed to produce a typing result and/or the lack of sufficiently sophisticated reporting software. Furthermore, inadequate resources to act may be contributing – the generation of the strain typing information was well-resourced but no funds were allocated to local TB teams to embed the information in their decision processes.

Secondly, we are evaluating the marginal impact of the TB-STS in a setting where a TB control programme already exists (which includes traditional TB control

strategies, such as stone in the pond contact tracing).[13,235] Measuring the impact of the TB-STS in isolation may therefore not be feasible and, even if we could, the improvement we are trying to observe may be incremental. In addition, the three-year evaluation period may have been too short for an impact to be observed given the delay between cause and effect in a complex system.[236]

Thirdly, there was limited implementation of the service due to delayed development of the STM as part of a well-integrated TB management system to capture linked cases, contacts and strains; and limited resources for local public health action. This also influenced the ability to evaluate the TB-STS, where suboptimal data collection systems meant that although some primary data collection was conducted, the evaluation was reliant on routine data sources to estimate model parameters.

8.8 The broader benefits of the TB-STS

There was no evidence to support the hypothesis that the current TB-STS would interrupt transmission – the TB-STS was not found to be effective or cost-effective in its current form – but the recommendations to PHE were not that the TB-STS should be discontinued, rather that it continues with substantial changes. This is because the qualitative data and the context of the service suggested that there were broader benefits to the service.

The broader benefits of the TB-STS are three-fold. Firstly, benefits to the TB service, such as better engagement with TB nurses and improved contact tracing practices, as well as the sense that it would be "regressive" to take this new tool away from public health practitioners. Secondly, there is the value of a national dataset combining clinical, epidemiological and molecular information for each TB patient. This will lead to the possibility of multiple future analyses, TB programme evaluations and research projects. Thirdly, the outcomes of such projects will lead to the far-reaching benefits of increased understanding about TB epidemiology, lineage, transmission and control. These benefits, however, may not require prospective strain typing, but could be gained through retrospective strain typing. Reading beyond the results of an

economic evaluation is an approach argued for recently: that cost-effectiveness alone should not be the determinant of health investment.[233]

Box 5 – Examples of the broader benefits of the TB-STS

- Higher quality, more standardised information collection during initial contact tracing of cases (see section on "Information Gathering" under User experience: Results of the semi-structured interviews, page 111)

- To understand the national and local epidemiology of TB: TB strain-typing in New York has enabled molecular epidemiological analyses to contribute to understanding of the TB epidemic and they have been able to tailor their public health response, especially amongst those with HIV or MDR-TB.[224] National strain-typing in the Netherlands has contributed greatly to their understanding of TB epidemiology.[82]

- To understand the molecular epidemiology of TB, thereby contributing to the global knowledge of TB.[225]

- To monitor and evaluate TB programmes: an outreach screening service in the homeless and drug-using population in Rotterdam was evaluated using strain-typing;[84] long-term trends are used to evaluate TB control strategies e.g. in San Francisco and the Netherlands.[82]

- To meet international obligations for molecular surveillance: ECDC Molecular surveillance of MDR-TB in Europe project.[226]

- To create a national repository of strain types: this can be used for national and local analyses, larger research projects, and provides the opportunity for national and international collaboration.

8.9 Recommendations

This evaluation has resulted in recommendations for the TB-STS, for future molecular typing services and future evaluations.

Recommendations for the TB-STS

The recommendations for the TB-STS were discussed, debated and finalised by a multidisciplinary expert steering group. They are listed in Box 6. Alternative designs for the public health arm of the TB-STS were considered:

a) The *status quo*, whereby clusters are investigated at the local and national level with cluster investigators taking the lead and instructing local investigations;
b) A locally run service where people at the local level decide whether to investigate clusters or not, so that there is no bottle neck created at the cluster investigator level;
c) A service that does not initiate any cluster investigations (unless an outbreak is identified) but is used to monitor and evaluate TB control at the local and national level;
d) A resource for informing policy and for further research at both the molecular and public health level, but no specific public health actions; or
e) Given the available evidence, cease the TB-STS altogether.

PHE have adopted many of the recommendations outlined in Box 6 to improve the TB-STS, making this evaluation an important contribution to the TB strain typing strategy in England. Table 41 shows how the recommendations have influenced the TB-STS to date (as of December 2014). This suggests that the recommendations were judged to be appropriate and implementable, which is an important feature of an evaluation.

Box 6 – Recommendations made by the Evaluation Group for the TB-STS

1. The timely universal typing of all culture-confirmed TB cases should be continued. The resulting database of strain types linked to national surveillance data should be analysed nationally and locally, and be fully accessible across Public Health England (PHE), the NHS, UK universities and for international collaborations. The database could be used for the following:
    a. To access typing results in response to local or national incidents of suspected transmission, enabling the prospective, proactive, local-led application of strain typing for TB control and public health protection;
    b. To understand the national and local epidemiology of TB, including the identification of risk groups for TB attributable to recent transmission;
    c. To understand the molecular epidemiology of TB, including circulating strains, lineages and virulence;
    d. To monitor TB programmes by analysing the trends in estimates of recent transmission;
    e. To meet international obligations for molecular surveillance, Europe-wide and globally;
    f. To create a national repository of strain types.
2. The epidemiological analysis of the data should be prioritised. Findings should be reported back to local HPUs and NHS partners.
3. Cluster investigations should be reconsidered. The evaluation found no evidence to suggest that cluster investigations were effective or cost effective. However, as acknowledged in the limitations, this may be due to insufficient evidence.

    Local It is recommended that cluster investigations are no longer led by CIs but are initiated from the local level in response to local demand. Under this scenario, the CIs and Field Epidemiologists should be available to assist Local HPTs when they choose to launch a cluster investigation.

    Regional It is recommended that regional cluster investigations are discontinued as they appear to add little value.

    National It is recommended that the routine investigations of national cluster investigations are discontinued and that national cluster investigations be limited to clusters that have been identified to be of public health importance, e.g. rapidly growing clusters and clusters of drug resistant TB. Under this scenario, CIs and Field Epidemiologists should be available to support these investigations.
4. The STM should be released as a priority.
5. Public health and laboratory quality assurance should continue.
    a. The actions and outcomes of all cluster investigations that *are* conducted should be routinely recorded and be accessible for future evaluations.
    b. A false positive TB isolation identification and reporting protocol should be agreed between the reference laboratories.
    c. The completeness of typing data (i.e. the proportion of all isolates typed and the availability of full 24-loci typing profiles) for meaningful analysis and interpretation should be improved.
6. A review of the human resources and capacity across the TB-STS is recommended. This should include any potential impact the TB-STS has on the TB service more broadly. Moving forward, there is a need to recognise the potential capacity available to implement a complex intervention such as the TB-STS.
7. The key driver for the effectiveness and cost-effectiveness of TB control identified in this evaluation was diagnostic delay. The TB service should focus on and invest in interventions and TB control strategies that will lead to the earlier diagnosis of TB.

Table 41 – PHE's responses to the recommendations of this evaluation

	Recommendation	PHE response
1.	The timely universal typing of all culture-confirmed TB cases should be continued	PHE continues to strain type all culture-confirmed TB cases
2.	The epidemiological analysis of the data should be prioritised	The epidemiological analysis of the first three years of the TB-STS was presented in the PHE TB annual report in 2014,[3] other papers for publication are in the pipeline (personal correspondence with Maeve Lalor, PHE national CI)
3.	Cluster investigations should be reconsidered: a) Local investigations should be initiated from the local level b) Regional investigations should be discontinued c) National investigations should be limited to clusters that have been identified to be of public health importance	a) Local investigations are now initiated at the local level b) Regional investigations have stopped c) National clusters are investigated where identified to be of particular public health importance
4.	The STM should be released as a priority	The STM was released in November 2013
5.	Public health and laboratory quality assurance should continue: a) Cluster investigation activities and outcomes recorded b) False positive isolation recorded c) Completeness of typing reported	a) Cluster investigation activities and outcomes are recorded on the STM b) Data on false positive TB incidents are collected from laboratories and clinics using a form. The data are stored on a database held at Colindale. c) Completeness of typing is reported in the PHE TB annual report[3]
6.	Review the human resources and capacity across the TB-STS	Due to the recommendations from this evaluation, local cluster investigations are being led by HPUs who are responsible for reviewing their own resources. National cluster investigations have been de-prioritised, so far fewer clusters are being investigated.
7.	The TB service should focus on and invest in interventions and TB control strategies that will lead to the earlier diagnosis of TB	An analysis of the diagnostic delay was presented in the PHE TB annual report for the first time in 2014, suggesting that it has become an important indicator for TB control.

This evaluation has made an important contribution to our understanding of the TB-STS. It has already been helpful to other countries who are considering introducing or making changes to a strain typing service. Public Health Specialists from the Norwegian Institute of Public Health and the Robert Koch Institute in Germany have requested the results of this evaluation to inform the development of a strain typing service in their own countries, with Norwegian colleagues describing it as "so useful to us". It has identified key areas for improved data collection and reporting in the TB service, which would benefit future evaluations of a comprehensive TB control strategy (described below). In addition, this evaluation provides a benchmark for future evaluations of a typing service for TB control.

Recommendations for the design and implementation of a future service

The findings of this evaluation are important given the current trend to introduce and upgrade national typing services.[69,199] In the context of the rapid development of typing methodologies (e.g. WGS), political commitment to genomic analysis,[228] and the development of PHE's 2014-2019 National TB Strategy,[237] this evaluation provides important evidence for policy makers. Based on the findings of this evaluation, unless current or new typing and diagnostic techniques accelerate diagnosis (including through analysis of primary specimens), reduce diagnostic delay, dramatically reduce the time it takes to type, and/or are embedded in a user-friendly standardised TB management system, the adoption of such a method alone is unlikely to impact on TB control. Comprehensive TB control strategies that aim to reduce TB incidence over the next decades, need ongoing evaluation of proposed interventions. This includes evaluation of effective public health responses and appropriate use of strain typing, clinical and epidemiological information.

In order to better evaluate the impact of the TB-STS, or any future molecular surveillance strategy for TB, the following data are required from both before and after its introduction:

- The proportion of prevalent LTBI detected per year
- The diagnostic delay

- The proportion of TB cases that are detected as a result of cluster investigations, contact tracing and through self-presentation
- The proportion of contacts of TB cases that were identified and investigated through cluster investigations and traditional contact tracing
- The additional number of contacts with TB and LTBI detected because of expanded contact tracing based on a cluster investigations

Based on the learning from this evaluation, the following recommendations apply to the implementation of any future public health typing service for infectious diseases (e.g. a future TB molecular typing service such as WGS):

1. The objectives of the service should be well-defined from the outset, making explicit the public health implications of such a service, and should be SMART (specific, measurable, attainable, relevant and time-bound);
2. Service implementation should consider the evaluation design to enable the evaluation to capture the required data for a robust analysis (see next set of recommendations regarding evaluation);
3. The service should be designed involving all stakeholders so that the role of each group is plausible given the resources (e.g. can you expect HPUs to manage cluster investigations?);
4. Resources should be allocated for all elements of the service (e.g. if the NHS is required to contribute to the implementation of the service, this should be resourced);
5. The molecular method used to type isolates should be able to discriminate between strains that have not been transmitted between patients recently, and the molecular clock of the biomarker should be slow enough to identify cases in the same chain of recent transmission and the genetic diversity of the organism (e.g. *M.tuberculosis*) in the population. In addition, it should be stable, rapid, reproducible across laboratories (e.g. could it be performed in a

single national laboratory to reduce cross-laboratory variation?), easy to perform and interpret, and applicable to clinical samples;

6. The typing technique used must reduce diagnostic delay and/or dramatically reduce the time it takes to type in order for the intervention to impact on TB control. Alternatively, if a technology fulfills more than one purpose (for example case identification, typing and detection of drug resistance) it is more likely to be cost effective than if it is typing alone. For TB, this is especially true if you can identify resistance sooner, as MDR cases are expensive to treat;
7. The service needs to be able to adapt to newer technologies and typing methods introduced in the future;
8. Training should involve all those involved in delivering the service and expectations of each role in service delivery should be made clear;
9. Software development/information management is often a particularly risky but important component of a complex intervention. It should be well-resourced and managed, and would ideally be completed before the service is started; and
10. Data collection for quality control, monitoring and evaluation should be implemented from the start of the service and be an integrated part of the service.

Recommendations for the evaluation of a future service

This evaluation could be used as a template for future evaluations of other services relating to alternative infections, especially considering the importance of increasing the evidence base on public health interventions and the compatible evaluation methods. Considering the rapid decline in the cost of sequencing,[238] it will eventually be possible and affordable to sequence everything. It is worth learning from the experiences in TB how best to use this information to protect and improve public health.

This evaluation has led to the following recommendations for the evaluation of a future service:

1. Evaluation is difficult but worthwhile, and is an essential component of any intervention;
2. The evaluation should be conducted by an independent evaluation team and the dissemination of results should be timely, shared with stakeholders, and made available to the public;
3. It is important to work with those designing and implementing the service to ensure that their objectives are appropriate and SMART (see above recommendations for the design/implementation of a service);
4. Ideally, the evaluation will be able to influence the implementation of the service in order to design a robust, experimental evaluation study (e.g. ensuring time to collect baseline data, rolling out the service gradually by geographical area for a step-wedge design, or setting up a cluster randomised control trial);
5. Evaluation frameworks are useful for structuring the evaluation and encompassing the whole service. The Donabedian Framework[161] is particularly simple, flexible, and can be applied across any intervention. Alternative models may be appropriate depending on the nature of the intervention and its implementation (e.g. Theory of Change);[157]
6. Mixed-methods ensure that different data types can be collected and analysed appropriately. This is important for capturing the unexpected;
7. Where possible, it is recommended to consider the long term outputs of a service;
8. It is important to consider the impact of the evaluation on service delivery (e.g. data collection for the evaluation should not be a large burden on those delivering the service, unless additional resources are allocated for this purpose); and

9. To validate the quality, and to enable other countries, services and disciplines to learn from the results of future evaluations, findings should be available in the public domain, preferably published in peer-reviewed journals.

8.10 Conclusion

Evaluating a complex public health intervention requires a pragmatic approach, taking into account how the service has been implemented. The TB-STS was not found to be effective or cost-effective during its first three years; however, the broader benefits of such a service justify its continuation as a modified service. To reduce costs, improve efficiency and increase effectiveness, changes are recommended to the TB-STS, including discontinuing routine cluster investigations and focussing on reducing diagnostic delay across the TB programme. This evaluation of a complex intervention informs the future of strain typing in the era of rapidly-advancing technologies. Future evaluations of public health interventions should be established prior to the implementation of a national intervention. This would ensure that sufficient baseline data can be collected; enable the evaluation design to influence the roll-out of the intervention; and increase the evidence base on specific interventions and evaluation methods.

References

1 World Health Organization. Global Tuberculosis Report 2014. .

2 ECDC and WHO/Europe joint report on tuberculosis surveillance and monitoring in Europe. *Euro Surveill Bull Eur Sur Mal Transm Eur Commun Dis Bull* 2014; **19**.

3 Public Health England. Tuberculosis in the UK. 2014 report. London, 2014.

4 Health Protection Agency. Tuberculosis in London: Annual Review (2011 data). London, 2012.

5 Parrish NM, Dick JD, Bishai WR. Mechanisms of latency in Mycobacterium tuberculosis. *Trends Microbiol* 1998; **6**: 107–12.

6 Kaufmann SHE, McMichael AJ. Annulling a dangerous liaison: vaccination strategies against AIDS and tuberculosis. *Nat Med* 2005; **11**: S33–44.

7 Horsburgh CR Jr, O'Donnell M, Chamblee S, *et al.* Revisiting rates of reactivation tuberculosis: a population-based approach. *Am J Respir Crit Care Med* 2010; **182**: 420–5.

8 Mathema B, Kurepina NE, Bifani PJ, Kreiswirth BN. Molecular Epidemiology of Tuberculosis: Current Insights. *Clin Microbiol Rev* 2006; **19**: 658–85.

9 Crampin AC, Mwaungulu JN, Mwaungulu FD, *et al.* Recurrent TB: relapse or reinfection? The effect of HIV in a general population cohort in Malawi. *AIDS Lond Engl* 2010; **24**: 417–26.

10 Aldridge RW, Yates TA, Zenner D, White PJ, Abubakar I, Hayward AC. Pre-entry screening programmes for tuberculosis in migrants to low-incidence countries: a systematic review and meta-analysis. *Lancet Infect Dis* 2014; **14**: 1240–9.

11 Jit M, Stagg HR, Aldridge RW, White PJ, Abubakar I, Find and Treat Evaluation Team. Dedicated outreach service for hard to reach patients with tuberculosis in London: observational study and economic evaluation. *BMJ* 2011; **343**: d5376.

12 Erkens CGM, Kamphorst M, Abubakar I, *et al.* Tuberculosis contact investigation in low prevalence countries: a European consensus. *Eur Respir J* 2010; **36**: 925–49.

13 Veen J. Microepidemics of tuberculosis: the stone-in-the-pond principle. *Tuber Lung Dis* 1992; **73**: 73–6.

14 Samra Z, Kaufman L, Bechor J, Bahar J. Comparative Study of Three Culture Systems for Optimal Recovery of Mycobacteria from Different Clinical Specimens. *Eur J Clin Microbiol Infect Dis* 2000; **19**: 750–4.

15 Mossong J, Hens N, Jit M, *et al.* Social Contacts and Mixing Patterns Relevant to the Spread of Infectious Diseases. *PLoS Med* 2008; **5**: e74.

16 Fox GJ, Barry SE, Britton WJ, Marks GB. Contact investigation for tuberculosis: a systematic review and meta-analysis. *Eur Respir J* 2013; **41**: 140–56.

17 Mulder C, Klinkenberg E, Manissero D. Effectiveness of tuberculosis contact tracing among migrants and the foreign-born population. *Euro Surveill Bull Eur Sur Mal Transm Eur Commun Dis Bull* 2009; **14**. http://www.ncbi.nlm.nih.gov/pubmed/19317977 (accessed April 25, 2012).

18 Mulder C, van Deutekom H, Huisman EM, *et al.* Coverage and yield of tuberculosis contact investigations in the Netherlands. *Int J Tuberc Lung Dis* 2011; **15**: 1630–7.

19 Borraccino A, Migliore E, Piccioni P, Baussano I, Carosso A, Bugiani M. Yield of tuberculosis contact investigation in a low-incidence country. *J Infect* 2014; **68**: 448–54.

20 Golub JE, Bur S, Cronin WA, *et al.* Delayed tuberculosis diagnosis and tuberculosis transmission. *Int J Tuberc Lung Dis* 2006; **10**: 24–30.

21 Underwood BR, White VLC, Baker T, Law M, Moore-Gillon JC. Contact tracing and population screening for tuberculosis--who should be assessed? *J Public Health Med* 2003; **25**: 59–61.

22 Ansari S, Thomas S, Campbell IA, Furness L, Evans MR. Refined tuberculosis contact tracing in a low incidence area. *Respir Med* 1998; **92**: 1127–31.

23 Ormerod LP. Results of tuberculosis contact tracing: Blackburn 1982-90. *Respir Med* 1993; **87**: 127–31.

24 Saunders MJ, Koh GCKW, Small AD, Dedicoat M. Predictors of contact tracing completion and outcomes in tuberculosis: a 21-year retrospective cohort study. *Int J Tuberc Lung Dis Off J Int Union Tuberc Lung Dis* 2014; **18**: 640–6.

25 Aissa K, Madhi F, Ronsin N, *et al.* Evaluation of a Model for Efficient Screening of Tuberculosis Contact Subjects. *Am J Respir Crit Care Med* 2008; **177**: 1041–7.

26 Sloot R, Schim van der Loeff MF, Kouw PM, Borgdorff MW. Yield of tuberculosis contact investigations in Amsterdam: opportunities for improvement. *Eur Respir J* 2014; **44**: 714–24.

27 Sloot R, Schim van der Loeff MF, Kouw PM, Borgdorff MW. Risk of Tuberculosis After Recent Exposure: a 10-year Follow-up Study of Contacts in Amsterdam. *Am J Respir Crit Care Med* 2014; published online Sept 29. DOI:10.1164/rccm.201406-1159OC.

28 Behr MA, Hopewell PC, Paz EA, Kawamura LM, Schecter GF, Small PM. Predictive value of contact investigation for identifying recent transmission of Mycobacterium tuberculosis. *Am J Respir Crit Care Med* 1998; **158**: 465–9.

29 Davidow AL, Mangura BT, Wolman MS, *et al.* Workplace contact investigations in the United States. *Int J Tuberc Lung Dis* 2003; **7**: S446–52.

30 Jereb J, Etkind SC, Joglar OT, Moore M, Taylor Z. Tuberculosis contact investigations: outcomes in selected areas of the United States, 1999. *Int J Tuberc Lung Dis Off J Int Union Tuberc Lung Dis* 2003; **7**: S384–90.

31 Marks SM, Taylor Z, Qualls NL, Shrestha-Kuwahara RJ, Wilce MA, Nguyen CH. Outcomes of Contact Investigations of Infectious Tuberculosis Patients. *Am J Respir Crit Care Med* 2000; **162**: 2033–8.

32 Morán-Mendoza O, Marion SA, Elwood K, Patrick DM, FitzGerald JM. Tuberculin skin test size and risk of tuberculosis development: a large population-based study in contacts. *Int J Tuberc Lung Dis Off J Int Union Tuberc Lung Dis* 2007; **11**: 1014–20.

33 Reichler MR, Reves R, Bur S, et al. Evaluation of investigations conducted to detect and prevent transmission of tuberculosis. *JAMA* 2002; **287**: 991–5.

34 Sprinson JE, Flood J, Fan CS, *et al.* Evaluation of tuberculosis contact investigations in California. *Int J Tuberc Lung Dis Off J Int Union Tuberc Lung Dis* 2003; **7**: S363–8.

35 Saldana L, Abid M, McCarthy N, Hunter N, Inglis R, Anders K. Factors affecting delay in initiation of treatment of tuberculosis in the Thames Valley, UK. *Public Health* 2013; **127**: 171–7.

36 Yimer S, Bjune G, Alene G. Diagnostic and treatment delay among pulmonary tuberculosis patients in Ethiopia: a cross sectional study. *BMC Infect Dis* 2005; **5**: 112.

37 Sreeramareddy CT, Panduru KV, Menten J, Van den Ende J. Time delays in diagnosis of pulmonary tuberculosis: a systematic review of literature. *BMC Infect Dis* 2009; **9**: 91.

38 Storla DG, Yimer S, Bjune GA. A systematic review of delay in the diagnosis and treatment of tuberculosis. *BMC Public Health* 2008; **8**: 15.

39 Abubakar I, Laundy MT, French CE, Shingadia D. Epidemiology and treatment outcome of childhood tuberculosis in England and Wales: 1999-2006. *Arch Dis Child* 2008; **93**: 1017–21.

40 Lewis KE, Stephens C, Shahidi MM, Packe G. Delay in starting treatment for tuberculosis in east London. *Commun Dis Public Health PHLS* 2003; **6**: 133–8.

41 Llewelyn M, Cropley I, Wilkinson RJ, Davidson RN. Tuberculosis diagnosed during pregnancy: a prospective study from London. *Thorax* 2000; **55**: 129–32.

42 Rodger A, Jaffar S, Paynter S, Hayward A, Carless J, Maguire H. Delay in the diagnosis of pulmonary tuberculosis, London, 1998-2000: analysis of surveillance data. *BMJ* 2003; **326**: 909–10.

43 Wares DF, Ormerod LP, Morton S. Delay in diagnosis of tuberculosis: of remaining concern in England and Wales. *J Public Health Med* 1999; **21**: 355–6.

44 Craig SE, Bettinson H, Sabin CA, Gillespie SH, Lipman MCI. Think TB! Is the diagnosis of pulmonary tuberculosis delayed by the use of antibiotics? *Int J Tuberc Lung Dis Off J Int Union Tuberc Lung Dis* 2009; **13**: 208–13.

45 French CE, Kruijshaar ME, Jones JA, Abubakar I. The influence of socio-economic deprivation on tuberculosis treatment delays in England, 2000–2005. *Epidemiol Infect* 2009; **137**: 591–6.

46 Paynter S, Hayward A, Wilkinson P, Lozewicz S, Coker R. Patient and health service delays in initiating treatment for patients with pulmonary tuberculosis: retrospective cohort study. *Int J Tuberc Lung Dis* 2004; **8**: 180–5.

47 Sultan H, Haroon S, Syed N. Delay and completion of tuberculosis treatment: a cross-sectional study in the West Midlands, UK. *J Public Health* 2012; published online June 21. DOI:10.1093/pubmed/fds046.

48 NICE. Tuberculosis: NICE guideline. Clinical diagnosis and management of tuberculosis, and measures for its prevention and control. Natl. Inst. Health Clin. Excell. 2011. http://www.nice.org.uk/ (accessed Sept 20, 2012).

49 Abubakar I, Pimpin L, Ariti C, *et al.* Systematic review and meta-analysis of the current evidence on the duration of protection by bacillus Calmette-Guérin vaccination against tuberculosis. *Health Technol Assess Winch Engl* 2013; **17**: 1–372, v – vi.

50 Schürch AC, van Soolingen D. DNA fingerprinting of Mycobacterium tuberculosis: from phage typing to whole-genome sequencing. *Infect Genet Evol J Mol Epidemiol Evol Genet Infect Dis* 2012; **12**: 602–9.

51 Mathema B, Kurepina N, Fallows D, Kreiswirth BN. Lessons from molecular epidemiology and comparative genomics. *Semin Respir Crit Care Med* 2008; **29**: 467–80.

52 Van Soolingen D. Molecular epidemiology of tuberculosis and other mycobacterial infections: main methodologies and achievements. *J Intern Med* 2001; **249**: 1–26.

53 Field N, Cohen T, Struelens MJ, *et al.* Strengthening the Reporting of Molecular Epidemiology for Infectious Diseases (STROME-ID): an extension of the STROBE statement. *Lancet Infect Dis* 2014; **14**: 341–52.

54 Van Embden JD, Cave MD, Crawford JT, *et al.* Strain identification of Mycobacterium tuberculosis by DNA fingerprinting: recommendations for a standardized methodology. *J Clin Microbiol* 1993; **31**: 406–9.

55 Kamerbeek J, Schouls L, Kolk A, *et al.* Simultaneous detection and strain differentiation of Mycobacterium tuberculosis for diagnosis and epidemiology. *J Clin Microbiol* 1997; **35**: 907–14.

56 Supply P, Lesjean S, Savine E, Kremer K, van Soolingen D, Locht C. Automated high-throughput genotyping for study of global epidemiology of Mycobacterium tuberculosis based on mycobacterial interspersed repetitive units. *J Clin Microbiol* 2001; **39**: 3563–71.

57 Cole ST, Brosch R, Parkhill J, *et al.* Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. *Nature* 1998; **393**: 537–44.

58 Kremer K, Arnold C, Cataldi A, *et al.* Discriminatory Power and Reproducibility of Novel DNA Typing Methods for Mycobacterium tuberculosis Complex Strains. *J Clin Microbiol* 2005; **43**: 5628–38.

59 BioNumerics version 6.5 created by Applied Maths NV. Available from http://www.applied-maths.com. .

60 Supply P. MIRU-VNTR typing: the new international standard for TB molecular epidemiology. Symposium of the Institut Pasteur de Tunisia, 2010.

61 Mazars E, Lesjean S, Banuls AL, *et al.* High-resolution minisatellite-based typing as a portable approach to global analysis of Mycobacterium tuberculosis molecular epidemiology. *Proc Natl Acad Sci U S A* 2001; **98**: 1901–6.

62 Bauer J, Andersen AB, Kremer K, Miörner H. Usefulness of spoligotyping to discriminate IS6110 low-copy-number Mycobacterium tuberculosis complex strains cultured in Denmark. *J Clin Microbiol* 1999; **37**: 2602–6.

63 Kremer K, van Soolingen D, Frothingham R, *et al.* Comparison of methods based on different molecular epidemiological markers for typing of

Mycobacterium tuberculosis complex strains: interlaboratory study of discriminatory power and reproducibility. *J Clin Microbiol* 1999; **37**: 2607–18.

64 Cowan LS, Diem L, Monson T, *et al.* Evaluation of a two-step approach for large-scale, prospective genotyping of Mycobacterium tuberculosis isolates in the United States. *J Clin Microbiol* 2005; **43**: 688–95.

65 Tsolaki AG, Gagneux S, Pym AS, *et al.* Genomic deletions classify the Beijing/W strains as a distinct genetic lineage of Mycobacterium tuberculosis. *J Clin Microbiol* 2005; **43**: 3185–91.

66 Comas I, Homolka S, Niemann S, Gagneux S. Genotyping of genetically monomorphic bacteria: DNA sequencing in Mycobacterium tuberculosis highlights the limitations of current methodologies. *PloS One* 2009; **4**: e7815.

67 Supply P, Mazars E, Lesjean S, Vincent V, Gicquel B, Locht C. Variable human minisatellite-like regions in the Mycobacterium tuberculosis genome. *Mol Microbiol* 2000; **36**: 762–71.

68 GeneScan™ Available from http://www.appliedbiosystems.com/. .

69 De Beer JL, van Ingen J, de Vries G, *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**: 1193–8.

70 Gopaul KK, Brown TJ, Gibson AL, Yates MD, Drobniewski FA. Progression toward an improved DNA amplification-based typing technique in the study of Mycobacterium tuberculosis epidemiology. *J Clin Microbiol* 2006; **44**: 2492–8.

71 Barlow REL, 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**: 2453–7.

72 Allix-Béguec C, Fauville-Dufaux M, Supply P. Three-year population-based evaluation of standardized mycobacterial interspersed repetitive-unit-variable-number tandem-repeat typing of Mycobacterium tuberculosis. *J Clin Microbiol* 2008; **46**: 1398–406.

73 Oelemann MC, Diel R, Vatin V, *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**: 691–7.

74 Grant A, Arnold C, Thorne N, Gharbia S, Underwood A. Mathematical modelling of Mycobacterium tuberculosis VNTR loci estimates a very slow mutation rate for the repeats. *J Mol Evol* 2008; **66**: 565–74.

75 Reyes JF, Tanaka MM. Mutation rates of spoligotypes and variable numbers of tandem repeat loci in Mycobacterium tuberculosis. *Infect Genet Evol J Mol Epidemiol Evol Genet Infect Dis* 2010; **10**: 1046–51.

76 Tanaka MM, Reyes JF. Mutation rate of VNTR loci in Mycobacterium tuberculosis: Response to Supply et al. *Infect Genet Evol* 2011; **11**: 1189–90.

77 Supply P, Niemann S, Wirth T. On the mutation rates of spoligotypes and variable numbers of tandem repeat loci of Mycobacterium tuberculosis. *Infect Genet Evol* 2011; **11**: 251–2.

78 Walker TM, Ip CL, Harrell RH, *et al.* Whole-genome sequencing to delineate Mycobacterium tuberculosis outbreaks: a retrospective observational study. *Lancet Infect Dis* 2013; **13**: 137–46.

79 Gardy JL, Johnston JC, Ho Sui SJ, *et al.* Whole-genome sequencing and social-network analysis of a tuberculosis outbreak. *N Engl J Med* 2011; **364**: 730–9.

80 Gardy J.L. Investigation of disease outbreaks with genome sequencing. Lancet Infect. Dis. 2013.

81 Bryant JM, Harris SR, Parkhill J, *et al.* Whole-genome sequencing to establish relapse or re-infection with Mycobacterium tuberculosis: a retrospective observational study. *Lancet Respir Med* 2013; **1**: 786–92.

82 Bryant JM, Schürch AC, van Deutekom H, *et al.* Inferring patient to patient transmission of Mycobacterium tuberculosis from whole genome sequencing data. *BMC Infect Dis* 2013; **13**: 110.

83 Ford CB, Shah RR, Maeda MK, *et al.* Mycobacterium tuberculosis mutation rate estimates from different lineages predict substantial differences in the emergence of drug-resistant tuberculosis. *Nat Genet* 2013; **45**: 784 – +.

84 Wetterstrand KA. DNA sequencing costs: data from the NHGRI Large-Scale Genome Sequencing Program. 2011. www.genome.gov/sequencingcosts.

85 Crook D, colleagues. UK CRC Modernising Medical Microbiology Consortium — UK CRC Modernising Medical Microbiology. http://www.modmedmicro.ac.uk/ (accessed April 14, 2012).

86 Small PM, Hopewell PC, Singh SP, *et al.* The epidemiology of tuberculosis in San Francisco. A population-based study using conventional and molecular methods. *N Engl J Med* 1994; **330**: 1703–9.

87 Borgdorff MW, van den Hof S, Kremer K, *et al.* Progress towards tuberculosis elimination: secular trend, immigration and transmission. *Eur Respir J Off J Eur Soc Clin Respir Physiol* 2010; **36**: 339–47.

88 Clark CM, Driver CR, Munsiff SS, *et al.* Universal genotyping in tuberculosis control program, New York City, 2001-2003. *Emerg Infect Dis* 2006; **12**: 719–24.

89 De Vries G, van Hest RAH, Richardus JH. Impact of mobile radiographic screening on tuberculosis among drug users and homeless persons. *Am J Respir Crit Care Med* 2007; **176**: 201–7.

90 Shorten RJ, Gillespie SH, Sule O, Lipman M, McHugh TD. Molecular strain typing of M. tuberculosis isolates from a suspected outbreak involving a faulty bronchoscope. *J Hosp Infect* 2005; **61**: 86–7.

91 Maguire H, Brailsford S, Carless J, *et al.* Large outbreak of isoniazid-monoresistant tuberculosis in London, 1995 to 2006: case-control study and recommendations. *Euro Surveill Bull Eur Sur Mal Transm Eur Commun Dis Bull* 2011; **16**. http://www.ncbi.nlm.nih.gov/pubmed/21489373 (accessed April 12, 2012).

92 Van Soolingen D. Whole-genome sequencing of Mycobacterium tuberculosis as an epidemiological marker. *Lancet Respir Med* 2014; **2**: 251–2.

93 Houben RMGJ, Glynn JR. A systematic review and meta-analysis of molecular epidemiological studies of tuberculosis: development of a new tool to aid interpretation. *Trop Med Int Health* 2009; **14**: 892–909.

94 Mears J, Abubakar I, Cohen T, McHugh TD, Sonnenberg P. Effect of study design and setting on tuberculosis clustering estimates using Mycobacterial Interspersed Repetitive Units-Variable Number Tandem Repeats (MIRU-VNTR): a systematic review. *BMJ Open* 2015; **5**: e005636.

95 Godfrey-Faussett P, Sonnenberg P, Shearer SC, *et al.* Tuberculosis control and molecular epidemiology in a South African gold-mining community. *Lancet* 2000; **356**: 1066–71.

96 Fok A, Numata Y, Schulzer M, FitzGerald MJ. Risk factors for clustering of tuberculosis cases: a systematic review of population-based molecular epidemiology studies [Review Article]. *Int J Tuberc Lung Dis* May; **12**: 480–92.

97 McNabb SJN, Kammerer JS, Hickey AC, *et al.* Added epidemiologic value to tuberculosis prevention and control of the investigation of clustered genotypes of Mycobacterium tuberculosis isolates. *Am J Epidemiol* 2004; **160**: 589–97.

98 Cronin WA, Golub JE, Magder LS, *et al.* Epidemiologic Usefulness of Spoligotyping for Secondary Typing of Mycobacterium Tuberculosis Isolates with Low Copy Numbers of IS6110. *J Clin Microbiol* 2001; **39**: 3709–11.

99 Ruddy MC, Davies AP, Yates MD, *et al.* Outbreak of isoniazid resistant tuberculosis in north London. *Thorax* 2004; **59**: 279–85.

100 Van Soolingen D, Borgdorff MW, de Haas PE, *et al.* Molecular epidemiology of tuberculosis in the Netherlands: a nationwide study from 1993 through 1997. *J Infect Dis* 1999; **180**: 726–36.

101 Aspler A, Chong H, Kunimoto D, *et al.* Sustained Intra- and Inter-jurisdictional Transmission of Tuberculosis within a Mobile, Multi-ethnic Social Network: Lessons for Tuberculosis Elimination. *Can J Public Health* 2010; **101**: 205–9.

102 Kik SV, Verver S, van Soolingen D, *et al.* Tuberculosis Outbreaks Predicted by Characteristics of First Patients in a DNA Fingerprint Cluster. *Am J Respir Crit Care Med* 2008; **178**: 96–104.

103 Driver CR, Macaraig M, McElroy PD, *et al.* Which Patients' Factors Predict the Rate of Growth of Mycobacterium tuberculosis Clusters in an Urban Community? *Am J Epidemiol* 2006; **164**: 21–31.

104 Mulder C, Erkens CGM, Kouw PM, *et al.* Missed opportunities in tuberculosis control in The Netherlands due to prioritization of contact investigations. *Eur J Public Health* 2011; published online March 7. DOI:10.1093/eurpub/ckr017.

105 Verver S, Warren RM, Munch Z, *et al.* Proportion of tuberculosis transmission that takes place in households in a high-incidence area. *Lancet* 2004; **363**: 212–4.

106 Houben RMGJ, Crampin AC, Ndhlovu R, *et al.* Human immunodeficiency virus associated tuberculosis more often due to recent infection than reactivation of latent infection. *Int J Tuberc Lung Dis Off J Int Union Tuberc Lung Dis* 2011; **15**: 24–31.

107 Alexander DC, Guthrie JL, Pyskir D, *et al.* Mycobacterium tuberculosis in Ontario, Canada: Insights from IS6110 restriction fragment length polymorphism and mycobacterial interspersed repetitive-unit-variable-number tandem-repeat genotyping. *J Clin Microbiol* 2009; **47**: 2651–4.

108 De Boer AS, Borgdorff MW, Vynnycky E, Sebek MM, van Soolingen D. Exogenous re-infection as a cause of recurrent tuberculosis in a low-incidence area. *Int J Tuberc Lung Dis Off J Int Union Tuberc Lung Dis* 2003; **7**: 145–52.

109 Glynn JR, Crampin AC, Yates MD, *et al.* The Importance of Recent Infection with Mycobacterium tuberculosis in an Area with High HIV Prevalence: A Long-Term Molecular Epidemiological Study in Northern Malawi. *J Infect Dis* 2005; **192**: 480–7.

110 Cohen T, Wilson D, Wallengren K, Samuel EY, Murray M. Mixed-strain Mycobacterium tuberculosis infections among patients dying in a hospital in KwaZulu-Natal, South Africa. *J Clin Microbiol* 2011; **49**: 385–8.

111 Mallard K, McNerney R, Crampin AC, *et al.* Molecular detection of mixed infections of Mycobacterium tuberculosis strains in sputum samples from patients in Karonga District, Malawi. *J Clin Microbiol* 2010; **48**: 4512–8.

112 Vynnycky E, Fine PE. The natural history of tuberculosis: the implications of age-dependent risks of disease and the role of reinfection. *Epidemiol Infect* 1997; **119**: 183–201.

113 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* 2001; **358**: 1687–93.

114 Burman WJ, Reves RR. Review of False-Positive Cultures for Mycobacterium tuberculosis and Recommendations for Avoiding Unnecessary Treatment. *Clin Infect Dis* 2000; **31**: 1390–5.

115 De Boer AS, Blommerde B, de Haas PEW, *et al.* False-positive mycobacterium tuberculosis cultures in 44 laboratories in The Netherlands (1993 to 2000): incidence, risk factors, and consequences. *J Clin Microbiol* 2002; **40**: 4004–9.

116 Small PM, McClenny NB, Singh SP, Schoolnik GK, Tompkins LS, Mickelsen PA. Molecular strain typing of Mycobacterium tuberculosis to confirm cross-contamination in the mycobacteriology laboratory and modification of procedures to minimize occurrence of false-positive cultures. *J Clin Microbiol* 1993; **31**: 1677–82.

117 Martín A, Herranz M, Lirola MM, Fernández RF, Bouza E, de Viedma DG. Optimized molecular resolution of cross-contamination alerts in clinical mycobacteriology laboratories. *BMC Microbiol* 2008; **8**: 30.

118 Ruddy M, McHugh TD, Dale JW, *et al.* Estimation of the Rate of Unrecognized Cross-Contamination with Mycobacterium tuberculosis in London Microbiology Laboratories. *J Clin Microbiol* 2002; **40**: 4100–4.

119 Jasmer RM, Roemer M, Hamilton J, *et al.* A Prospective, Multicenter Study of Laboratory Cross-Contamination of Mycobacterium tuberculosis Cultures. *Emerg Infect Dis* 2002; **8**: 1260–3.

120 Braden CR, Templeton GL, Stead WW, Bates JH, Cave MD, Valway SE. Retrospective Detection of Laboratory Cross-Contamination of Mycobacterium Tuberculosis Cultures with Use of DNA Fingerprint Analysis. *Clin Infect Dis* 1997; **24**: 35–40.

121 Frieden TR, Woodley CL, Crawford JT, Lew D, Dooley SM. The molecular epidemiology of tuberculosis in New York City: the importance of nosocomial transmission and laboratory error. *Int J Tuber Lung Dis* 1996; **77**: 407–13.

122 Slump E, Erkens C, van Hunen R, van Rest J, Schimmel H, Van Soolingen D. Tuberculose in Nederland 2012 - Surveillancerapport. .

123 Public Health England. Tuberculosis in the UK: Annual report on tuberculosis surveillance in the UK, 2013. London, 2013.

124 Bauer J, Yang Z, Poulsen S, Andersen AB. Results from 5 years of nationwide DNA fingerprinting of Mycobacterium tuberculosis complex isolates in a country with a low incidence of M. tuberculosis infection. *J Clin Microbiol* 1998; **36**: 305–8.

125 Kamper-Jørgensen Z, Andersen AB, Kok-Jensen A, *et al.* Clustered tuberculosis in a low-burden country: Nationwide genotyping through 15 years. *J Clin Microbiol* 2012; **50**: 2660–7.

126 Lillebaek T, Andersen ÅB, Bauer J, *et al.* Risk of Mycobacterium tuberculosis Transmission in a Low-Incidence Country Due to Immigration from High-Incidence Areas. *J Clin Microbiol* 2001; **39**: 855–61.

127 Lambregts-van Weezenbeek CSB, Sebek MMGG, van Gerven PJHJ, *et al.* Tuberculosis contact investigation and DNA fingerprint surveillance in The Netherlands: 6 years' experience with nation-wide cluster feedback and cluster monitoring. *Int J Tuberc Lung Dis Off J Int Union Tuberc Lung Dis* 2003; **7**: S463–70.

128 Borgdorff MW, Nagelkerke N, van Soolingen D, de Haas PE, Veen J, van Embden JD. Analysis of tuberculosis transmission between nationalities in the Netherlands in the period 1993-1995 using DNA fingerprinting. *Am J Epidemiol* 1998; **147**: 187–95.

129 Borgdorff MW, Nagelkerke NJD, De Haas PEW, Van Soolingen D. Transmission of Mycobacterium Tuberculosis Depending on the Age and Sex of Source Cases. *Am J Epidemiol* 2001; **154**: 934–43.

130 Dahle UR, Eldholm V, Winje BA, Mannsåker 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**: 930–5.

131 Dahle UR, Sandven P, Heldal E, Mannsaaker T, Caugant DA. Deciphering an Outbreak of Drug-Resistant Mycobacterium tuberculosis. *J Clin Microbiol* 2003; **41**: 67–72.

132 Dahle UR, Sandven P, Heldal E, Caugant DA. Continued Low Rates of Transmission of Mycobacterium tuberculosis in Norway. *J Clin Microbiol* 2003; **41**: 2968–73.

133 Bidovec-Stojkovic U, Zolnir-Dovc M, Supply P. One year nationwide evaluation of 24-locus MIRU-VNTR genotyping on Slovenian Mycobacterium tuberculosis isolates. *Respir Med* 2011; **105 Suppl 1**: S67–73.

134 Zolnir-Dovc M, Poljak M, Erzen D, Sorli J. Molecular epidemiology of tuberculosis in Slovenia: results of a one-year (2001) nation-wide study. *Scand J Infect Dis* 2003; **35**: 863–8.

135 Haddad MB, Diem LA, Cowan LS, *et al.* Tuberculosis genotyping in six low-incidence States, 2000-2003. *Am J Prev Med* 2007; **32**: 239–43.

136 Centers for Disease Control and Prevention. New CDC Program for Rapid Genotyping of Mycobacterium Tuberculosis Isolates. *JAMA J Am Med Assoc* 2005; **293**: 2086–2086.

137 CDC - Appendix D - Reported Tuberculosis in the United States, 2012 - TB. http://www.cdc.gov/tb/statistics/reports/2012/appendixD.htm (accessed Dec 1, 2014).

138 Guide to the Application of Genotyping to Tuberculosis Prevention and Control. Handbook for TB Controllers, Epidemiologists, Laboratorians, and Other Program Staff. 2004; published online June. http://www.cdc.gov/tb/programs/genotyping/manual.htm (accessed Dec 1, 2014).

139 Department of Health; Donaldson SL. Stopping Tuberculosis in England: An action plan from the Chief Medical Officer. 2004; published online Oct 7. http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_4090417 (accessed April 11, 2012).

140 National TB Strain Typing Project Board HPA. Project Initiation Document (version 1.7). 2011; published online Feb.

141 Evans JT, Gardiner S, Smith EG, Webber R, Hawkey PM. Global Origin of Mycobacterium tuberculosis in the Midlands, UK. *Emerg Infect Dis* 2010; **16**: 542–5.

142 Evans JT, Serafino Wani RL, Anderson L, *et al.* A geographically-restricted but prevalent Mycobacterium tuberculosis strain identified in the West Midlands Region of the UK between 1995 and 2008. *PloS One* 2011; **6**: e17930.

143 Love J, Sonnenberg P, Glynn JR, *et al.* Molecular epidemiology of tuberculosis in England, 1998. *Int J Tuberc Lung Dis Off J Int Union Tuberc Lung Dis* 2009; **13**: 201–7.

144 Cambell D, Stanley J. Experimental and Quasi-Experimental Designs for Research. Chicago: Rand McNally, 1966.

145 Joint Committee on Standards for Educational Evaluation. 1980.

146 Patton M. Utilization-focussed evaluation. London: Sage, 1978.

147 Patton M. Practical Evaluation. London: Sage, 1982.

148 Guba E, Lincoln Y. Fourth generation evaluation. CA: Sage, 1989.

149 Rossi PH, Freeman H. Evaluation: A systematic approach, 3rd edn. CA: Sage, 1985.

150 Pawson R, Tilley N. Realistic Evaluation. SAGE, 1997.

151 Wallace WL. The Logic of Science in Sociology. New Brunswick: Transaction Publishers, 1971.

152 Bauer J, Kok-Jensen A, Faurschou P, Thuesen J, Taudorf E, Andersen AB. A prospective evaluation of the clinical value of nation-wide DNA fingerprinting of tuberculosis isolates in Denmark. *Int J Tuberc Lung Dis* 2000; **4**: 295–9.

153 Turner P. An evaluation of the national tuberculosis strain typing strategy. 2006.

154 Craig P, Dieppe P, Macintyre S, *et al.* Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ* 2008; **337**: a1655.

155 Campbell M, Fitzpatrick R, Haines A, *et al.* Framework for design and evaluation of complex interventions to improve health. *BMJ* 2000; **321**: 694–6.

156 Chen H-T, Rossi PH. Issues in the theory-driven perspective. *Eval Program Plann* 1989; **12**: 299–306.

157 Chen H-T. A theory-driven evaluation perspective on mixed methods research. *Res Sch* 2006; **13**: 75–83.

158 Weiss CH. Nothing as Practical as Good Theory: Exploring Theory-Based Evaluation for Comprehensive Community Initiatives for Children and Families. In: New Approaches to Evaluating Community Initiatives Volume 1 Concepts, Methods, and Contexts. Washington DC: The Aspen Institute, 1995.

159 Kernick D. Assessing complex interventions. Guidelines perpetuate inappropriate methods. *BMJ* 2008; **337**: a2473.

160 Anderson R. New MRC guidance on evaluating complex interventions. *BMJ* 2008; **337**: a1937.

161 Donabedian A. Evaluating the quality of medical care. 1966. *Milbank Q* 2005; **83**: 691–729.

162 Basch CE, Sliepcevich EM, Gold RS, Duncan DF, Kolbe LJ. Avoiding type III errors in health education program evaluations: a case study. *Health Educ Q* 1985; **12**: 315–31.

163 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; published online May 23. DOI:10.1093/aje/kwr061.

164 Glynn JR, Bauer J, de Boer AS, *et al.* Interpreting DNA fingerprint clusters of Mycobacterium tuberculosis. European Concerted Action on Molecular Epidemiology and Control of Tuberculosis. *Int J Tuberc Lung Dis Off J Int Union Tuberc Lung Dis* 1999; **3**: 1055–60.

165 Hanekom M, van der Spuy GD, Gey van Pittius NC, *et al.* Discordance between mycobacterial interspersed repetitive-unit-variable-number tandem-repeat typing and IS6110 restriction fragment length polymorphism genotyping for analysis of Mycobacterium tuberculosis Beijing strains in a setting of high incidence of tuberculosis. *J Clin Microbiol* 2008; **46**: 3338–45.

166 Supply P, Allix C, Lesjean S, *et al.* Proposal for standardization of optimized mycobacterial interspersed repetitive unit-variable-number tandem repeat typing of Mycobacterium tuberculosis. *J Clin Microbiol* 2006; **44**: 4498–510.

167 Hunter PR, Gaston MA. Numerical index of the discriminatory ability of typing systems: an application of Simpson's index of diversity. *J Clin Microbiol* 1988; **26**: 2465–6.

168 WHO | TB data. WHO. http://www.who.int/tb/country/en/index.html (accessed Dec 12, 2012).

169 Kruijshaar ME, Pimpin L, Abubakar I, *et al.* The burden of TB-HIV in the EU: how much do we know? A survey of surveillance practices and results. *Eur Respir J* 2011; **38**: 1374–81.

170 World Health Organization. WHO Tuberculosis Country Profiles. http://www.who.int/tb/country/data/profiles/en/.

171 Freeman MF, Tukey JW. Transformations Related to the Angular and the Square Root. *Ann Math Stat* 1950; **21**: 607–11.

172 Gurjav U, Jelfs P, McCallum N, Marais BJ, Sintchenko V. Temporal dynamics of Mycobacterium tuberculosis genotypes in New South Wales, Australia. *BMC Infect Dis* 2014; **14**: 455–455.

173 Asgharzadeh M, Kafil HS, Roudsary AA, Hanifi GR. Tuberculosis transmission in Northwest of Iran: using MIRU-VNTR, ETR-VNTR and IS6110-RFLP methods. *Infect Genet Evol J Mol Epidemiol Evol Genet Infect Dis* 2011; **11**: 124–31.

174 Allix-Béguec C, Supply P, Wanlin M, Bifani P, Fauville-Dufaux M. Standardised PCR-based molecular epidemiology of tuberculosis. *Eur Respir J Off J Eur Soc Clin Respir Physiol* 2008; **31**: 1077–84.

175 Tuite AR, Guthrie JL, Alexander DC, *et al.* Epidemiological evaluation of spatiotemporal and genotypic clustering of mycobacterium tuberculosis in Ontario, Canada. *Int J Tuberc Lung Dis* 2013; **17**: 1322–7.

176 Guang-ming D, Zhi-guo Z, Peng-ju D, *et al.* Differences in the population of genetics of Mycobacterium tuberculosis between urban migrants and local residents in Beijing, China. *Chin Med J (Engl)* 2013; **126**: 4066–71.

177 Zmak L, Obrovac M, Katalinic Jankovic V. First insights into the molecular epidemiology of tuberculosis in Croatia during a three-year period, 2009 to 2011. *Scand J Infect Dis* 2014; **46**: 123–9.

178 Tessema B, Beer J, Merker M, *et al.* Molecular epidemiology and transmission dynamics of Mycobacterium tuberculosis in Northwest Ethiopia: new phylogenetic lineages found in Northwest Ethiopia. *BMC Infect Dis* 2013; **13**: 131.

179 Smit PW, Haanpera M, Rantala P, *et al.* Molecular Epidemiology of Tuberculosis in Finland, 2008-2011. *Plos One* 2013; **8**: e85027.

180 Roetzer A, Schuback S, Diel R, *et al.* Evaluation of Mycobacterium tuberculosis typing methods in a 4-year study in Schleswig-Holstein, Northern Germany. *J Clin Microbiol* 2011; **49**: 4173–8.

181 Ojo OO, Sheehan S, Corcoran DG, *et al.* Molecular epidemiology of Mycobacterium tuberculosis clinical isolates in Southwest Ireland. *Infect Genet Evol J Mol Epidemiol Evol Genet Infect Dis* 2010; **10**: 1110–6.

182 Aleksic E, Merker M, Cox H, *et al.* First Molecular Epidemiology Study of Mycobacterium tuberculosis in Kiribati. *Plos One* 2013; **8**: e55423.

183 Dymova MA, Liashenko OO, Poteiko PI, Krutko VS, Khrapov EA, Filipenko ML. Genetic variation of Mycobacterium tuberculosis circulating in Kharkiv Oblast, Ukraine. *BMC Infect Dis* 2011; **11**: 77.

184 Varghese B, Supply P, Shoukri M, *et al.* Tuberculosis Transmission among Immigrants and Autochthonous Populations of the Eastern Province of Saudi Arabia. *Plos One* 2013; **8**: e77635.

185 Lim LK-Y, Sng LH, Win W, *et al.* Molecular Epidemiology of Mycobacterium tuberculosis Complex in Singapore, 2006-2012. *Plos One* 2013; **8**: e84487.

186 Alonso-Rodriguez N, Martínez-Lirola M, Sánchez ML, *et al.* Prospective universal application of mycobacterial interspersed repetitive-unit-variable-number tandem-repeat genotyping to characterize Mycobacterium tuberculosis isolates for fast identification of clustered and orphan cases. *J Clin Microbiol* 2009; **47**: 2026–32.

187 Jonsson J., Hoffner S., Berggren I., *et al.* Comparison between RFLP and MIRU-VNTR genotyping of mycobacterium tuberculosis strains isolated in stockholm 2009 to 2011. PLoS ONE. 2014.

188 Muwonge A, Malama S, Johansen TB, *et al.* Molecular Epidemiology, Drug Susceptibility and Economic Aspects of Tuberculosis in Mubende District, Uganda. *PLoS ONE* 2013; **8**.

189 Sails AD, Barrett A, Sarginson S, *et al.* Molecular epidemiology of Mycobacterium tuberculosis in East Lancashire 2001-2009. *Thorax* 2011; **66**: 709–13.

190 Mandal S, Bradshaw L, Anderson LF, *et al.* Investigating transmission of Mycobacterium bovis in the United Kingdom in 2005 to 2008. *J Clin Microbiol* 2011; **49**: 1943–50.

191 Hamblion EL, Wynne-Edwards E, Anderson C, Anderson SR. A summary of strain typing and clustering of TB in London in 2010 and an analysis of the associated risk factors. *Thorax* 2011; **66**: A88–9.

192 Evans J. Analysis of prevalent Mycobacterium tuberculosis strains in the United Kingdom: detection, distribution and expansion of MIRU-VNTR profiles containing high numbers of isolates. In: European Society of Clinical Microbiology and Infectious Diseases. Vienna, Austria, 2010.

193 Nikolayevskyy VV, Brown TJ, Bazhora YI, Asmolov AA, Balabanova YM, Drobniewski FA. Molecular epidemiology and prevalence of mutations conferring rifampicin and isoniazid resistance in Mycobacterium tuberculosis strains from the southern Ukraine. *Clin Microbiol Infect Off Publ Eur Soc Clin Microbiol Infect Dis* 2007; **13**: 129–38.

194 Hang NTL, Maeda S, Lien LT, *et al.* Primary Drug-Resistant Tuberculosis in Hanoi, Viet Nam: Present Status and Risk Factors. *PLoS ONE* 2013; **8**: UNSP e71867.

195 Dymova MA, Kinsht VN, Cherednichenko AG, Khrapov EA, Svistelnik AV, Filipenko ML. Highest prevalence of the Mycobacterium tuberculosis Beijing genotype isolates in patients newly diagnosed with tuberculosis in the Novosibirsk oblast, Russian Federation. *J Med Microbiol* 2011; **60**: 1003–9.

196 Glynn JR, Vyonycky E, Fine PEM. Influence of Sampling on Estimates of Clustering and Recent Transmission of Mycobacterium tuberculosis Derived from DNA Fingerprinting Techniques. *Am J Epidemiol* 1999; **149**: 366–71.

197 TB Strain Typing Project Board HPA. TB Strain Typing Cluster Investigation Handbook for Health Protection Units 1st Edition. 2011; published online Sept. https://hpaintranet.hpa.org.uk/Content/ProgrammesProjects/HPAProgrammes/HPAKeyHealthProtectionProgrammes/Respiratory/TB/StrainTyping/ (accessed Nov 30, 2011).

198 Schürch AC, van Soolingen D. DNA fingerprinting of Mycobacterium tuberculosis: From phage typing to whole-genome sequencing. *Infect Genet Evol J Mol Epidemiol Evol Genet Infect Dis* 2011; published online Nov 3. DOI:10.1016/j.meegid.2011.08.032.

199 Walker TM, Monk P, Grace Smith E, Peto TEA. Contact investigations for outbreaks of Mycobacterium tuberculosis: advances through whole genome sequencing. *Clin Microbiol Infect Off Publ Eur Soc Clin Microbiol Infect Dis* 2013; published online Feb 21. DOI:10.1111/1469-0691.12183.

200 Pawson R. Evidence-Based Policy: A Realist Perspective. SAGE, 2006.

201 Davis FD. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Q* 1989; **13**: 319–40.

202 Mears J, Abubakar I, Crisp D, *et al.* Prospective evaluation of a complex public health intervention: lessons from an initial and follow-up cross-sectional survey of the tuberculosis strain typing service in England. *BMC Public Health* 2014; **14**: 1023.

203 Van Buynder P. Enhanced surveillance of tuberculosis in England and Wales: circling the wagons? *Commun Dis Public Health PHLS* 1998; **1**: 219–20.

204 Health Protection Agency Centre for Infections. Tuberculosis in the UK: Annual report on tuberculosis surveillance in the UK, 2010. 2010 http://www.hpa.org.uk/web/HPAweb&HPAwebStandard/HPAweb_C/1287143581697 (accessed Nov 24, 2011).

205 Office for National Statistics. Population Estimates for UK, England and Wales, Scotland and Northern Ireland, Mid-2001 to Mid-2010 Revised. http://www.ons.gov.uk/ons/rel/pop-estimate/population-estimates-for-uk--england-and-wales--scotland-and-northern-ireland/mid-2001-to-mid-2010-revised/index.html (accessed July 24, 2014).

206 Ritchie J, Lewis J, editors. Qualitative Research Practice: A Guide for Social Science Students and Researchers. London; Thousand Oaks, Calif.: SAGE Publications Ltd, 2003.

207 Health Protection Agency West Midlands Public Health Laboratory. Protocol for Laboratory, Clinical and Public Health Investigation and Management of Possible False Positive Microscopy or Cultures for Mycobacteria. 2009.

208 Vynnycky E, Sumner T, Fielding KL, *et al.* Tuberculosis control in South African goldmines: mathematical modelling of a trial of community-wide isoniazid preventive therapy. *Press*.

209 Office for National Statistics. Migration Indicators Tool. 2010.

210 National Institute for Health and Clinical Excellence. Methods for development of NICE public health guidance. 2009. http://www.nice.org.uk/phmethods2009 (accessed Dec 14, 2011).

211 Department of Health. NHS Reference Costs 2010-11 Collection Guidance. 2010; published online Dec 16. http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_122803 (accessed Nov 30, 2011).

212 PSSRU | Unit Costs of Health and Social Care 2011. http://www.pssru.ac.uk/project-pages/unit-costs/2011/index.php (accessed April 13, 2012).

213 Crofts JP, Pebody R, Grant A, Watson JM, Abubakar I. Estimating tuberculosis case mortality in England and Wales, 2001-2002. *Int J Tuberc Lung Dis Off J Int Union Tuberc Lung Dis* 2008; **12**: 308–13.

214 Brown T, Evans JT, Sails AD, *et al.* HPA Mycobacterium tuberculosis Strain Typing: A guide to data production and distribution. 2012; published online Aug.

215 TB Strain Typing Project Board HPA. The HPA National TB Strain Typing Project Q&A Sheet. 2011; published online Feb. https://hpaintranet.hpa.org.uk/Content/ProgrammesProjects/HPAProgrammes/HPAKeyHealthProtectionProgrammes/Respiratory/TB/StrainTyping/ (accessed Nov 30, 2011).

216 CDC. Tuberculosis Genotyping in the United States, 2004-2010. 2012; published online June.

217 Theaker S, TB Strategy Board, Leicester. Tuberculosis Summary Needs Assessment: Leicester, Leicestershire and Rutland. 2013; published online Dec.

218 Ghosh S, Moonan PK, Cowan L, Grant J, Kammerer S, Navin TR. Tuberculosis Genotyping Information Management System: Enhancing Tuberculosis Surveillance in the United States. *Infect Genet Evol* 2012; **12**: 782–8.

219 Pareek M, Abubakar I, White PJ, Garnett GP, Lalvani A. Tuberculosis screening of migrants to low-burden nations: insights from evaluation of UK practice. *Eur Respir J Off J Eur Soc Clin Respir Physiol* 2011; **37**: 1175–82.

220 Comstock GW, Ferebee SH, Hammes LM. A controlled trial of community-wide isoniazid prophylaxis in Alaska. *Am Rev Respir Dis* 1967; **95**: 935–43.

221 Ferebee SH. Controlled chemoprophylaxis trials in tuberculosis. A general review. *Bibl Tuberc* 1970; **26**: 28–106.

222 Anderson C, White J, Abubakar I, *et al.* Raising standards in UK TB control: introducing cohort review. *Thorax* 2013; published online July 18. DOI:10.1136/thoraxjnl-2013-203751.

223 Inamdar L, Kumbang J, Ingham R, Gajraj R, Mohammed H, Frizzell R. Birmingham or Bradford - TB Strain Typing Cluster Investigations are Resource Intensive. In: 5 Nations Health Protection Conference. Ireland, 2013.

224 Kumbang J, Inamdar L, Gajraj R, Coetzee N, Ingham R, Muhammad H. Challenges of strain type TB cluster investigation in Birmingham and Bradford. In: Health Protection Conference 2012. Warwick, UK, 2012.

225 Mears J, Vynnycky E, Lord J, *et al.* The prospective evaluation of the TB strain typing service in England: a mixed methods study. *Thorax* 2015; published online April 16. DOI:10.1136/thoraxjnl-2014-206480.

227 Wellcome Trust Health Innovation Challenge Fund. http://www.wellcome.ac.uk/Funding/Innovations/Funded-projects/Health-Innovation-Challenge-Fund/index.htm (accessed Feb 18, 2015).

228 Department of Health. 100,000 Genomes Project. www.genomicsengland.co.uk/100k-genome-project (accessed June 4, 2014).

229 Squire SB, Ramsay ARC, van den Hof S, *et al.* Making innovations accessible to the poor through implementation research [State of the art series. Operational research. Number 7 in the series]. *Int J Tuberc Lung Dis* 2011; **15**: 862–70.

230 Habicht JP, Victora CG, Vaughan JP. Evaluation designs for adequacy, plausibility and probability of public health programme performance and impact. *Int J Epidemiol* 1999; **28**: 10–8.

231 National Institute for Health and Clinical Excellence. Tuberculosis: NICE guideline for hard-to-reach groups. http://publications.nice.org.uk/identifying-and-managing-tuberculosis-among-hard-to-reach-groups-ph37 (accessed Sept 20, 2012).

232 Pareek M, Watson JP, Ormerod LP, *et al.* Screening of immigrants in the UK for imported latent tuberculosis: a multicentre cohort study and cost-effectiveness analysis. *Lancet Infect Dis* 2011; **11**: 435–44.

233 Horton R. Offline: What have economists ever done for global health? *The Lancet* 2014; **383**: 1024.

234 Mann G, Squire SB, Bissell K, *et al.* Beyond accuracy: creating a comprehensive evidence base for TB diagnostic tools [State of the art]. *Int J Tuberc Lung Dis* 2010; **14**: 1518–24.

235 NICE. Tuberculosis: clinical diagnosis and management of tuberculosis, and measures for its prevention and control. Costing report. 2006; published online March.

236 Atun R. Health systems, systems thinking and innovation. *Health Policy Plan* 2012; **27**: iv4–8.

237 Public Health England. National TB Strategy 2014-2019 Consultation. http://www.hpa.org.uk/Publications/InfectiousDiseases/Tuberculosis/1403TBstrategyconsultation2014/ (accessed July 1, 2014).

238 Didelot X, Bowden R, Wilson DJ, Peto TEA, Crook DW. Transforming clinical microbiology with bacterial genome sequencing. *Nat Rev Genet* 2012; **13**: 601–12.

Appendices

Appendix 1: Systematic review search strategy

Appendix 2: Initial and follow-up cross-sectional survey questions

Appendix 3: False positive TB isolation follow-up surveys to source laboratories

Appendix 4: Rate of cluster growth equations

Appendix 5: Transmission model equations

Appendix 6: Cost-effectiveness model QALY estimates

Appendix 7: Publications and conference presentations

Appendix 1: Systematic review search strategy (Medline/Embase)

1. (tubercle adj3 (bacillus or bacilli)).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]
2. ((mycobacterium or mycobacteria) adj3 (bovis or africanum or microti or canetti)).mp.
3. exp tuberculosis/ or mycobacterium tuberculosis/ or tuberculosis.mp. or tb.mp. or Mtb.mp. or "M tuberculosis complex".mp.
4. or/1-3
5. Minisatellite Repeats/ or Genotype/ or Interspersed Repetitive Sequences/ or DNA Fingerprinting/ or Bacterial Typing Techniques/
6. "miru".mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]
7. "vntr".mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]
8. (miru adj3 vntr).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]
9. (mycobacterial adj3 interspersed adj3 repetitive adj3 units).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]
10. (dna adj3 fingerprinting).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]
11. ((strain adj3 type) or (strain adj3 typing) or (strain adj3 types)).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]
12. ((molecular adj3 typing) or (molecular adj3 strain adj3 typ*)).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]
13. (genotype or genotyping or genotypes).ti,ab.
14. (minisatellite adj3 repeat*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]

15. molecular epidemiology/mt or (molecular adj3 epidemiology).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]
16. or/5-15
17. exp disease outbreaks/ or (outbreak adj3 analysis).mp. or (outbreak adj3 investigation).mp. or (outbreak adj3 management).mp. or (tuberculosis adj3 outbreak).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]
18. exp contact tracing/ or (contact adj3 tracing).mp. or (contact* adj3 traced).mp. or (contact adj3 screen*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]
19. exp case management/ or (case adj3 management).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]
20. exp Risk Factors/
21. (risk adj3 factor*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]
22. exp Epidemiologic Factors/
23. infectious disease transmission.mp. or exp Disease Transmission, Infectious/
24. exp case management/ or (case adj3 management).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]
25. program evaluation/ or evaluation studies as topic/ or (program adj3 evaluation).mp. or (programme adj3 evaluation).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]
26. public health practice/ or (public adj3 health).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]
27. ((tuberculosis adj3 control) or (tb adj3 control)).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]
28. (molecular adj3 surveillance).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]

29. exp cluster analysis/ or (cluster* adj3 rate*).mp. or (cluster* adj3 growth).mp. or (cluster* adj3 analysis).mp. or (cluster adj3 investigation).mp. or (proportion adj3 cluster*).mp. or (molecular adj3 cluster*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]
30. ((recent adj3 transmission) or (transmission adj3 event*) or (transmission adj3 rate*) or (chain adj3 transmission) or (transmission adj3 setting*)).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]
31. or/17-30
32. 4 and 16
33. 32 and 31
34. limit 33 to yr="1998-Current"
35. limit 34 to english language
36. animals/
37. humans/
38. 36 not 37
39. 35 not 38

Appendix 2: Initial and follow-up survey questions

Survey Questions

Questions that are numbered without a letter were asked to all participants. Where a letter follows the question number the question was asked to that particular group (x=nurses, y=physicians, z=public health staff). The questions are divided by target group for Sections C and D.

TB Control and Strain Typing: A Questionnaire for Nurses / Physicians / HPU staff

The Health Protection Agency has commissioned an evaluation of the National TB Strain Typing Service (STS). We are interested in how much time nurses / physicians / staff from HPUs spend on cluster investigations and their uptake and experience of using strain typing data. This is the second part of a two-part questionnaire (the first part was conducted in November 2010). All the information collected in this form will be anonymised and stored according to the Data Protection Act; we do ask for your name and contact details so we can link this to your response - oryour predecessor's response - to the first part of the questionnaire.

THANK YOU FOR AGREEING TO TAKE PART IN THIS QUESTIONNAIRE

You can save the questionnaire and return to it another time by clicking "save" at the bottom of any page and following the instructions.

If you have any queries please do not hesitate to contact Jessica Mears, evaluation scientist - jessica.mears@hpa.org.uk, 0208 327 6265

Section A -- This section asks basic information about where you work

1. Name
2. Email address
3. Telephone number

4. Job title

5.

    x. Which Health Protection Unit does your workplace fall under? (multiple selections available) *List of 26 HPUs*

    y. Which Health Protection Unit does your workplace fall under? (multiple selections available) *List of 26 HPUs*

    z. Do you work at the national level, regional level, or Health Protection Unit level (or other)? *National / Regional / HPU / other*

Section B -- The following questions are about your current awareness and use of TB strain typing

1. Have you heard of the National TB Strain Typing Service (aside from this questionnaire)? *Y/N*

2. How do you access strain typing data? (multiple selections possible) *Don't get data / Send request for specific cases / Automatically receive for each case / Contact labs directly / Cluster reports from lab / Cluster report/phone call from HPU / Web-based ETS/LTBR / Contacted by a cluster investigator / Other*

*If you have answered "*Don't get data"* in Section B Q2, skip to BQ7

3.

    x. - How many weeks after a sample is sent to the lab do you receive the individual strain type? (weeks)

    y. - How many weeks after a sample is sent to the lab do you receive the individual strain type? (weeks)

    z. Do you usually access/get hold of strain typing data before or after you have finished contact tracing? *Before / After*

4. What do you use strain typing for? (multiple selections possible) *Don't know / Identify clusters and links between cases / Disprove clusters and links between cases / Justify extended contact tracing / Justify stopping contact tracing / To provide more information / Other (please specify)*

5. How often do you use strain typing data in your case management or outbreak investigation? *Never / For few cases / For about half of cases / For many cases / For every case*

Any comments you have about how often you use the strain typing data would be helpful

6. Do you find the strain typing information useful? *Very useful / Quite useful / Not very useful / Useless*

Any comments you have about how useful you find the strain typing data would be helpful

*Skip to BQ8

7. If you had access to strain typing data, what would you use it for? (multiple selections possible) *Don't know / Identify clusters and links between cases / Disprove clusters and links between cases / Justify extended contact tracing / Justify stopping contact tracing / To provide more information / Other (please specify)*

8. Have you received any training on using strain typing data? *Yes (please specify) / No*

9. Do you have access to any resources or tools to help interpret and use strain typing information? (multiple selections possible) *No / Handbook / Training workshops / Colleagues / Webcast/online training / Other (please specify)*

10. How would you rate your knowledge of how to interpret MIRU-VNTR data? *(no knowledge) 1 / 2 / 3 / 4 / 5 (Excellent knowledge)*

Section C -- The following questions are about your current TB-related workload

CX: Questions for Nurses

1. What are your working hours? *Full time / Part time / Other (please specify)*
2. Do you work on anything other than TB? *Yes / No*

*If you have answered "No" in Section CX Q2, skip to CXQ4

3. What proportion of your average week do you spend on TB-related work? *%*
4. Are you involved in contact tracing? *Yes / No*

*If you have answered "No" to CXQ4, skip to Section DX

We are now going to ask a few questions about contact tracing and the contacts of TB patients.

5. How do you do your contact tracing work? *Dedicated time in clinic (% of contact tracing done in this way) / Telephone interviews (%) / Home visit (%) / Other (please specify)*
6. In total, how many contacts have you screened in the last month? (using any of these methods: Mantoux, X-ray, IGRA, symptom check) (Number of people)
7. How many hours did this take? (Hours)
8. In the last month, what proportion (%) of your total working time have you spent on tracing and screening contacts in and out of clinic? (% of time)

We are now going to ask you some questions about newly diagnosed TB patients.

9. How many new active TB cases have you managed in the last full month? (cases)
10. How many of these required contact tracing? (cases)

This section asks you about 5 recent index cases for whom contact tracing has been completed. It might help to have your case management notes in front of you before starting this section.

11. Please fill in the following table based on 5 recent index cases whose contacts have been screened:

	Case 1	Case 2	Case 3	Case 4	Case5
ETS / LTBR No.					
Age / year of birth					
Site of disease	*Pulmonary / extra-pulm*	*Pulmonary / extra-pulm*	*Pulmonary / extra-pulm*	*Pulmonary / extra-pulm*	*Pulmonary / extra-pulm*
Smear test result	*Positive / Negative / Unknown*	*Positive / Negative / Unknown*	*Positive / Negative / Unknown*	*Positive / Negative / Unknown*	*Positive / Negative / Unknown*
Culture result	*Positive / Negative / Unknown*	*Positive / Negative / Unknown*	*Positive / Negative / Unknown*	*Positive / Negative / Unknown*	*Positive / Negative / Unknown*
When was the strain type available to you?	*Before contact investigation / During contact investigation / After contact investigation / Never*	*Before contact investigation / During contact investigation / After contact investigation / Never*	*Before contact investigation / During contact investigation / After contact investigation / Never*	*Before contact investigation / During contact investigation / After contact investigation / Never*	*Before contact investigation / During contact investigation / After contact investigation / Never*
No. contacts identified					
No. contacts screened					
No. contacts with active TB disease					
No. contacts with latent TB infection					

CY: Questions for Physicians

1. What are your working hours? *Full time / Part time / Other (please specify)*
2. What kind of physician are you?
3. What proportion of your average week do you spend on TB-related work? (%)
4. On average throughout the year, how many new TB cases do you personally start on treatment per month? (cases)
5. How often are you called to an incidence meeting to discuss contact tracing/screening? *Once a month or more / Once every 1-3 months / Once every 3-6 months / Less than once every 6 months / Never*
6. Has the strain type of the case under discussion ever been relevant in an incident meeting? *Yes / No*

If yes, please could you give some more details about the incident(s)

CZ: Questions for Public Health Staff

1. What are your working hours? *Full time / Part time / Other (please specify)*
2. Do you work on anything other than TB? *Yes / No*

*If you answered "No" to Section CZ Q2 skip to CZQ5

3. What proportion of your average week do you spend on TB-related work? (%)
4. On average, how many hours per week do you spend on TB-related work? (Hours)
5. In the last 3 months, how many TB cases have been reported in your workplace (HPU, region or nationally)? (cases)
6. In the last 3 months, how many extended contact investigations were initiated because the cases were known contacts (regardless of strain type)? (extended contact investigations)

7. In how many of these was strain typing used to provide more information?

8. In how many of these investigations – investigations that were initially opened because cases were known contacts of each other – did strain typing influence your investigation? *Number of investigations lengthened or reopened / Number of investigations stopped*

If possible, please provide more information about these incidents.

9. In the last 3 months, how many cluster investigations were initiated because of strain typing data (i.e. where links between cases were not identified through initial contact tracing)? (cluster investigations)

10. In how many of these investigations that were initiated because of the strain typing data were epidemiological links found between cases following the cluster investigation?

If possible, please provide more information about these incidents.

11. In the last month, what proportion (%) of your time have you spent on cluster investigation activities (investigating epidemiologically linked or strain type-linked cases)? (%)

12. If you have any other comments about how you use (or don't use) strain typing data and how useful you find it please write in the box below.

Section D -- The following questions are about investigating transmission

DX: Questions for Nurses

In the last year have you investigated any clusters of 2 or more TB cases because…

1. …they were contacts of each other (not because of strain type)? *Yes / No*

Comments welcome

2. …they were linked by strain type along? *Yes / No*

Comments welcome

3. In the last year have you stopped investigating possible transmission because the cases involved had different strain types? *Yes / No*

If yes, please could you give some more details about the incident.

4. If you have any other comments about how you use (or don't use) strain typing data and how useful you find it please write in the box below.

DY: Questions for Physicians

In the questions below, please estimate the number of TB investigations you have been involved in during the last 12 months.

1. In the last year, have you investigated any clusters of 2 or more TB cases because they were contacts of each other (not because of strain type)? *Yes / No*

If yes, approximately how many investigations were launched based on information gained through contact tracing only?

2. In the last year have you investigated any clusters of 2 or more TB cases because they were linked by strain type along? *Yes / No*

If yes, approximately how many investigations were launched based on strain type alone?

3. In the last year have you stopped investigating possible transmission because the cases involved had different strain types? *Yes / No*

If yes, how many incidents does this apply to? Where possible, please provide more information about the incident(s)

End

Appendix 3: False positive TB isolation follow-up surveys to source laboratories

Below are email questionnaires that were sent to source laboratories regarding A) possible, B) unconfirmed, and C) unlikely cross contamination events.

A) Possible cross contamination event

Evaluation of the National TB Strain Typing Service: query

To:
Cc:

Dear XXX,

My name is Jessica and I am the scientist for the evaluation of the National TB Strain Typing Service (TB-STS). This prospective evaluation was included in the original three-year STS project to assess the effectiveness and cost-effectiveness of the service.

We are evaluating the effectiveness of the TB-STS (i.e. the strain typing carried out by the national and regional mycobacterial reference laboratories) at identifying false positive TB isolates which will form a critical component of the cost-effectiveness model of the evaluation of the TB-STS.

We would like to follow up on xxx reported possible cross contamination event(s):

Reference laboratory notified you by letter dated xx/xx/xxxx of a possible cross contamination incident involving xxx TB cases (source lab numbers: xxx (ref lab number xxx) and xxx (ref lab number xxx)). According to the reference laboratory records this was a likely case of cross contamination. I would be grateful if you would answer the following questions about the incident. This information will only be used for the evaluation of the TB-STS; it will not be part of any general service evaluation.

	Please mark either Yes or No with 'x'	Yes	No
1	Were you aware of this potential false positive isolation before receiving the letter mentioned above?		
2	What impact did this incidence have on the patient? Had they started treatment? Had contact screening already been carried out? Other: *Write here*		
3	Has this incident led to improvements in your laboratory protocol? *Please specify here*		
4	Has the strain typing helped you to identify any other false positive isolation incidents?		

B) Unconfirmed cross contamination event

Evaluation of the National TB Strain Typing Service: query

To:
Cc:

Dear XXX,

My name is Jessica and I am the scientist for the evaluation of the National TB Strain Typing Service (TB-STS). This prospective evaluation was included in the original three-year STS project to assess the effectiveness and cost-effectiveness of the service.

We are evaluating the effectiveness of the TB-STS (i.e. the strain typing carried out by the national and regional mycobacterial reference laboratories) at identifying false positive TB isolates which will form a critical component of the cost-effectiveness model of the evaluation of the TB-STS.

We would like to follow up on a reported possible cross contamination event:

The reference laboratory notified you by letter on the xx/xx/xxxx of a possible cross contamination incident involving xxx TB cases (source lab numbers: xxxx (ref lab number xxx) and xxx (ref lab number xxx)). According to the reference laboratory records this has not yet been confirmed as a likely cross contamination incident or not. I would be grateful if you would answer the following questions about the incident. This information will only be used for the evaluation of the TB STS; it will not be part of any general service evaluation.

	Please mark Yes or No with 'x'	Yes	No
1	Were you aware of this potential false positive isolation before receiving the letter mentioned above?		
2	Have you carried out an investigation into this incident?		
3	What were the conclusions of your investigation? Likely cross contamination Unlikely cross contamination Other: *Write here*		

C) Unlikely cross contamination event

Evaluation of the National TB Strain Typing Service: query

To:
Cc:

Dear XXX,

My name is Jessica and I am the scientist for the evaluation of the National TB Strain Typing Service (TB-STS). This prospective evaluation was included in the original three-year STS project to assess the effectiveness and cost-effectiveness of the service.

We are evaluating the effectiveness of the TB-STS (i.e. the strain typing carried out by the national and regional mycobacterial reference laboratories) at identifying false positive TB isolates which will form a critical component of the cost-effectiveness model of the evaluation of the TB-STS.

We would like to follow up on a reported possible cross contamination event:

The reference laboratory notified you by letter on the xx/xx/xxxx of a possible cross contamination incident involving xxx TB cases (source lab numbers: xxx (ref lab number xxx) and xxx (ref lab number xxx)). According to the reference laboratory records this was not a cross contamination incident. I would be grateful if you would answer the following questions about the incident. This information will only be used for the evaluation of the TB STS; it will not be part of any general service evaluation.

	Please mark either Yes or No with 'x'	Yes	No
1	Were you aware of this potential false positive isolation before receiving the letter mentioned above?		
2	If yes, had you already investigated the incident?		
3	Have you identified any other false positive isolation incidents with the aid of strain typing? *Please specify here*		

Appendix 4: Rate of cluster growth equations

I prepared the data for analysis and the univariate and multivariate linear regression analyses were conducted by Ted Cohen and Leonid Chindelevitch.

—

Data from the cluster monitoring database on the number of cases in a cluster, the date cases were added to a cluster and the date cluster investigations started were merged with the clinical and demographic data collected on ETS for the years 2010 and 2011. After cleaning the data and excluding the first case in each of the 113 clusters, there were a total of 949 cases. The final set of variables used in the multivariate linear regression were: sex, UK-born, age group (treated as an ordinal variable), site of disease, case order (the order in which cases were added to the cluster), after (our indicator of cases discovered after a cluster investigation was started), and cluster level (local, regional and national) and lineage, the latter being factor variables.

It was assumed that each cluster had an intrinsic growth rate, which depends only on the scale of the cluster (either local, regional or national) and the lineage, and that it can only be altered by the opening of a cluster investigation. To fully specify the model, f is defined as a function mapping $i \in [K]$ to the scale and lineage of the ith cluster. For the jth case in the ith cluster, the regression variables are:

$f_{ij} = 1 \rightarrow$ the case is female

$n_{ij} = 1 \rightarrow$ the case is not UK born

$e_{ij} = 1 \rightarrow$ the case has extrapulmonary TB

$c_{ijm} = 1 \rightarrow$ the case was the mth case added to the cluster

$a_{ijk} = 1 \rightarrow$ the case belongs to the kth age group

$t_{ij} = 1 \rightarrow i \in S$ and the case is diagnosed after the investigation starts

The following regression equation was performed:

$$T_{fj} - T_{i(j-1)} \sim \tau_{f(i)} + \alpha t_{ij} + \beta f_{ij} + \gamma n_{ij} + \epsilon e_{ij} + \lambda_k a_{ijk} + \omega_m c_{ijm}$$,[a]

where

$\tau_{f(i)}$ is the rate of cluster growth with a scale and lineage $f(i)$

α is the effect of opening an investigation

β is the effect of gender

γ is the effect of being UK born or not UK born

ϵ is the effect of the site of TB

λ_k is the effect of being in age group k

$\omega_m$is the effect of being the mth case in a cluster

[a] L. Forsberg White and M. Pagano, "A Likelihood-Based Method for Real-Time Estimation of the Serial Interval and Reproductive Number of an Epidemic," *Statistics in Medicine* 27, no. 16 (2008): 2999–3016, doi:10.1002/sim.3136.

Appendix 5: Transmission model equations

The transmission model was based on a model previously developed by Emilia Vynnycky.[1,2] *The model was adapted by Emilia for this evaluation.*

—

Overview

Figure 24 in the thesis shows the general structure of the model. The model is age structured, with the population stratified into single year age groups and deterministic, describing what happens on average over time, using weekly time steps (see below for further details, the difference equations and input parameters). The model includes immigration and emigration and considers the following three epidemiological scenarios:

Scenario 1. Low incidence, comparable to that in the white UK population. The predicted TB incidence increases with increasing age, reaching about seven per 100,000 for those aged ≥55 years (figure 41A in the thesis), which is consistent with observed data (two to five and four to nine per 100,000 per year in 2011).[3] Here the infection risk is assumed to have declined since 1950[4] and has remained roughly constant since 1980, a small proportion (<10%) of those aged <55 years are assumed to have been infected, as compared with 50% on average of those aged ≥55 years (figure 41B in the thesis). The proportion of disease that is attributable to recent transmission decreases steadily with increasing age, reaching <10% for those aged ≥55 years (figure 41C in the thesis).

Scenario 2. Medium incidence, comparable to that in the non-white UK-born. The disease incidence is about 20 per 100,000 per year, as compared with nine to 55 per 100,000 in the observed 2009 data.[5] The annual risk of infection (ARI) is assumed to have been constant over time at 0.1% per year, with a low proportion of individuals who have been infected (average of <20% for those aged ≥55 years).

Scenario 3. High incidence, comparable to that in the non-UK born. The disease incidence is about 120 per 100,000 year, which is comparable to observed data

(notification rates of 59-273 per 100,000 in 2009, depending on the ethnic group). The ARI is assumed to have been constant over time at one percent per year, similar to that in some developing countries, with proportion of individuals who have been infected increasing with increasing age to reach an average of 20% for those aged ≥55 years**.**

For scenarios 2 and 3, the assumed in- and out- migration rates are eight and six per 1000 per year respectively, based on data from the period 2000-2010.[6] In-migrants are assumed to be aged 15-54 years; the assumed out-migration rate is identical for all ages. The TB prevalence among in-migrants is assumed to be 0.02%, which is consistent with the predicted prevalence in the model for an ARI of one percent per year. The TB incidence in these individuals in their native populations is similar to that shown in (figure 41A in the thesis) for an ARI of one percent per year, which is similar to that in the non-UK born population in the UK,[3] but slightly lower than that estimated among immigrants, shortly after entering the UK (320-400 per 100,000 in 1998).[7] Based on recent data, we assume that no cases are detected when entering the UK.[8] The model parameters are shown in Supplementary Table 1.

Supplementary Table 1 – Summary of assumed parameter values and their ranges.

Definition	Symbol	Base case value	Source/comment
Transmission			
Number of people effectively contacted by each smear-positive case in a) Low incidence (similar to white UK population) b) Medium incidence (Non-white, UK-born population) c) High incidence (Non-white, non-UK-born)	c_e		Calculated to reproduce incidence consistent with observed notification rates
Infectiousness of smear-negative TB cases, compared to that of smear-positives	f	22%	[9]
Force of infection at time t	$\lambda(t)$		See text
Disease onset			
Rate of disease onset following recent infection at time s_i since first infection among those not PT among of age a	$d_{p,z-,a}(s_i)$	Cumulative risk over 5 years: 4% (children) 14% (adults), increases linearly between ages 10 and 20 years	[10,1]
Rate of disease onset following recent infection at time s_i since first infection among those on PT among of age a	$d_{p,z+,a}(s_i)$	Calculated as $d_{p,z-,a}(s_i)\pi_{d,z+}$	
Rate of disease onset at time s_r following reinfection among those not on PT of age a	$d_{x,z-,a}(s_r)$	Cumulative risk over 5 years: 8%	[10,1]
Rate of disease onset at time s_r following reinfection among those on PT of age a	$d_{x,z+,a}(s_r)$	Calculated as $d_{x,z-,a}(s_i)\pi_{d,z+}$	
Annual rate of developing disease through reactivation (%/year) among those not on PT of age a	$d_{n,z-,a}$	0.03%/year	[10,1]
Annual rate of developing disease through reactivation (%/year) among those on PT of age a	$d_{n,z+,a}$	Calculated as $d_{n,z-,a}\pi_{d,z+}$	
Percentage of respiratory TB disease that is smear-positive among those of age a	$o_{s+,a}$	10% (children)	Public Health England (PHE)

		65% (adults)	Enhanced Surveillance database and data in [1]. Follows the age-specific pattern in [1].
Duration that people spend in the reinfected compartment (experiencing the risk of disease given reinfection before being transferred to the latent compartment	T_R	5 years	--
Duration that people spend in the infected compartment (experiencing the high risk of disease given infection before being transferred to the latent compartment	T_E	5 years	--
Demography			
Annual birth rate per 1000 per year		13.1	Office for National Statistics[6]
Annual general population mortality rates	$m_{tb-,a}$	Age-dependent	Office for National Statistics[6]
Inmigration rate		0.8%/year	Office for National Statistics[6]
Outmigration rate for those of age *a*	μ_a	0.6%/year	Office for National Statistics[6]
TB prevalence among immigrants		0.02%	Consistent with model predictions based on an ARI of 1%/year
Case detection			
% of immigrant TB cases with smear status *s* that are detected on entry to the UK	$p_{in,f,s}$	0%	[8]
Average time from disease onset to detection (among non-immigrants) at time *t*	$T_{detect}(t)$	10 weeks (before the start of the TB-STS); varied thereafter	PHE Enhanced Surveillance database
Average rate at which cases are found	$r_f(t)$	Calculated as $1/T_{detect}$	

Maximum duration of that people spend in the detected (found) compartment before being distributed to the undetected compartment, if they have not started TB treatment in the meantime.	$T_{f_{max}}$	6 months	--
Preventive treatment			
Proportion of infections that are detected at time *t*	$p_{i,det}(t)$	Varied between 3% and 13% per year	No data available. Assumed to differs before and after the start of the strain-typing service
Percentage of eligible contacts (TST/IGRA+ and aged <35 years) that start PT	$p_{z+,start}$	95%	Plausible value, based on national policy[11]
Proportion of infected people that start PT at time t	$i_{z+}(t)$	Calculated as: $p_{i,det}\ p_{z+,start}$	--
Protection provided by PT against disease whilst taking PT	$\pi_{d,z+}$	65%	[12,13]
Proportion of those taking PT who complete PT	$p_{z+,stop}$	85%	[11]
Rate at which those taking PT stop taking PT	i_{z-}	1.3%/week	Based on $p_{z+,stop}$ = 0.85
Maximum duration of PT	$T_{Z_{max}}$	3 months	--
Treatment			
Average time from detection to start of TB treatment	$T_{treat,\ start}$	2 weeks	--
Rate at which cases start TB treatment at time s_f following detection	$\tau(s_f)$	s_f< 4 weeks: 35%/week $s_f \geq$ 4 weeks: 0%/week	Calculated so that 82% of detected cases complete treatment (see text)
Percentage of detected cases that complete TB treatment		82%	PHE Enhanced Surveillance database
Percentage of detected cases who default from treatment		5.5%	PHE Enhanced Surveillance database
Mortality rate among TB cases (before and during TB treatment)	$m_{tb+,a}$	7%	PHE Enhanced Surveillance

			database
Percentage of detected cases that are lost to follow-up		5.5%	PHE Enhanced Surveillance database
Duration of TB treatment	$T_{\tau_{max}}$	26 weeks	--

The subscript z- and z+ refer to those not on PT and on PT respectively, a refers to the age group. The abbreviations sm- and sm+ refer to those who are smear-negative and smear-positive respectively.

Model assumptions

Individuals are assumed to be born uninfected and are infected at a rate $\lambda(t)$ (the force or risk of infection). The force of infection depends on the prevalence of infectious individuals and is calculated as the product of the prevalence of infectious individuals and the effective contact rate, c_e, defined as the average number of individuals effectively contacted by each infectious person per unit time. An effective contact is defined as one that is sufficient to lead to transmission if it occurs between an uninfected ("susceptible") person and an infectious person.[14] The effective contact rate is calculated so it leads to given values for the annual risk of infection (see below).

Following infection, individuals are assumed to face an increased rate of developing disease during the first five years after infection ("primary" disease), which decreases with time since infection, after which they can either experience disease through reactivation or following reinfection. The rates at which they develop disease through the various mechanisms are age-dependent and are identical to those estimated in previous work.[1] The rate of disease onset following reinfection is less than that following new "primary" infection, due to some immunity resulting from previous infection.[10,1]

As in previous versions of the model, the proportion of disease that is sputum smear or culture-positive (infectious) is assumed to increase with age, based on observed data.[1] For simplicity, females are not modelled explicitly in the model. For simplicity, the effects of HIV are also not modelled, given the low prevalence of HIV (2.4 per 1000) in England and Wales by 2008.[15]

Following disease onset, cases are assumed to be detected at a constant rate, with an average time to detection of 10 weeks. Given this relatively short time to detection, progression from smear negative to smear positive TB is not modelled explicitly.

Following detection, cases are assumed to start TB treatment after an average period of two weeks, so that the average time from disease onset to detection is 12 weeks, as observed in the strain typing data. 82% of those who start treatment are assumed to complete it, with the remainder dying (7%), defaulting from treatment (5.5%) or being lost to follow-up (5.5%).[16] Those who default from treatment are assumed to return to the undetected category and remain infectious. TB treatment is assumed to last a fixed period of six months. TB treatment is assumed to clear infection and individuals can develop disease subsequently only following reinfection. The rate at which they develop disease following reinfection is assumed to be identical to the rate at which those who have been infected for at least five years (described as those in the "latent" category in figure 24 in the thesis) develop disease following reinfection.

Based on observed data, 95% of those aged <35 years who are identified as having been infected, according to TST/IGRA, are assumed to start preventive treatment (PT) for three or six months, with 85% of these completing the full course.[8] National policy dictates that PT is not given to those under 35 years.[11] PT is assumed to provide 65% protection against disease whilst individuals are taking it.[12,13] Given complete compliance, the full course of PT is assumed to fully cure the infection, so that individuals can only develop disease subsequently following reinfection. It is also assumed that individuals who have either previously had TB treatment or PT would not be provided PT again.

In the absence of the TB-STS, a small percentage (3%) of all infected individuals is assumed to have been detected and treated each year. This proportion is unknown, but was probably very low, as implied by the number of tuberculin-positive contacts of tuberculosis cases that were identified for each tuberculosis cases that was investigated . For example, data on contact tracing activity suggested that after the introduction of the TB-STS, on average, about four contacts of each identified

tuberculosis case who was not in a cluster, was traced, with one of the contacts being tuberculin positive. Since approximately 9000 cases were reported in England in 2009,[5] this suggests that about 9000 tuberculin-positive people were identified. If the average prevalence of tuberculous infection in England is less than 10% and given a population of 55 million in England and Wales,[17] then the proportion of prevalent infections that is detected each year is likely to be less than 1%. An analogous calculation suggests that if the average prevalence of tuberculous infection in England was less than one percent, then the proportion of prevalent infections that is detected each year is likely to be less than two percent.

The amount by which the proportion of infections that were detected after the introduction of the TB-STS increased is also poorly understood. However, it is unlikely to have increased substantially, given that the number of contacts that were screened per TB case for cases who were in a cluster was similar to that for cases who were not in a cluster. We here assume that it increases by a factor of three, i.e. to 13% per year, which is likely to be close to or exceed the upper limit on the likely value.

The proportion of those eligible who take up preventive treatment, once detected, is also unknown, as is the proportion of those who start taking preventive treatment who complete it. We have assumed values of 95% (minimum and maximum values of 30% and 95% respectively) for the former and values of 85% (minimum and maximum values of 50% and 100% respectively) for the proportion of those starting preventive treatment who complete it. These values are plausible, and are consistent with those used in previous decision analyses,[18] although their accuracy is unclear. Studies of contact tracing activities in the USA from the period 1996-7 found that about 74% of tuberculin-positive positive contacts of tuberculosis cases started preventive treatment, with 56% completing it.[19] Similar data from the UK are limited. For example, studies have sometimes reported the numbers or proportions of contacts who started preventive treatment, without providing the numbers who were eligible or who completed preventive treatment.[20]

Model equations

The model was set up using weekly time steps using the difference equations below. The model was written using the C programming language. Supplementary Table 1 provides the main parameters and variables; Supplementary Table 2 summarizes the definitions of the compartments and variables in the model; any additional parameters are defined below.

Supplementary Table 2 – Definitions of the compartments and variables in the model.

Symbol	Definition
$U_a(t)$	Number of people of age a at time t who have never been infected.
$E_{z-,a}(t,s_i)$	Number of people of age a who have been infected for duration s_i at time t, who have never had PT.
$E_{z+,a}(t,s_i,s_z)$	Number of people of age a who have been infected for duration s_i and have been on PT for duration s_z at time t.
$E_{z_p,a}(t,s_i)$	Number of people of age a who have been infected for duration s_i at time t, who have previously had PT.
$L_{z-,a}(t)$	Number of people of age a in the latent category at time t, who have never had PT.
$L_{z+,a}(t,s_z)$	Number of people of age a in the latent category at time t, who have been on PT for duration s_z.
$P_{e+,a}(t)$	Number of people of age a who have previously had PT, cleared their infection and have not been reinfected since clearing their infection.
$P_{e-,a}(t)$	Number of people of age a who have had PT, have not cleared their infection and have not been reinfected during the previous five years
$R_{z-,a,}(t,s_r)$	Number of people of age a who have been reinfected for duration s_r at time t, who have never had PT.
$R_{z+,a}(t,s_r,s_z)$	Number of people of age a who have been reinfected for duration s_r and have been on PT for duration s_z at time t.
$R_{z_p,a}(t,s_r)$	Number of people of age a who have been reinfected for duration s_r at time t, who have previously had PT.
$D_{p,s,a}(t,s_o)$	Number of undetected cases of age a and smear status s who have had disease because of recent (primary) infection for duration s_o at time t, if $s_o < T_{o_{\max}}$. If $s_o = T_{o_{\max}}$, $D_{p,s,a}(t,s_o)$ represents the number of cases of age a, smear status s who have had disease because of recent (primary) infection for at least time $T_{o_{\max}}$ at time t

$D_{n,s,a}(t,s_o)$	Number of undetected cases of age *a* and smear status *s* who have had disease through (endogenous) reactivation for duration s_o at time *t*, if $s_o < T_{o_{\max}}$. If $s_o = T_{o_{\max}}$, $D_{n,s,a}(t,s_o)$ represents the number of cases of age *a*, smear status *s* who have had disease through (endogenous) reactivation for at least time $T_{o_{\max}}$.
$D_{x,s,a}(t,s_o)$	Number of undetected cases of age *a* and smear status *s* who have had disease because of (exogenous) reinfection for duration s_o at time *t*, if $s_o < T_{o_{\max}}$. If $s_o = T_{o_{\max}}$, $D_{x,s,a}(t,s_o)$ represents the number of cases of age *a*, smear status *s* who have had disease because of (exogenous) reinfection for at least time $T_{o_{\max}}$ at time *t*
$F_{s,a}(t,s_f)$	Number of cases of smear status *s*, age *a*, who have been detected ("found") for duration s_f at time *t* and have not yet started TB treatment.
$C_a(t,s_\tau)$	Number of cases of age *a*, who have been on TB treatment for duration s_τ at time *t*.
$V_{z-,a}(t)$	Number of people of age *a* who are in the recovered category at time *t* who are not on PT.
$V_{z+,a}(t,s_z)$	Number of people of age *a*, who are in the recovered category at time *t* who have been taking PT for duration s_z.
$M_{in,U,a}(t)$	Number of new immigrants at time *t*, who are of age *a*, and not infected
$M_{in,L,a}(t)$	Number of new immigrants at time *t*, who are of age *a*, and in the latent category.
$M_{in,E,a}(t,s_i)$	Number of new immigrants at time *t* who are of age *a*, and who have been newly infected for duration s_i
$M_{in,R,a}(t,s_r)$	Number of new immigrants at time *t* who are of age *a*, and who have been reinfected for duration s_r
$M_{in,D_p,s,a}(t,s_o)$	Number of new immigrants at time *t* who are of age *a*, who have been experiencing disease because of endogenous reactivation for duration s_o, and have smear status *s*.
$M_{in,D_n,s,a}(t,s_o)$	Number of new immigrants at time *t* who are of age *a*, who have been experiencing disease because of recent (primary) infection for duration s_o, and have smear status *s*.
$M_{in,D_x,s,a}(t,s_o)$	Number of new immigrants at time *t* of age *a*, who have been experiencing disease through exogenous reinfection for duration s_o, and have smear status *s*.
$M_{in,V,a}(t)$	Number of new immigrants at time *t* of age *a*, who have previously had TB, been treated and have not been reinfected since then.

People were allowed to experience the benefits of PT (i.e. reduced rates of disease onset) or lack of benefit in the same week as they started or stopped PT respectively. To simplify the equations whilst allowing this to occur, the population in the PT-related compartments was transferred into subsequent strata at the end of each time step, once other transitions had been accounted for.

Uninfected compartment

$$U_a(t+\delta t) = U_a(t)(1-\lambda(t)-\mu_a - m_{tb-,a}) \quad \text{Equation 1}$$

Recently (primary) infected compartment

Recently (primary) infected people who are not on PT

$$E_{z-,a}(t+\delta t,0) = (1-i_{z+}(t))\lambda(t)U_a(t)(1-d_{p,z-,a}(0)) \quad \text{Equation 2a}$$

$$E_{z-,a}(t+\delta t,s_i+\delta s_i) = (1-i_{z+}(t))E_{z-,a}(t,s_i)(1-d_{p,z-,a}(s_i)-\mu_a - m_{tb-,a}) + M_{in,E,a}(t,s_i) \quad \text{Equation 2b}$$

Recently (primary) infected people who are on PT

$$E_{z+,a}(t+\delta t,s_i+\delta s_i,0) = E_{z-,a}(t,s_i)i_{z+,a}(t)(1-d_{p,z+,a}(s_i)-\mu_a - m_{tb-,a}) \quad \text{Equation 3a}$$

$$E_{z+,a}(t+\delta t,0,s_z) = i_{z+,a}(t)\lambda(t)U_a(t) \quad \text{Equation 3b}$$

$$E_{z+,a}(t+\delta t,s_i+\delta s_i,s_z+\delta s_z) = E_{z+,a}(t,s_i,s_z) - E_{z+,a}(t,s_i,s_z)(d_{p,z+,a}(s_i)+m_{tb-,a}+\mu_a) \quad \text{Equation 3c}$$

Recently (primary) infected people who have previously been on PT

$$E_{z_p,a}(t+\delta t,s_i+\delta s_i) = E_{z_p,a}(t,s_i) - E_{z_p,a}(t,s_i)(d_{p,z-,a}(s_i)+m_{tb-,a}+\mu_a) \quad \text{Equation 4}$$

Latent and Reinfected compartments

To ensure that no one in the population could start PT multiple times, the latent and reinfected compartments are subdivided according to whether or not they have been

on PT previously. For simplicity, this detail is omitted from the model diagram (figure 24 in the thesis). However, the disease-related compartments have not been stratified according to previous PT – this simplification is unlikely to affect conclusions since a negligible proportion of the model population is likely to experience PT twice and treatment for tuberculosis disease.

People with Latent infection

$$L_{z-,a}(t+\delta t) = L_{z-,a}(t)(1-i_{z+,a}(t))(1-d_{n,z-,a}-\lambda(t)-m_{tb-,a}-\mu_a) + M_{in,L,a}(t) + R_{z-,a}(t,T_R) + E_{z-,a}(t,T_E) \quad \text{Equation 5a}$$

$$L_{z+,a}(t+\delta t,0) = L_{z-,a}(t) i_{z+,a}(t)(1-d_{n,z+,a}-\lambda(t)-m_{tb-,a}-\mu_a) \quad \text{Equation 5b}$$

$$L_{z+,a}(t+\delta t, s_z+\delta s_z) = L_{z+,a}(t,s_z) - L_{z+,a}(t,s_z)(d_{n,z+,a}+\lambda(t)+m_{tb-,a}+\mu_a) + E_{z+,a}(t,T_E,s_z) + R_{z+,a}(t,T_R,s_z) \quad \text{Equation 5c}$$

People who have completed PT but have not been reinfected in the previous 5 years

$$P_{e+,a}(t+\delta t) = P_{e+,a}(t) - (\lambda(t)+m_{tb-,a}+\mu_a)P_{e+,a}(t) + V_{z+,a,}(t,T_{z_{max}}) + L_{z+,a}(t,T_{z_{max}}) \quad \text{Equation 6a}$$

$$P_{e-,a}(t+\delta t) = P_{e-,a}(t) - P_{e-,a}(t)(d_{n,z-,a}+\lambda(t)+m_{tb-,a}+\mu_a) + R_{z_p}(t,T_R) + E_{z_p}(t,T_E) \quad \text{Equation 6b}$$

Reinfected people who are not on PT

$$R_{z-,a}(t+\delta t,0) = (1-i_{z+}(t))\lambda(t)(L_{z-,a}(t)+V_{z-,a}(t))(1-d_{x,z-,a}(0)) \quad \text{Equation 7a}$$

$$R_{z-,a}(t+\delta t, s_r+\delta s_r) = (1-i_{z+}(t))R_{z-,a}(t,s_r)(1-d_{x,z-,a}(s_r)-m_{tb-,a}-\mu_a) + M_{in,R,a}(t,s_r) \quad \text{Equation 7b}$$

Reinfected people who are on PT

$$R_{z+,a}(t+\delta t, s_r+\delta s_r, 0) = R_{z-,a}(t,s_r) i_{z+,a}(t)(1-d_{x,z+,a}(s_r)-m_{tb-,a}-\mu_a) \quad \text{Equation 8a}$$

$$R_{z+,a}(t+\delta t, 0, s_z) = \lambda(t)(L_{z+,a}(t,s_z)+V_{z+,a}(t,s_z))(1-d_{x,z+,a}(0)) \quad \text{Equation 8b}$$

$$R_{z+,a}(t+\delta t, s_r+\delta s_r, s_z+\delta s_z) = R_{z+,a}(t,s_r,s_z) - R_{z+,a}(t,s_r,s_z)(d_{x,z+,a}(s_r)+m_{tb-,a}+\mu_a) \quad \text{Equation 8c}$$

Reinfected people who have previously been on PT

$$R_{z_p,a}(t+\delta t,0) = \lambda(t)(P_{e+,a}(t)+P_{e-,a}(t))(1-d_{x,z-,a}(0)) \quad \text{Equation 9a}$$

$$R_{z_p,a}(t+\delta t,s_r+\delta s_r) = R_{z_p,a}(t,s_r) - R_{z_p,a}(t,s_r)(d_{x,z-,a}(s_r)+m_{tb-,a}+\mu_a) \quad \text{Equation 9b}$$

Cases who have not yet been detected

To allow calculation of the proportion of tuberculosis cases that have been reinfected recently, cases which have not yet been detected are further stratified according to the mechanism by which they are experiencing disease (i.e. (exogenous) reinfection or (endogenous) reactivation). Once detected ("found"), cases remain in the detected compartments for a maximum period of 6 months (denoted by $T_{f_{\max}}$), unless they start treatment in the meantime, after which they are redistributed into the undetected compartments, according to their relative size. Considering cases experiencing disease through endogenous reactivation, this redistribution is calculated using the equation

$$p_{Dn,s,a} = \frac{D_{n,s,a}(t,T_{o_{\max}})}{D_{p,s,a}(t,T_{o_{\max}})+D_{n,s,a}(t,T_{o_{\max}})+D_{x,s,a}(t,T_{o_{\max}})}$$

The equation considering cases of primary or exogenous disease is analogous.

Cases experiencing disease because of primary infection, who have not yet been detected

$$D_{p,s+,a}(t+\delta t,0) = o_{s+,a}\sum_{s_i=0}^{T_E} d_{p,z-,a}(s_i)(E_{z-,a}(t,s_i)+E_{z_p,a}(t,s_i)) + o_{s+,a}\sum_{s_i=0}^{T_E}\sum_{s_z=0}^{T_{Z_{\max}}} E_{z+,a}(t,s_i,s_z)d_{p,z+,a}(s_i) \quad \text{Equation 10a}$$

$$D_{p,s-,a}(t+\delta t,0) = (1-o_{s+,a})\sum_{s_i=0}^{T_E} d_{p,z-,a}(s_i)(E_{z-,a}(t,s_i)+E_{z_p,a}(t,s_i)) + (1-o_{s+,a})\sum_{s_i=0}^{T_E}\sum_{s_z=0}^{T_{Z_{\max}}} E_{z+,a}(t,s_i,s_z)d_{p,z+,a}(s_i) \quad \text{Equation 10b}$$

$$D_{p,s,a}(t+\delta t, s_o+\delta s_o) = D_{p,s,a}(t,s_o) + (1-p_{in,f,s})M_{in,D_p,s,a}(t,s_o) - (r_f(t)+m_{tb+}+m_{tb-,a}+\mu_a)D_{p,s,a}(t,s_o)$$ Equation 10c

$$s_o < T_{o_{\max}}$$

$$D_{p,s,a}(t+\delta t, s_o+\delta s_o) = D_{p,s,a}(t,s_o) - (r_f(t)+m_{tb+}+m_{tb-,a}+\mu_a)D_{p,s,a}(t,s_o) + (1-p_{in,f,s})M_{in,D_p,s,a}(t,s_o) + p_{D_p,s,a}F_{s,a}(t,T_{f_{\max}})$$ Equation 10d

$$s_o = T_{o_{\max}}$$

Cases experiencing disease because of reactivation, who have not yet been detected

$$D_{n,s-,a}(t+\delta t,0) = (1-o_{s+,a})(L_{z-,a}(t)d_{n,z-,a} + L_{z+,a}(t)d_{n,z+,a})$$ Equation 11a

$$D_{n,s+,a}(t+\delta t,0) = o_{s+,a}(L_{z-,a}(t)d_{n,z-,a} + L_{z+,a}(t)d_{n,z+,a})$$ Equation 11b

$$D_{n,s,a}(t+\delta t, s_o+\delta s_o) = D_{n,s,a}(t,s_o) - D_{n,s,a}(t,s_o)(r_f(t)+m_{tb+}+m_{tb-,a}+\mu_a) + (1-p_{in,f,s})M_{in,D_n,s,a}(s_o)$$ Equation 11c

$$s_o < T_{o_{\max}}$$

$$D_{n,s,a}(t+\delta t, s_o+\delta s_o) = D_{n,s,a}(t,s_o) - (r_f(t)+m_{tb+}+m_{tb-,a}+\mu_a)D_{n,s,a}(t,s_o) + (1-p_{in,f,s})M_{in,D_n,s,a}(t,s_o) + p_{D_n,s,a}F_{s,a}(t,T_{f_{\max}})$$ Equation 11d

$$s_o = T_{o_{\max}}$$

Cases experiencing disease because of reinfection, who have not yet been detected

$$D_{x,s-,a}(t+\delta t,0) = (1-o_{s+,a})\sum_{s_r=0}^{T_R} d_{x,z-,a}(s_r)(R_{z-,a}(t,s_r) + R_{z_p,a}(t,s_r)) + (1-o_{s+,a})\sum_{s_r=0}^{T_R}\sum_{s_z=0}^{T_{Z_{\max}}} R_{z+,a}(t,s_r,s_z)d_{x,z+,a}(s_r)$$ Equation 12a

$$D_{x,s+,a}(t+\delta t,0) = o_{s+,a}\sum_{s_r=0}^{T_R} d_{x,z-,a}(s_r)(R_{z-,a}(t,s_r)+R_{z_p,a}(t,s_r)) + o_{s+,a}\sum_{s_r=0}^{T_R}\sum_{s_z=0}^{T_{Z_{\max}}} R_{z+,a}(t,s_r,s_z)d_{x,z+,a}(s_r)$$ Equation 12b

$$D_{x,s,a}(t+\delta t,s_o+\delta s_o) = D_{x,s,a}(t,s_o) + (1-p_{in,f,s})M_{in,D_x,s,a}(t,s_o) - (r_f(t)+m_{tb+}+m_{tb-,a}+\mu_a)D_{x,s,a}(t,s_o)$$ Equation 12c

$$s_o < T_{o_{\max}}$$

$$D_{x,s,a}(t+\delta t,s_o+\delta s_o) = D_{x,s,a}(t,s_o) - (r_f(t)+m_{tb+}+m_{tb-,a}+\mu_a)D_{x,s,a}(t,s_o) + (1-p_{in,f,s})M_{in,D_x,s,a}(t,s_o) + p_{D_x,s,a}F_{s,a}(t,T_{f_{\max}})$$ Equation 12d

$$s_o = T_{o_{\max}}$$

Detected cases

$$F_{s-,a}(t+\delta t,0) = \sum_{s_o=0}^{T_{0_{\max}}}(r_f(t)(D_{p,s-a}(t,s_o)+D_{n,s-a}(t,s_o)+D_{x,s-,a}(t,s_o)) + p_{in,f,s-}\sum_{s_o=0}^{T_{0_{\max}}}(M_{in,D_p,s-,a}(t,s_o)+M_{in,D_x,s-,a}(t,s_o)+M_{in,D_x,s-,a}(t,s_o))$$ Equation 13a

$$F_{s-,a}(t+\delta t,s_f+\delta s_f) = F_{s-,a}(t,s_f) - (m_{tb+}+m_{tb-,a}+\mu_a+\tau(s_f))F_{s-,a}(t,s_f)$$ Equation 13b

$$F_{s+,a}(t+\delta t,0) = \sum_{s_o=0}^{T_{0_{\max}}} r_f(t)(D_{p,s+,a}(t,s_o)+D_{n,s+,a}(t,s_o)+D_{x,s+,a}(t,s_o)) + p_{in,f,s+}\sum_{s_o=0}^{T_{0_{\max}}}(M_{in,D_p,s+,a}(t,s_o)+M_{in,D_n,s+,a}(t,s_o)+M_{in,D_x,s+,a}(t,s_o))$$ Equation 13c

$$F_{s+,a}(t+\delta t,s_f+\delta s_f) = F_{s+,a}(t,s_f) - F_{s+,a}(t,s_f)(m_{tb+}+m_{tb-,a}+\mu_a+\tau(s_f))$$ Equation 13d

$$0 < s_f < T_{f_{\max}}$$

Cases undergoing TB treatment

$$C_a(t+\delta t,0) = \sum_{s_f=0}^{T_{f\max}} \tau(s_f)(F_{s-,a}(t,s_f) + F_{s+,a}(t,s_f))$$ Equation 14a

$$C_a(t+\delta t, s_\tau + \delta s_\tau) = C_a(t,s_\tau) - C_a(t,s_\tau)(m_{tb+} + m_{tb-,a} + \mu_a)$$ Equation 14b

$$0 < s_\tau < T_{\tau_{\max}}$$

People who have recovered from TB disease

$$V_{z-,a}(t+\delta t) = V_{z-,a}(t) - V_{z-,a}(t)(i_{z+,a}(t) + \lambda(t) + m_{tb-,a} + \mu_a) + M_{in,V,a}(t) + C(T_{\tau_{\max}})$$ Equation 15a

$$V_{z+,a}(t+\delta t,0) = V_{z-,a}(t) i_{z+,a}(t)$$ Equation 15b

$$V_{z+,a}(t+\delta t, s_z + \delta s_z) = V_{z+,a}(t,s_z) - V_{z+,a}(t,s_z)(i_{z-,a}(t) + \lambda(t) + m_{tb-,a} + \mu_a)$$ Equation 15c

$$0 < s_z < T_{z_{\max}}$$

<u>Transitions at the end of each time step</u>

$$P_{e+,a}(t+\delta t) = P_{e+,a}(t) + L_{z+,a}(t,T_{z_{\max}}) i_{z-}(T_{z_{\max}}) + \sum_{s_z=0}^{T_{z\max}} V_{z+,a}(t,s_z) i_{z-}(s_z) + \sum_{s_r \geq T_{z_{\max}}} R_{z+,a}(t,s_r,T_{z_{\max}}) i_{z-}(T_{z_{\max}}) + \sum_{s_i \geq T_{z_{\max}}} E_{z+,a}(t,s_i,T_{z_{\max}}) i_{z-}(T_{z_{\max}})$$ Equation 16

$$P_{e-,a}(t+\delta t) = P_{e-,a}(t) + \sum_{s_z < T_{z_{\max}}} L_{z+,a}(t,s_z) i_{z-}(s_z)$$ Equation 17

$$L_{z+,a}(t+\delta t, s_z + \delta s_z) = L_{z+,a}(t,s_z)(1 - i_{z-}(s_z))$$ Equation 18

$$s_z < T_{z_{\max}}$$

$$E_{z_p,a}(t+\delta t, s_i+\delta s_i) = E_{z_p,a}(t,s_i) + E_{z+,a}(t,s_i,T_{z_{\max}}) + \sum_{s_z<T_{z_{\max}}} E_{z+,a}(t,s_i,s_z) i_{z-}(s_z)$$ Equation 19a

$$s_i \neq T_{z_{\max}}$$

$$E_{z_p,a}(t+\delta t, s_i+\delta s_i) = E_{z_p,a}(t,s_i) + \sum_{s_z<T_{z_{\max}}} E_{z+,a}(t,s_i,s_z) i_{z-}(s_z)$$ Equation 19b

$$s_i = T_{z_{\max}}$$

$$E_{z+,a}(t,s_i+\delta s_i, s_z+\delta s_z) = E_{z+,a}(t,s_i,s_z)(1-i_{z-}(s_z))$$ Equation 20

$$s_z < T_{z_{\max}}$$

$$R_{z_p,a}(t+\delta t, s_r+\delta s_r) = R_{z_p,a}(t,s_r) + R_{z+,a}(t,s_r,T_{z_{\max}}) + \sum_{s_z<T_{z_{\max}}} R_{z+,a}(t,s_r,s_z) i_{z-}(s_z)$$ Equation 21a

$$s_r \neq T_{z_{\max}}$$

$$R_{z_p,a}(t+\delta t, s_r+\delta s_r) = R_{z_p,a}(t,s_r) + \sum_{s_z<T_{z_{\max}}} R_{z+,a}(t,s_r,s_z) i_{z-}(s_z)$$ Equation 21b

$$s_r = T_{z_{\max}}$$

$$R_{z+,a}(t,s_r+\delta s_r, s_z+\delta s_z) = R_{z+,a}(t,s_r,s_z)(1-i_{z-}(s_z))$$ Equation 22

$$s_z < T_{z_{\max}}$$

$$V_{z+,a}(t+\delta t, s_z+\delta s_z) = V_{z+,a}(t,s_z)(1-i_{z-}(s_z)) \qquad s_z < T_{z_{\max}}$$ Equation 23a

$$V_{z+,a}(t+\delta t, s_z+\delta s_z) = V_{z+,a}(t,s_z)(1-i_{z-}(s_z)) + L_{z+,a}(t,s_z)(1-i_{z-}(s_z)) + \sum_{s_r \geq T_{z\max}} R_{z+,a}(t,s_r,T_{z_{\max}})(1-i_{z-}(T_{z_{\max}}))$$ Equation 23b

$$s_z = T_{z_{\max}}$$

The force or risk of infection

The force of infection at time *t* is given in Equation 24 in terms of the effective contact rate (c_e) (defined as the average number of individuals effectively contacted by each infectious case), the total number of smear-negative and smear-positive cases ($I_{s-}(t)$ and $I_{s+}(t)$ respectively), the population size ($N(t)$) and the relative infectiousness of smear-negative, compared to smear-positive cases (f). The latter equals 22%, consistent with molecular epidemiological data(*2*).

$$\lambda(t) = \frac{c_e(fI_{s-}(t) + I_{s+}(t))}{N(t)}$$ Equation 24

Extending the definition used for acute infections, an effective contact is defined as one that is sufficient to lead to transmission if it occurs between an infectious individual and someone with either a "latent" infection or who has never been infected.[14]

The total number of smear-positive individuals is given by the following equation

$$I_{s+}(t) = \sum_a \sum_{s_o=0}^{T_{o\max}} (D_{p,s+,a}(t,s_o) + D_{n,s+,a}(t,s_o) + D_{x,s+a}(t,s_o)) + \sum_{s_f=0}^{T_{f\max}} F_{s+,a}(t,s_f)$$ Equation 25

The equation for smear-negative cases is analogous.

The rate at which detected cases start TB treatment

The rate at which cases start treatment in the model was calculated so that the average time until cases had started treatment equalled 2 weeks and 82% of cases did not eventually start treatment. Cases who had not started treatment within 6 months were returned to the undetected categories (see above). These rates were calculated as the values for $\tau(s_f)$ satisfying the following equations:

$$u(s_f + 1) = u(s_f) - \tau(s_f)u(s_f) \qquad \text{Equation 26}$$

$$u(4) = 0.18$$

where: $\tau(s_f)$ is the rate at which cases start TB treatment in week s_f after detection (assumed to be constant in each month); $u(s_f)$ is the estimated proportion of those detected who are still untreated s_f weeks after detection.

References

1 Vynnycky E, Fine PE. The natural history of tuberculosis: the implications of age-dependent risks of disease and the role of reinfection. *Epidemiol Infect* 1997; **119**: 183–201.

2 Vynnycky E, Sumner T, Fielding KL, *et al.* Tuberculosis control in South African goldmines: mathematical modelling of a trial of community-wide isoniazid preventive therapy. *submitted*.

3 Health Protection Agency Centre for Infections. Tuberculosis in the UK: 2011 report. http://www.hpa.org.uk/webw/HPAweb&HPAwebStandard/HPAweb_C/1317131784267 (accessed April 11, 2012).

4 Vynnycky E, Fine PE. The annual risk of infection with *Mycobacterium tuberculosis* in England and Wales since 1901. *Int J Tuberc Lung Dis Off J Int Union Tuberc Lung Dis* 1997; **1**: 389–96.

5 Health Protection Agency Centre for Infections. Tuberculosis in the UK: Annual report on tuberculosis surveillance in the UK, 2010. 2010 http://www.hpa.org.uk/web/HPAweb&HPAwebStandard/HPAweb_C/1287143581697 (accessed Nov 24, 2011).

6 Office for National Statistics. Migration Indicators Tool. 2010.

7 Rose A, Watson J, Graham C, *et al.* Tuberculosis at the end of the 20th century in England and Wales: results of a national survey in 1998. *Thorax* 2001; **56**: 173–9.

8 Pareek M, Abubakar I, White PJ, Garnett GP, Lalvani A. Tuberculosis screening of migrants to low-burden nations: insights from evaluation of UK practice. *Eur Respir J Off J Eur Soc Clin Respir Physiol* 2011; **37**: 1175–82.

9 Behr MA, Warren SA, Salamon H, *et al.* Transmission of *Mycobacterium tuberculosis* from patients smear-negative for acid-fast bacilli. *Lancet* 1999; **353**: 444–9.

10 Sutherland I, Svandová E, Radhakrishna S. The development of clinical tuberculosis following infection with tubercle bacilli. 1. A theoretical model for the development of clinical tuberculosis following infection, linking from data on the risk of tuberculous infection and the incidence of clinical tuberculosis in the Netherlands. *Tubercle* 1982; **63**: 255–68.

11 National Collaborating Centre for Chronic Conditions (UK), Centre for Clinical Practice at NICE (UK). Tuberculosis: Clinical Diagnosis and Management of Tuberculosis, and Measures for Its Prevention and Control. London: National Institute for Health and Clinical Excellence (UK), 2011 http://www.ncbi.nlm.nih.gov/books/NBK97852/ (accessed June 4, 2014).

12 Comstock GW, Ferebee SH, Hammes LM. A controlled trial of community-wide isoniazid prophylaxis in Alaska. *Am Rev Respir Dis* 1967; **95**: 935–43.

13 Ferebee SH. Controlled chemoprophylaxis trials in tuberculosis. A general review. *Bibl Tuberc* 1970; **26**: 28–106.

14 ABBEY H. An examination of the Reed-Frost theory of epidemics. *Hum Biol* 1952; **24**: 201–33.

15 Presanis AM, Gill ON, Chadborn TR, *et al*. Insights into the rise in HIV infections, 2001 to 2008: a Bayesian synthesis of prevalence evidence. *AIDS Lond Engl* 2010; **24**: 2849–58.

16 Health Protection Agency. Tuberculosis in the UK: Annual report on tuberculosis surveillance in the UK 2009. 2009 www.hpa.org.uk/Publications/InfectiousDiseases/Tuberculosis/0912Tuberculosisintheuk/.

17 Office for National Statistics. Population Estimates for UK, England and Wales, Scotland and Northern Ireland, Mid-2001 to Mid-2010 Revised. http://www.ons.gov.uk/ons/rel/pop-estimate/population-estimates-for-uk--england-and-wales--scotland-and-northern-ireland/mid-2001-to-mid-2010-revised/index.html (accessed July 24, 2014).

18 Pareek M, Bond M, Shorey J, *et al*. Community-based evaluation of immigrant tuberculosis screening using interferon γ release assays and tuberculin skin testing: observational study and economic analysis. *Thorax* 2013; **68**: 230–9.

19 Marks SM, Taylor Z, Qualls NL, Shrestha-Kuwahara RJ, Wilce MA, Nguyen CH. Outcomes of Contact Investigations of Infectious Tuberculosis Patients. *Am J Respir Crit Care Med* 2000; **162**: 2033–8.

20 Ansari S, Thomas S, Campbell IA, Furness L, Evans MR. Refined tuberculosis contact tracing in a low incidence area. *Respir Med* 1998; **92**: 1127–31.

Appendix 6: Cost-effectiveness model

The cost-effectiveness model was developed by Jo Lord.

—

Objective

To estimate the cost-effectiveness of the TB-STS as an addition to the current system for TB control in England.

Methods

The analysis followed the methods recommended by the National Institute for Health and Clinical Excellence (NICE) for evaluation of public health interventions.[1]

Perspective – A public sector perspective was used for costing, and included costs and savings attributable to the TB-STS for the NHS, Local Authorities, Department of Health and other public bodies. The majority of costs and savings from the TB-STS fall on the Public Health England (PHE) centre, regions, Health Protection Units (HPU), laboratories and NHS TB services.

Measure of health effects – Health benefits attributable to the TB-STS were estimated in the form of Quality Adjusted Life Years (QALYs) gained for index cases, their contacts, and for people benefiting from prevention of onward transmission of TB (as estimated from the transmission model). QALY estimates included TB-related mortality and morbidity.

Time horizon – Costs and health effects resulting from operating the TB-STS were estimated over a period of 20 years.

Incremental analysis – The results are presented in the form of an Incremental Cost Effectiveness Ratio (ICER), which is the additional cost per additional QALY gained with the TB-STS. Thus we estimated the expected difference in costs and in health effects with/without the TB-STS. Any costs or health effects incurred under both systems were ignored. The resulting ICER was compared with the NICE-recommended upper threshold of £30,000 per QALY gained.[2]

Uncertainty – Deterministic sensitivity analysis was used to test the impact of uncertainty over input parameters on the cost-effectiveness results.

Discounting – Costs and QALYs were both discounted at the NICE recommended rate of 3.5% per year. The impact of using the Department of Health recommended discount rates of 3.5% for costs and 1.5% discount rates for QALYs were tested in sensitivity analysis.

The conceptual model underlying the economic analysis is illustrated in figure 25 in the thesis. It was hypothesised that the introduction of the TB-STS might influence outcomes or health care expenditure through the following mechanisms:

TB-STS infrastructure – The TB-STS has imposed capital and revenue costs for the reference laboratories and national, regional and local Health Protection Services (HPS). These include direct costs of the tests, but also costs of establishing the infrastructure to request tests, report results and perform quality assurance.

Detection of false positives – One potential benefit of strain typing is earlier identification of the false positive TB cases that can be caused by laboratory contamination. In addition to the avoidance of anxiety for patients and their families, earlier identification of such cases has health and financial implications if treatment is avoided or reduced. There might also be benefits in earlier detection of alternative diagnoses (e.g. lung cancer), but these are difficult to quantify, and have not been included in our analysis.

Case finding activity – Introduction of the TB-STS might in theory have increased or decreased case finding activity and related costs. As contact tracing is usually completed before the strain type result is available (Figure 39, page 145 in the main thesis), one would not expect it to impact on the initial number of contacts traced by TB clinics. However, it is possible that it could have affected decisions by health protection staff to initiate or extend investigations of potential outbreaks. If strain typing identifies otherwise unsuspected clusters of cases, the number of contacts followed up could increase, increasing costs. But strain typing might also have the

effect of disproving links between epidemiologically linked cases, thus reducing case finding activity and costs.

Case finding yield – Regardless of the impact on the volume of case finding activity, we hypothesised that strain typing would improve the yield of case finding; increasing the number of cases of active disease and latent infection identified per case of TB. If true, this would have a number of benefits:

Earlier detection of active disease – It seems plausible that cases detected through the TB-STS-enhanced cluster investigations might benefit from earlier diagnosis and initiation of treatment, and that earlier treatment might be associated with a reduction in QALY loss from TB.

Increased detection of latent infection – One might also expect an increase in the detection of latent infection resulting from strain typing. Individuals diagnosed with Latent TB Infection (LTBI) who are suitable for and accept prophylactic treatment should then have a reduced risk of developing active disease themselves, avoiding QALY loss and NHS costs. However, there are costs and side effects of prophylaxis, which will offset its benefits to some extent.

Prevention of onward transmission – Both earlier detection of active disease and increased prophylactic treatment should help to prevent transmission. If so, this would lead to further QALY gains and cost savings.

In addition to the above direct effects, the TB-STS may well provide more indirect benefits. For example, the availability of a national information resource on the distribution and growth of clusters might benefit future tuberculosis research and service development (see Box 5 in the thesis). Such effects are hard to quantify or value, and so were not been included in the economic analysis, but they were discussed and taken into consider by the evaluation expert group.

<u>Estimated impact on false positive identification</u>

The survey of reference laboratories identified 70 possible false positive TB tests, of which 59 (84%) had a known outcome. 30 of the incidents with a known outcome

(51%) were confirmed as false positive results attributed to cross contamination. Of these, 17 (57%) were not known about by the source laboratory before they were contacted by the reference lab. For eight of the 30 confirmed cross-contamination incidents (27%), the patient commenced treatment. For the economic analysis, it was assumed that five cases of unnecessary treatment would be avoided per year due to the TB-STS (ten cases per year was tested in sensitivity analysis).

Estimated impact on case finding activity and yield

Evidence on the impact of the TB-STS on the volume and yield of case finding activity was sparse. There was some evidence of an increase in time spent on cluster investigations reported in the survey of health protection staff: from a mean of 2.7% before to 7.1% after implementation.[3] However, TB specialist nurses did not report any significant increase in time spent on contact tracing. In the economic analysis, an opportunity cost for additional time spent by HPU staff on cluster investigations was assumed: 4.4% Whole Time Equivalent (WTE) for each of 26 Consultants in Communicable Disease Control (CCDC) at £99,000pa, costing a total of £113,256 per year (total annual cost of £50,000 per year and £500,000 per year tested in sensitivity analysis).

There was no clear evidence of whether introduction of the TB-STS resulted in an increase in the number of contacts screened, or in the yield of contacts with active disease or latent infection found. Analysis of the contact tracing database, national dataset and cluster monitoring dataset showed that a greater number of contacts were screened and more contacts with latent infection were identified in cases that were clustered and investigated compared with unique cases. However, there were no significant differences in the numbers of contacts screened or cases of latent infection identified for clustered cases that were investigated compared with clustered cases that were not investigated. Similarly, evidence for a change in the rate of cluster growth after the initiation of an investigation or for a change in the duration of diagnostic delay was equivocal.

It is unclear whether these negative results reflect the absence of an effect of the TB-STS, or the difficulties in obtaining evidence. We therefore conducted a scenario

analysis, in which we estimated the cost-effectiveness of the TB-STS under a series of assumptions about its possible effects.

Population assumptions

Results were estimated across the population of England (53m) and took account of the age distribution of the population (age groups <15, 15-34, 35-54, 55+).[4] The results were based on an epidemiological scenario with a medium TB incidence (similar to that in the non-white UK-born population in which the average infection risk was constant over time at 0.1%.) This was chosen to reflect an average risk level across the community.

Scenarios investigated

The transmission model was used to estimate the number of cases prevented under a range of assumptions about the effect of the TB-STS on: a) the proportion of previously infected individuals detected; and b) the mean length of time between onset of symptoms and treatment initiation.

The base case scenario (S0) was intended to reflect the expected costs and outcomes of the TB control system in the absence of the TB-STS. This was modelled assuming that three percent of previously infected individuals are detected per year and that the mean time from onset of symptoms to the start of treatment is 12 weeks. The transmission model results for this base case scenario are summarised in Supplementary Table 1 for the population of England over 20 years, and assuming a constant risk of infection of 0.1% per annum. The estimated number of cases diagnosed exceeds the number of new cases in each year, as there is a pool of cases who have previously not been diagnosed or who have defaulted from treatment.

The cost-effectiveness of the TB-STS was then estimated under a range of assumptions about its effect on identification of cases of LTBI found and the diagnostic delay for active cases. The results of the transmission model under these scenarios are summarised in table 37 in the thesis. They suggest that increases in the proportion of people with latent infection identified and treated have a relatively modest impact on TB incidence: if an additional 10% of prevalent infections were

detected each year, the number of new TB cases would fall by an estimated 736 cases per year (11%). In contrast, reductions in diagnostic delay for active cases were estimated to have a much bigger impact on TB incidence: a one week reduction in the time from onset of symptoms to treatment was associated with an estimated reduction of 1,650 TB cases (25%). Furthermore, if such a reduction in diagnostic delay could be achieved, it would also be accompanied by a reduction in the number of people requiring prophylactic treatment.

Supplementary Table 1 – Summary of transmission model results for the base case scenario (S0)

Scenario	Year	LTBI diagnosed	LTBI starting treatment	New TB cases	TB cases diagnosed	TB cases starting treatment
S0	Year 1	9,069	8,616	6,730	7,568	6,698
	Year 2	9,060	8,607	6,723	7,561	6,691
	Year 3	9,051	8,599	6,717	7,554	6,685
	Year 4	9,043	8,590	6,711	7,547	6,679
	Year 5	9,034	8,582	6,705	7,540	6,673
	Year 6	9,025	8,574	6,698	7,532	6,666
	Year 7	9,016	8,566	6,692	7,526	6,660
	Year 8	9,008	8,557	6,686	7,519	6,654
	Year 9	8,999	8,549	6,680	7,512	6,648
	Year 10	8,991	8,541	6,674	7,505	6,642
	Year 11	8,982	8,533	6,668	7,498	6,636
	Year 12	8,974	8,525	6,662	7,491	6,630
	Year 13	8,966	8,517	6,656	7,485	6,624
	Year 14	8,957	8,509	6,650	7,478	6,618
	Year 15	8,949	8,502	6,644	7,472	6,612
	Year 16	8,941	8,494	6,638	7,465	6,607
	Year 17	8,933	8,486	6,633	7,458	6,601
	Year 18	8,925	8,478	6,627	7,452	6,595
	Year 19	8,916	8,471	6,621	7,446	6,589
	Year 20	8,908	8,463	6,615	7,439	6,584
	Total	179,747	170,760	133,431	150,046	132,794
	Mean pa	8,987	8,538	6,672	7,502	6,640

Estimated number of cases by year for population of England (53m) over 20 years, assuming constant ARI of 0.1%.

Cost estimates

The estimated costs of establishing and running the national strain typing programme were estimated from financial information obtained from Public Health England and TB Reference Laboratories. Capital expenditure was converted to an equivalent annual cost assuming a 10 year lifetime of the investment and 3.5% annual discount rate. Total costs were estimated at just under £1m per year (Table 9 in the thesis, page 96).

The estimated costs of screening, diagnosis and treatment are shown in Supplementary Table 2. The average quantities of resource items per patient were based on standard treatment protocols, informed by expert judgement. Unit costs per resource item were taken from standard national sources: Department of Health Reference Costs 2010-11 for Tuberculosis Specialist Nurse visits, outpatient consultations (respiratory clinic), and inpatient admissions; British National Formulary, Sept 2012 for medications; and published sources for tests.[5,6]

The costs of treatment for latent and active disease were estimated at £743 and £1,114 respectively for a full course, or £669 and £1,002 respectively allowing for drop out from treatment: assuming that 15% of patients drop out, after a mean of one month for latent infection and 2 months for active disease. Patients with TB who drop out are likely to be identified and offered treatment again at a later time. Such repeat cases are included in the transmission model estimates of the number of people diagnosed per year, and incur additional costs for diagnosis and treatment in the cost effectiveness analysis. For simplicity, it is assumed that the cost of diagnosis and treatment is the same for new and repeat cases.

The cost per contact screened was estimated at £234 (including contact tracing, TST and IGRA testing, and initial rule-out of active disease). The total cost of contact screening was estimated as a function of the number of people diagnosed with LTBI, as estimated by the transmission model. The study of the yield of cluster investigations above (Table 28 in the thesis, page 150) found that on average (across unique and clustered cases) 2.6 contacts were screened and 0.7 cases of LTBI were identified per TB case. Therefore, it was assumed that to diagnose one case of LTBI, 3.97 contacts would need to be screened, on average, at a cost of £963. The cost of further follow-up and investigations for each contact suspected of having active disease was estimated at £434. It was assumed that 20% of individuals investigated for active TB would receive a positive diagnosis (based on estimates of contact tracing yield reported here and expert opinion) so the estimated cost to diagnose one case of TB was £2,170 (5 x £434). The costs of treatment for latent and active disease were estimated at £743 and £1,114 respectively for a full course, or £669 and

£1,002 respectively allowing for drop out from treatment: assuming that 15% of patients drop out,[7] after a mean of one month for LTBI and 2 months for active disease.

Supplementary Table 2 – Estimated costs of screening, diagnosis and treatment

	Quantity	Unit cost	Cost	Source
Contact screening and follow up				
TB specialist nurse - non face to face	1	£27	£27	Ref cost 2011[8]
TB specialist nurse - face to face	2	£62	£124	Ref cost 2011
Mantoux test	1	£1.22	£1.22	NICE 2011[9]
IGRA test	0.5	£56	£28	Pareek 2011[10]
Outpatient appointment for IGRA +	0.25	£187	£47	Ref cost 2011
Chest X-ray (to rule out active disease)	0.25	£28	£7	NICE 2010[11]
Per contact screened			£234	
Per person diagnosed with LTBI [a]			£963	
Diagnosis of active disease				
TB specialist nurse - face to face	3	£62	£186	Ref cost 2011
Outpatient appointment for diagnosis	1	£187	£187	Ref cost 2011
Chest X-ray	1	£28	£28	NICE 2010
Sputum smear microscopy	1	£1.56	£1.56	Dowdy 2008[12]
Culture & MDR identification	1	£30	£30	Dinnes 2007[13]
Liver function test	1	£1	£1	Ref cost 2011
Per contact with suspected TB			£434	
Per person diagnosed with TB [b]			£2,170	
Management of LTBI				
Follow-up appointments nurse only	3	£62	£186	Ref cost 2011
Follow-up appointments nurse & consultant	2	£185	£370	Ref cost 2011
Isoniazid 300 mg daily (per month)	3	£41	£124	BNF 2012[14]
Rifampicin 600 mg daily (per month)	3	£21	£63	BNF 2012
B6 pyridoxine 10mg tablets (per month)	3	£0.5	£1	BNF 2012
Per person completing treatment			£743	
Per person starting treatment [c]			£669	
Management of active disease				
Admission	5%	£2,949	£147	Ref cost 2011
Follow-up appointments nurse only	5	£62	£310	Ref cost 2011
Follow-up appointments nurse & consultant	2	£185	£370	Ref cost 2011
Rifater (R,I ,P) 6 tablets daily for 2 months	2	£37	£74	BNF 2012
Ethambutol 15 mg/kg for 2 months	2	£63	£126	BNF 2012
Rifanah (R,I) 300/150 2 tab daily for 4 months	4	£21	£84	BNF 2012
B6 pyridoxine (per month)	6	£0.5	£3	BNF 2012
Per person completing treatment			£1,114	
Per person starting treatment [d]			£1,002	

[a] Assumes 4.11 contacts screened per person diagnosed with LTBI (Table 28, page 150 in the thesis); [b] Assumes 5 people investigated per person diagnosed with TB (expert opinion); [c] Assumes that 15% do not complete chemoprophylaxis, after an average 1 month of treatment (expert opnion); [d] Assumes that 15% do not complete treatment,[7] after an average 2 months of treatment (expert opinion); BNF British National Formulary

QALY estimates

Estimates of the QALY loss per case of TB are shown in Supplementary Table 3. At ages of 15 and older, TB-related mortality contributed more to estimated QALY loss than TB-related morbidity. On average across all ages, a loss of 0.5 QALYs was attributable to case fatality out of a total estimated 0.62 QALYs lost. Estimates of QALY loss due to morbidity are based on simple assumptions about the duration and quality of life reduction in three periods of time:

a. **Pre-treatment period**: from onset of symptoms to initiation of treatment, which is assumed to last for 12 weeks in scenarios S0 to S10, and is reduced according to the diagnostic delay in scenarios S11 to S14. During this time, patients are assumed to have a utility equal to 90% of that of the general population of the same age.

b. **Acute period**: assumed to last for 2 months from diagnosis, during which patients have a utility value of 0.675.[15]

c. **Post-acute period:** from after the acute period to the end of treatment (4 months), during which patients have a utility value of 0.813.[15]

Overall QALY losses per case of TB are estimated to be 0.19, 0.40, 0.59 and 1.18, respectively, for patients in age group 0-14, 15-34, 35-54, and 55 plus. It is assumed that after treatment completion there is no lasting effect of TB on quality of life or mortality risk, although within the transmission model, individuals can be re-infected, potentially incurring another QALY loss associated with a new TB incidence.

The QALY impacts of adverse effects of treatment are assumed to be incorporated in the above utility multipliers for active disease. Patients with a false positive TB diagnosis who start treatment, are assumed to be treated for four months on average, and during this time they are assumed to experience a utility loss of 0.1 due to the inconvenience and adverse side effects of TB treatment. Thus, the avoidance of

treatment for a false positive is associated with a mean QALY gain of 0.03. The QALY loss associated with the adverse effects of prophylactic treatment is estimated based on the assumption that 10% of patients experience some side effects, and that these last for one month on average, incurring a mean utility loss of 0.1. Thus the mean QALY loss per person treated with prophylaxis is 0.0008. It is assumed that there are no lasting consequences from adverse reactions to treatment for active disease or LTBI.

Supplementary Table 3 – Calculation of QALY loss per case of TB

	Incidence		Case fatality	Life expectancy		Utility				Discounted to age of incidence											
										TB mortality						TB morbidity			Total QALYs lost		
										LYL per fatality		QALYs lost per fatality *		QALYs lost per case		QALYs lost per case					
Age	Male	Female		Male	Female	No TB	Pre- treat	Acute	Post acute	Male	Female	Male	Female	Male	Female	Pre-treat	Acute	Post acute	Mortality	Morbidity	Total
0-4	62	78	0.3%	78.0	82.1	0.94	0.84	0.63	0.76	26.62	26.88	24.98	25.22	0.07	0.08	0.023	0.051	0.058	11	19	29
5-9	47	46	0.2%	73.5	77.5	0.94	0.84	0.63	0.76	26.29	26.59	24.67	24.94	0.05	0.05	0.023	0.051	0.058	5	12	17
10-14	77	103	0.2%	68.5	72.5	0.94	0.84	0.63	0.76	25.86	26.21	24.27	24.60	0.05	0.05	0.023	0.051	0.058	9	24	33
15-19	192	173	1.2%	63.5	67.6	0.94	0.84	0.63	0.76	25.36	25.78	23.80	24.19	0.29	0.29	0.023	0.051	0.058	105	48	153
20-24	582	406	1.2%	58.7	62.6	0.94	0.84	0.63	0.76	24.77	25.26	23.25	23.70	0.28	0.28	0.023	0.051	0.058	278	130	408
25-29	717	547	1.2%	53.9	57.7	0.94	0.84	0.63	0.76	24.09	24.65	22.60	23.13	0.27	0.28	0.023	0.051	0.058	346	167	513
30-34	686	500	1.2%	49.0	52.8	0.91	0.82	0.62	0.74	23.28	23.93	21.29	21.88	0.26	0.26	0.023	0.050	0.057	307	153	459
35-39	531	365	1.2%	44.3	47.9	0.91	0.82	0.61	0.74	22.34	23.08	20.26	20.93	0.24	0.25	0.023	0.049	0.057	221	114	335
40-44	416	294	1.2%	39.6	43.1	0.88	0.79	0.60	0.72	21.24	22.08	18.74	19.48	0.22	0.23	0.022	0.048	0.055	162	88	250
45-49	343	227	4.8%	34.9	38.3	0.86	0.78	0.58	0.70	19.97	20.92	17.25	18.07	0.83	0.87	0.022	0.047	0.054	481	68	549
50-54	290	223	4.8%	30.3	33.6	0.83	0.75	0.56	0.68	18.50	19.58	15.44	16.34	0.74	0.78	0.021	0.045	0.052	390	59	449
55-59	245	180	4.8%	25.9	29.1	0.82	0.74	0.55	0.67	16.85	18.05	13.86	14.84	0.67	0.71	0.021	0.045	0.051	291	48	340
60-64	200	163	4.8%	21.7	24.6	0.81	0.73	0.54	0.66	15.03	16.33	12.13	13.18	0.58	0.63	0.020	0.044	0.050	220	41	260
65-69	164	126	17.6%	17.7	20.4	0.80	0.72	0.54	0.65	13.05	14.40	10.49	11.58	1.85	2.04	0.020	0.044	0.050	560	30	590
70-74	168	133	17.6%	14.1	16.4	0.78	0.70	0.53	0.63	10.98	12.29	8.55	9.57	1.50	1.69	0.019	0.042	0.049	477	31	508
75-79	167	109	17.6%	10.8	12.6	0.75	0.68	0.51	0.61	8.86	10.08	6.67	7.59	1.17	1.34	0.019	0.041	0.047	342	27	369
80-84	127	82	17.6%	8.0	9.4	0.70	0.63	0.47	0.57	6.86	7.87	4.79	5.49	0.84	0.97	0.017	0.038	0.044	186	19	205
85-89	77	35	17.6%	5.8	6.7	0.65	0.58	0.44	0.53	5.16	5.87	3.35	3.81	0.59	0.67	0.016	0.035	0.040	69	9	78
90+	21	21	17.6%	4.2	4.6	0.65	0.58	0.44	0.53	3.80	4.17	2.47	2.71	0.43	0.48	0.016	0.035	0.040	19	4	23
	5,112	3,811																	4,477	1,092	5,569
						Utility multipliers	0.9	0.675	0.813												
						Duration of health states (months)	3	2	4												

0-14	0.06	0.13	0.19
15-34	0.27	0.13	0.40
35-54	0.47	0.12	0.59
55+	1.07	0.10	1.18
All ages	0.50	0.12	0.62

Approximation - assumes current quality of life persists for life
Utility values from Kruijshaar et al 2010,[15] reported in NICE guidance 2011[16]
LYL years of life lost

References

1 National Institute for Health and Clinical Excellence. Methods for development of NICE public health guidance. 2009. http://www.nice.org.uk/phmethods2009 (accessed Dec 14, 2011).

2 National Institute for Health and Clinical Excellence. Social value judgements. Principles for the development of NICE guidance, Second. London: NICE, 2008 http://www.nice.org.uk/ (accessed Feb 27, 2013).

3 Mears J, Abubakar I, Crisp D, *et al.* Prospective evaluation of a complex public health intervention: lessons from an initial and follow-up cross-sectional survey of the tuberculosis strain typing service in England. *BMC Public Health* 2014; **14**: 1023.

4 2011 Census - Population and Household Estimates for England and Wales, March 2011. Off. Natl. Stat. 2012; published online July 16. http://www.ons.gov.uk/ons/rel/census/2011-census/population-and-household-estimates-for-england-and-wales/stb-e-w.html (accessed March 25, 2013).

5 Hughes R, Wonderling D, Li B, Higgins B. The cost effectiveness of Nucleic Acid Amplification Techniques for the diagnosis of tuberculosis. *Respir Med* 2012; **106**: 300–7.

6 Pareek M, Bond M, Shorey J, *et al.* Community-based evaluation of immigrant tuberculosis screening using interferon γ release assays and tuberculin skin testing: observational study and economic analysis. *Thorax* 2013; **68**: 230–9.

7 Public Health England. Tuberculosis in the UK. 2014 report. London, 2014.

8 Health D of. NHS reference costs 2009-2010. 2011; published online Jan 13. http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_123459 (accessed April 13, 2012).

9 NICE. Tuberculosis: costing statement. Update for 2011 Guidelines. 2011; published online March. http://guidance.nice.org.uk/CG117/CostingStatement/pdf/English (accessed Nov 30, 2011).

10 Pareek M, Abubakar I, White PJ, Garnett GP, Lalvani A. Tuberculosis screening of migrants to low-burden nations: insights from evaluation of UK practice. *Eur Respir J Off J Eur Soc Clin Respir Physiol* 2011; **37**: 1175–82.

11 National Institute for Health and Clinical Excellence. Tuberculosis: NICE guideline for hard-to-reach groups. http://publications.nice.org.uk/identifying-

and-managing-tuberculosis-among-hard-to-reach-groups-ph37 (accessed Sept 20, 2012).

12 Dowdy DW, O'Brien MA, Bishai D. Cost-effectiveness of novel diagnostic tools for the diagnosis of tuberculosis. *Int J Tuberc Lung Dis Off J Int Union Tuberc Lung Dis* 2008; **12**: 1021–9.

13 Dinnes J, Deeks J, Kunst H, *et al.* A systematic review of rapid diagnostic tests for the detection of tuberculosis infection. *Health Technol Assess Winch Engl* 2007; **11**: 1–196.

14 Committee JF. British National Formulary (BNF) 64, 64 edition. London: Pharmaceutical Press, 2012.

15 Kruijshaar ME, Lipman M, Essink-Bot M-L, *et al.* Health status of UK patients with active tuberculosis. *Int J Tuberc Lung Dis Off J Int Union Tuberc Lung Dis* 2010; **14**: 296–302.

16 NICE. Tuberculosis: NICE guideline. Clinical diagnosis and management of tuberculosis, and measures for its prevention and control. Natl. Inst. Health Clin. Excell. 2011. http://www.nice.org.uk/ (accessed Sept 20, 2012).

Appendix 7: Publications and conference presentations

Publications

Mears J, Abubakar I, Cohen T, McHugh TD, Sonnenberg P. Effect of study design and setting on tuberculosis clustering estimates using Mycobacterial Interspersed Repetitive Units-Variable Number Tandem Repeats (MIRU-VNTR): a systematic review. *BMJ Open* 2015; **5**: e005636.

Mears J, Abubakar I, Crisp D, *et al.* Prospective evaluation of a complex public health intervention: lessons from an initial and follow-up cross-sectional survey of the tuberculosis strain typing service in England. *BMC Public Health* 2014; **14**: 1023.

Mears J, Vynnycky E, Lord J, *et al.* Evaluation of the Tuberculosis Strain Typing Service (TB-STS) in England. *Thorax*; published online April 16 2015.

Conference presentations

Mears J, Vynnycky E, Lord J, et al. Evaluation of the Tuberculosis Strain Typing Service (TB-STS) in England. The Lancet 2013; 382, Supplement 3: S73.

Mears J, on behalf of the TB Strain Typing Evaluation Group, Evaluation of the National TB Strain Typing Service: Results from a Baseline Survey. Health Protection 2011, Warwick

BMJ Open

Effect of study design and setting on tuberculosis clustering estimates using Mycobacterial Interspersed Repetitive Units-Variable Number Tandem Repeats (MIRU-VNTR): a systematic review

Jessica Mears,[1] Ibrahim Abubakar,[1,2,3] Theodore Cohen,[4] Timothy D McHugh,[5] Pam Sonnenberg[1]

To cite: Mears J, Abubakar I, Cohen T, *et al*. Effect of study design and setting on tuberculosis clustering estimates using Mycobacterial Interspersed Repetitive Units-Variable Number Tandem Repeats (MIRU-VNTR): a systematic review. *BMJ Open* 2015;**5**:e005636. doi:10.1136/bmjopen-2014-005636

▸ Prepublication history and additional material is available. To view please visit the journal (http://dx.doi.org/10.1136/bmjopen-2014-005636).

Received 12 May 2014
Revised 23 December 2014
Accepted 30 December 2014

For numbered affiliations see end of article.

Correspondence to
Dr Pam Sonnenberg;
p.sonnenberg@ucl.ac.uk

ABSTRACT

Objectives: To systematically review the evidence for the impact of study design and setting on the interpretation of tuberculosis (TB) transmission using clustering derived from Mycobacterial Interspersed Repetitive Units-Variable Number Tandem Repeats (MIRU-VNTR) strain typing.

Data sources: MEDLINE, EMBASE, CINHAL, Web of Science and Scopus were searched for articles published before 21st October 2014.

Review methods: Studies in humans that reported the proportion of clustering of TB isolates by MIRU-VNTR were included in the analysis. Univariable meta-regression analyses were conducted to assess the influence of study design and setting on the proportion of clustering.

Results: The search identified 27 eligible articles reporting clustering between 0% and 63%. The number of MIRU-VNTR loci typed, requiring consent to type patient isolates (as a proxy for sampling fraction), the TB incidence and the maximum cluster size explained 14%, 14%, 27% and 48% of between-study variation, respectively, and had a significant association with the proportion of clustering.

Conclusions: Although MIRU-VNTR typing is being adopted worldwide there is a paucity of data on how study design and setting may influence estimates of clustering. We have highlighted study design variables for consideration in the design and interpretation of future studies.

Strengths and limitations of this study

- This is a timely evaluation of the impact of study design on estimates of tuberculosis clustering using Mycobacterial Interspersed Repetitive Units-Variable Number Tandem Repeats strain typing because it has been incorporated into national typing services globally.
- The strength of this meta-analysis was limited by the lack of detail reported by the included studies, highlighting the need for better quality reporting in primary studies.
- We have shown that the proportion of clustering derived from MIRU-VTNR typing is influenced by the number of loci typed, whether consent is required to type isolates, TB incidence in the study setting, and the maximum cluster size, highlighting these as important considerations in the design and interpretation of future studies.

INTRODUCTION

The introduction of molecular typing methods has improved our understanding of *Mycobacterium tuberculosis* (TB) transmission and has changed local and national control policies.[1–5] The proportion of cases that are clustered is often used to estimate the amount of ongoing transmission within the population, based on the assumption that cases with indistinguishable strain types are part of a chain of transmission. TB molecular typing methodology is changing rapidly and it is important that we better understand how to interpret the outputs and thus act.

TB molecular typing methods include Spoligotyping,[6] insertion sequence *6110* (IS*6110*) restriction fragment length polymorphism (RFLP) analysis (the recent gold standard),[7] Mycobacterial Interspersed Repetitive Units-Variable Number Tandem Repeats (MIRU-VNTR) typing,[8] and whole genome sequencing.[9–11] Published reviews have identified factors that might influence or bias clustering by IS*6110* RFLP.[12 13] No study has repeated this analysis using more up-to-date typing methods, which is important for understanding of the epidemiology of TB and to shape the application of molecular typing to improve TB control.

Published meta-analyses and modelling studies using IS*6110* RFLP data show that the proportion of clustering observed can be affected by (1) study design (affecting the proportion of eligible cases that are included in the study); (2) features of the typing method (such as the ability to type isolates with low copy numbers); and (3) study setting (such as characteristics of the study population). For example, the proportion of clustering increases when the fraction of the total data sampled increases[13–15] and when study duration increases.[16]

MIRU-VNTR is currently the preferred method of molecular typing,[17–21] and can be used together with Spoligotyping.[8] Relative to IS*6110* RFLP, MIRU-VNTR does not have to exclude isolates with a low IS*6110* copy number, has a faster turnaround time, is high throughput and the numeric strain types are more easily compared. MIRU-VNTR strain typing is increasingly being adopted worldwide,[1 22–27] yet unlike IS*6110* RFLP, the evidence for the interpretation of the findings such as the impact of study design and setting on clustering have not been reviewed. Although the two typing methods have been shown to have a similar discriminatory value, the markers evolve independently and at different rates, resulting in a difference in clustering between the two methods.[28] This suggests that there could be differences in the way study design, typing method and setting affects clustering by the two methods. We conducted a systematic review to assess the evidence for the impact of study design and setting on the interpretation of TB transmission using clustering derived from MIRU-VNTR strain typing—as has been shown using IS*6110* RFLP typing.

METHODS

Five electronic databases were searched (EMBASE, ISI Web of Science, CINHAL, Scopus and Medline (Ovid)) up to 20 October 2014. The search strategy combined the following terms with Boolean operators: Tuberculosis, strain typing, and transmission (see online supplementary appendix 1). The search was limited to studies using the standard MIRU-VNTR method,[8] in humans only, and in English.

All titles and abstracts from each of the searches were examined. The full text of each paper was obtained and reviewed if the study reported MIRU-VNTR strain typing of *M. tuberculosis* complex isolates with at least 15 of the standardised 24 loci (Exact Tandem Repeat A, B, C, D, E; MIRU 2, 10, 16, 20, 23, 24, 26, 27, 39, 40; VNTR 424, 1955, 2163b, 2347, 2401, 3171, 3690, 4052, 4156).[8 29 30]

Studies using fewer than 15 loci were not included because the level of discrimination is inadequate for epidemiological use (n=121).[8] Studies that used loci different to the standardised 15 and 24 set were not included in the analysis in order to reduce the heterogeneity between studies (n=19). All publication types were included in this first screen to ensure that no relevant data were missed.

Reviews, letters, editorials, outbreaks or case reports (n=103) were excluded in the second screen. Studies that used incomplete sampling (eg, random samples, studies using subsets of populations such as multidrug-resistant patients; n=47) and studies that had a sample size of less than 50 (n=4) were also excluded.

A reviewer (JM) extracted the following data items from all included studies using a form developed in Excel (Microsoft 2010): publication details (year, authors, study country), study details (study duration, loci typed, secondary typing method, study population, whether participant consent was required (a characteristic of the study design that was used as proxy for sampling fraction, assuming that where consent was required the sampling fraction was low)), the number of clustered and unique isolates and the covariates of interest: the maximum size of clusters; the proportion of clusters containing two cases; the proportion of the population that was culture positive; the proportion of culture positive isolates typed; risk factors for clustering; and the Hunter Gaston Discriminatory Index (HGDI)[31]). IA extracted data from 10% of the papers for external validity, disagreements were discussed and a consensus agreed on.

The main outcome measure—the proportion of TB isolates clustered by MIRU-VNTR strain typing—was calculated as the number of clustered isolates/number of clustered+unique isolates. Where there were uncertainties JM consulted with IA.

Authors were contacted if TB incidence rate was not reported. Where no response was received WHO country estimates of TB incidence for the study year were used.[32] As so few studies reported the proportion coinfected with TB/HIV, these estimates for the study country were taken from an European Union-wide survey and WHO country profiles.[33 34] Owing to poor recording of the sampling fraction (the number of isolates typed/the total number of culture positive TB cases diagnosed during the study period (n=19)), whether the

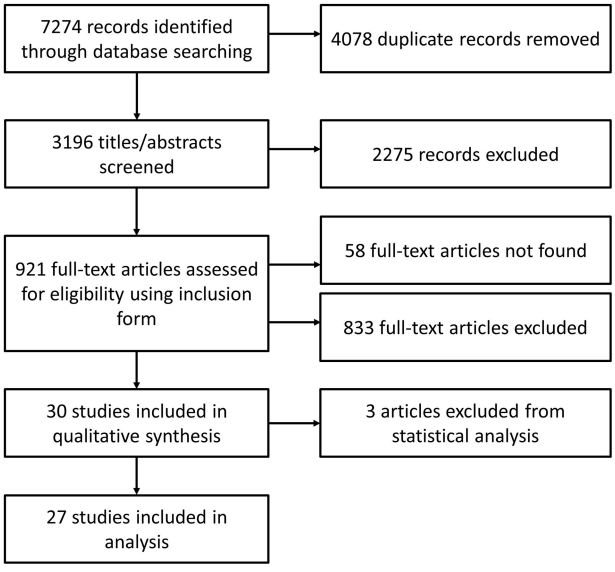


Figure 1 Results of systematic search, screening and data extraction.

Table 1 The study setting and design characteristics of the included articles

	Study setting							Study design								
Reference	Study area and country	TB incidence (per 100 000)	TB/HIV (per 100 000)‡	Previous TB treatment (%)	Pulmonary TB (%)	Maximum cluster size	Clusters of size 2 (%)	Study duration (months)	Study size (clustered +unique isolates)	Culture positive in study population (%)	Culture positive isolates typed (%)	Typing method§	Loci typed¶	Consent required	Risk of bias*	Clustering (%)†
51	New South Wales, Australia	6.7	0.2	0.0	63.7			36	1128			m24	N	no	low	20.1
40	Tabriz and Orumieh, Azarbaijan	26.0		5.2	87.0	5	81.8	12	156		94.5	m15	O	no	low	32.7
52	Brussels-Capital Region, Belgium	35.2	5.1	10.8		23	64.2	24	530	86.1	87.9	m24	N	no	low	29.6
53	Brussels-Capital Region, Belgium	35.2	5.1		100			39	802	81.8	84.7	m24s	N	no	low	28.8
54	Ontario, Canada	4.8	0.4			18	58.8	65	2016			m24s	N	no	low	23.1
37	Changping District, Beijing, China		0.3		100	0		30	318	31.5	94.6	m24	N	no	high	0.0
38	Croatia	19.0	0.1			45	48.3	36	1587			m15	N	no	high	62.8
55	Amhara region, Northwest Ethiopia		24.0	17.6	100	13		5	244			m24	N	yes	low	45.1
56	Finland	5.0	0.0			20		48	1048	75.4	99.4	m15s		no	low	33.9
57	Hamburg, Germany	12.7					45.5	12	154	78.2	91.1	m24s	N	no	low	22.1
45	Schleswig-Holstein, Germany	3.2	0.1			22	44.4	48	277			m24s	N	no	high	27.1
58	South West Ireland	15.3	3.3		82.7	12		36	171	79.5	96.1	m24s	N	no	low	27.5
59	South Tawara, Kiribati	370.0		4.1	100	25	55.6	24	73	45.4	98.6	m24s	N	yes	low	75.3
60	Netherlands	6.5	0.2				57.2	60	3978		100.1	m24	N	no	low	46.7
41	Kharkiv, Russia	94.0	3.8	63.3	100	10	50.0	3	98		100	m15	O	yes	high	31.6
61	Eastern province, Saudi Arabia	4.0			73.1	24	19.0	24	522			m24s	N	no	low	40.2
62	Singapore	40.5	1.2			21	48.0	24	1128	82.0	34.5	m24s	N	no	low	30.8
63	Slovenia	10.6	0.0			6		12	196	94.4	97.5	m24s	N	no	low	36.2
47	Almeria, Spain	26.0	6.0			8		27	281		81.9	m15	N	no	high	43.1
64	Sweden	4.8	0.1			10		36	406			m24s	N	no	low	21.2
65	Mubende, Uganda		86.0	31.1	87.8	11	70.0	6	67	21.5	90.5	m15s	N	yes	low	35.8
42	East Lancashire, UK	18.3	8.2			13	58.3	102	332	48.5	69.9	m15	O	no	low	42.8
39	UK		8.2		42.3	12	50.0	48	102	90.7	87.2	m15	O	no	low	30.4
66	London, UK	44.9	8.2					9	964	36.0	100	m24	N	no		37.0
43	Midlands, UK	15.0	8.2					48	4207	58.3	100	m15	O	no		61.2
44	Odessa and Nikolaev, Ukraine	80.4	3.9	34.2	100			4	225			m15	O	yes**	low	60.4
67	Hanoi, Vietnam	146.0	10.0	0.0	100			20	465	92.7	91.9	m15s	N	yes	low	55.3

*Risk of bias was assessed using the STROME-ID checklist. Studies scoring <20 were categorised as have a high risk of bias. See online supplementary appendix 2 for STROME-ID scores.
†The proportion of clustering was calculated as the number of clustered isolates/number of clustered+unique isolates.
‡Estimates from of the prevalence of TB/HIV coinfection in the study country.[33 34]
§15=15 MIRU-VNTR loci (made up of the 'old 12' or 'new 12' defined in the footnote below), 24=24 MIRU-VNTR loci (ETR A, B, C, D, E; MIRU 2, 10, 16, 20, 23, 24, 26, 27, 39, 40; VNTR 424, 1955, 2163b, 2347, 2401, 3171, 3690, 4052, 4156), S=with Spoligotyping.
¶O=old 12 MIRU loci (MIRU 2, 4, 10, 16, 20, 23, 24, 26, 27,30, 31, 39, 40), N=new 12 MIRU loci (MIRU 10, 16, 26, 31, 40 +Mtub 04, 21, 39+ETR A C+QUB 11b, 26).
**11.3% did not consent to being part of the study. The other studies that required consent for isolates to be typed did not report the refusal rate.
ETR, Exact Tandem Repeat; TB, tuberculosis.

study required the consent of participants (yes/no) was included as a proxy for (low/high) sampling fraction. The risk of bias within each study was assessed using the STROME-ID checklist.[35]

Data were analysed in Stata V.12. Where studies reported data from more than one set of loci, the method with the highest discriminatory value was included (ie, MIRU-VNTR 24 would be chosen over MIRU-VNTR 15, and MIRU-VNTR 15 plus Spoligotyping would be chosen over MIRU-VNTR 15 alone; n=8). This review was not concerned with summary measures of clustering, but factors that influenced clustering; therefore articles must have included at least one of the covariates. Continuous variables were transformed where the distribution was skewed. The proportion clustered was transformed using the Freeman Tukey transformation.[36] Study heterogeneity was assessed using a forest plot and the χ^2 test of heterogeneity. Univariable meta-regression analyses were carried out to determine the effect of the study design covariates on the proportion of clustered isolates. All covariates in the analysis were hypothesised to influence the proportion clustered a priori.

Sensitivity analyses were conducted to see the effect of removing studies reporting 0% clustering, with only extrapulmonary TB cases, only *Mycobacterium bovis* cases, studies using the 'old 12' MIRU loci as part of their 15 loci, and studies assessed as having a high likelihood of bias (STROME-ID score less than 20).

RESULTS

The search identified 7274 references resulting in 27 studies (25 journal articles and 2 conference abstracts) included after deduplication and title/abstract/full text screening (figure 1). The main characteristics of the included studies are shown in table 1. Studies were published between 2007 and 2014 and the clustering reported varied from 0%[37] to 62.8%.[38] In all studies, clustered isolates were defined as having identical strain types based on the MIRU-VNTR loci typed, with or without Spoligotyping. Seventeen studies included isolates from newly diagnosed TB cases, three studies reported including isolates from new and chronic cases of TB, and seven did not report this information. In addition, 10 studies did not include repeat isolates from the same patient, one study included a repeat isolate from one patient and the remaining 17 did not report whether repeat isolates were included or not. Furthermore, four studies included isolates with missing loci in the cluster analysis, whereas four excluded isolates with missing loci and the remaining 20 did not report how they dealt with missing loci. The number of studies reporting each variable of interest is shown in table 2. STROME-ID scores can be found in online supplementary appendix 2.

A forest plot shows the spread of clustering reported by number of loci and additional typing method (figure 2). Significant heterogeneity was identified between the studies (p<0.001), suggesting that a metaregression would be an appropriate analysis.

The univariable metaregression shows evidence for the proportion of clustering to decrease as the number of MIRU-VNTR loci typed increased from 15 to 24 (p=0.04; table 3), accounting for 14% of the between study variation, and to increase when the study participants consented to being included in the study (p=0.03), accounting for 14% of the between study variation. The proportion of clustering increased as the TB incidence in the population increased (p=0.007, adjusted R^2=26.7). There was also evidence for the proportion of clustering

Table 2 The number of studies that reported the variables of interest

	Reported	Missing
Study setting		
TB incidence	8	15
TB/HIV coinfection	5	22
Previous TB treatment	9	18
Proportion pulmonary TB	14	13
Maximum cluster size	19	8
Percentage of clusters with 2 cases	14	13
Study design		
Study duration	27	0
Study size	27	0
Percentage of population that is culture positive	15	12
Percentage of culture positive typed	19	8
24 loci (compared to 15)	27	0
Repeat isolates	12	15
Missing loci	8	19
Double alleles	1	26
Consent required	6*	21
Epidemiological information	6	21

*Only one study reported the consent rate.
TB, tuberculosis.

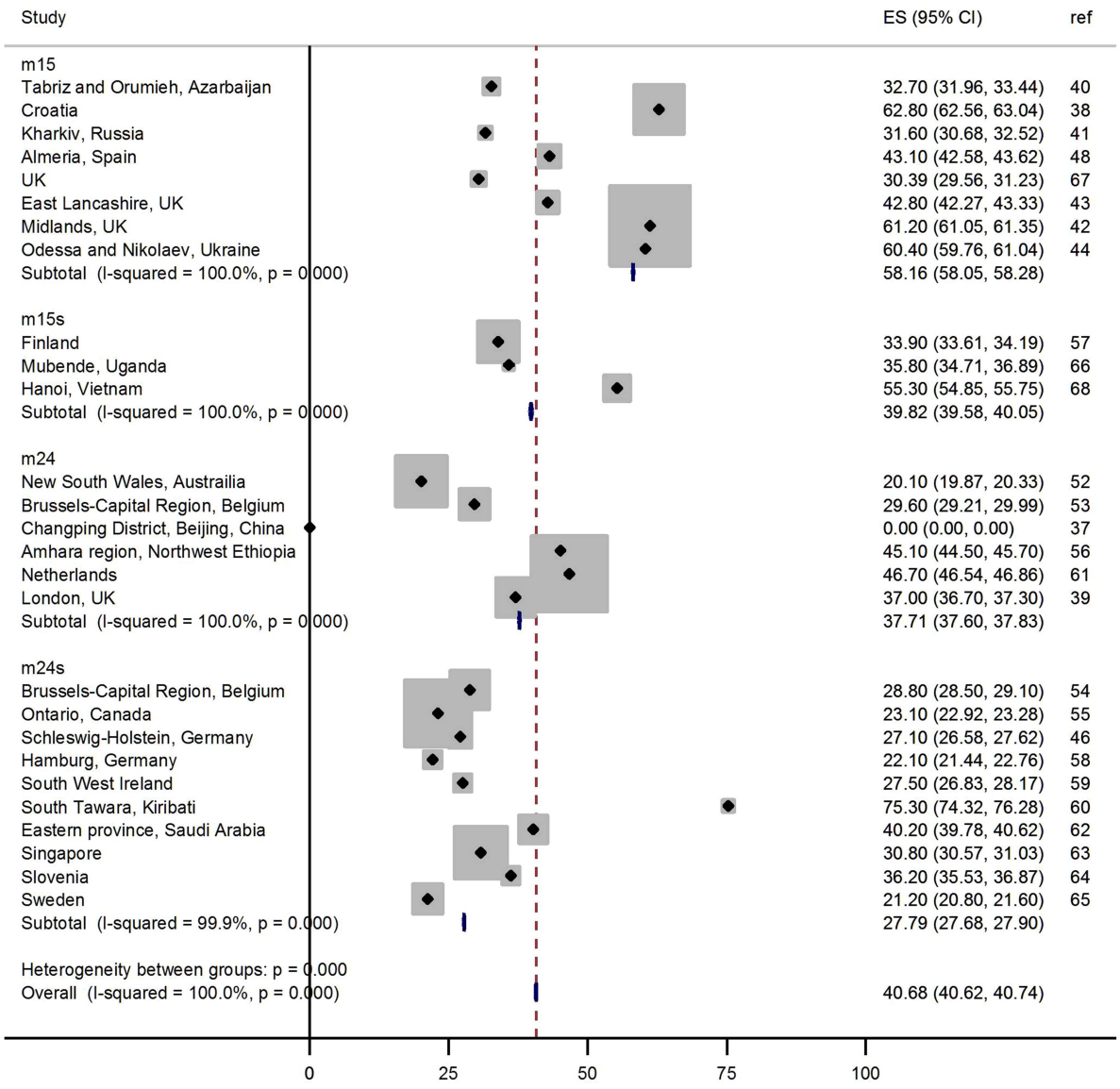


Figure 2 Forest plot showing the proportion of clustering reported in each study by the number of MIRU-VNTR loci typed. The number of loci typed is categorised into 15 loci (m15), 15 loci with Spoligotyping (m15 s), 24 loci (m24) and 24 loci with Spoligotyping (m24 s). The study reference is shown in the right hand column.

to increase as the maximum cluster size increased (p=0.001), accounting for 48% of between study variation. There was no evidence of the other study design or study setting variables significantly influencing the proportion clustered. Though non-significant (p>0.05), the TB/HIV coinfection rate in the population explained 2% of the between study variation. Too few studies included information on the proportion of clusters containing two cases, proportion of the study sample with previous TB or with pulmonary TB, so these could not be included in the analysis (table 2).

Sensitivity analyses to examine the effect of excluding studies reporting 0% clustering,[37] only *M. bovis* cases,[39] studies using the 'old 12' MIRU loci,[39–44] and studies assessed as having a high risk of bias,[37 38 45–47] did not generally change the results. The proportion of culture-positive TB in the population remained insignificant but explained 2.6% of the between study variation when excluding 0% clustering (p=0.278 and adjusted R^2=2.62). Similarly, the proportion of culture-positive TB in the population remained insignificant but explained 2.6% of the between study variation when excluding studies with the highest risk of bias (p=0.278 and adjusted R^2=2.62). The number of loci typed became non-significant, but explained 9.6% and 10.5% of the between study variation when excluding studies using the 'old 12' loci and the highest risk of bias, respectively (p=0.106, adjusted R^2=9.63; p=0.111, adjusted R^2=10.51, respectively).

DISCUSSION

This review identified 27 studies that met the inclusion criteria. We illustrate that the interpretation of studies

Table 3 Univariable metaregression showing the coefficients for change in the proportion of clustering and the percentage of between-study variation explained by variables describing the study design and setting

	n	Coefficient*	CI	p Value	Adjusted R^2†
Study setting					
TB incidence	23	0.14	0.04 to 0.24	0.007	26.74
TB/HIV coinfection	23	0.04	−0.03 to 0.11	0.246	2.00
Maximum cluster size	19	0.20	0.09 to 0.30	0.001	48.20
Study design					
Study duration	27	−0.02	−0.09 to 0.06	0.677	−3.37
Percentage of population that is culture positive	15	0.34	−1.23 to 1.96	0.661	−5.92
Percentage of culture positive typed	19	0.22	−1.08 to 1.52	0.725	−5.41
Study size	27	0.03	−0.11 to 0.16	0.702	−3.31
24 loci (compared to 15)	27	−0.30	−0.59 to −0.01	0.04	13.58
Consent required	27	0.38	0.04 to 0.72	0.029	14.41

*Coefficients for the change in the proportion of clustering for each covariate. For example, for a one unit increase in maximum cluster size, the proportion of clustering increases by 0.2.
†The proportion of between-study variation explained by the univariate metaregression.
TB, tuberculosis.

using MIRU-VNTR to estimate clustering is subject to bias relating to study design and setting; however, there were insufficient data available to fully explore this impact.

As expected, we found that the proportion of clustering decreased with a greater number of MIRU-VNTR loci typed, with increasing TB incidence and with increasing maximum cluster size. We found that requiring consent to type patient isolates increased the proportion of clustering, which is not expected, given that the sampling fraction would be lower in these studies.

The other study design variables included in this analysis, such as study duration, did not significantly influence the proportion of isolates that were clustered, contrary to previous findings.[12] This is likely to be because of a lack of good quality evidence: of the 27 studies that met the inclusion criteria for the review, none reported all the variables of interest, reducing the power of the analysis and precluding multivariable metaregression (table 2). Importantly, key details of cluster analyses were not reported consistently across the studies, such as whether repeat isolates from the same patients were included, or typing profiles with missing loci were included, introducing new, unmeasured biases. In addition, the range of the variables may have been too limited to show any impact on clustering estimates. For example, the proportion of culture-positive isolates typed ranged from 34.5% to 100%, with 17 of the 19 studies reporting this variable from 81.9% to 100%. Furthermore, most of the studies (17/27=63%) were from low TB burden settings and therefore may be reflecting the rate at which imported cases have matching strain types by chance, rather than rates of recent transmission.

The sensitivity analysis suggested that, when excluding the studies with the greatest risk of bias, the culture-positivity in the population might explain a small amount of the between-study variation. This is consistent with estimates of the influence of sampling on the proportion of clustering using *IS*6110 RFLP typing.[48] In the sensitivity analysis excluding studies that used the 'old 12' loci, the effect of the number of loci typed becomes non-significant. This is likely because studies using the 'old 12' accounted for six out of 10 studies reporting 15 loci, reducing the number of studies and the power of the model.

This study is a timely evaluation of the impact of study design on estimates of TB clustering using MIRU-VNTR strain typing because it has been incorporated into national typing services globally.[23 49] The findings are relevant where strain typing is used to evaluate TB control systems across different settings because the proportion of clustering is influenced by the number of loci typed, the TB incidence and the maximum cluster size. Given that strain typing methods are advancing beyond MIRU-VNTR typing and that the application of whole genome sequencing to TB control and public health strategies has been demonstrated,[9–11 50] it is important that the biases in the analysis of such methods are explored and compared. Understanding how to design and compare research studies for public health will greatly improve the benefit gained from newer technologies.

The strength of this meta-analysis was limited by the (lack of) detail reported by the included studies. This review has highlighted the need for better quality reporting in primary studies to enable future reviews to be more robust. Recently published standards for reporting of molecular epidemiology for infectious diseases should improve the quality of reporting.[35] This review is further limited by our inability to access 58 of the title/abstract screened articles for full text screening.

The use of TB strain typing as a public health tool in TB control programmes is increasing globally. We have identified a lack of good quality studies that can contribute to our understanding in interpreting the molecular typing of TB. We have also shown that the proportion of clustering derived from MIRU-VTNR typing is influenced by the number of loci typed, whether consent is

required to type isolates, TB incidence in the study setting and the maximum cluster size, highlighting these as important considerations in the design and interpretation of future studies.

Author affiliations
[1]Department of Infection and Population Health, University College London, London, UK
[2]Centre for Infectious Disease Surveillance and Control, Public Health England, London, UK
[3]Clinical Trials Unit, Medical Research Council, London, UK
[4]Division of Global Health Equity, Brigham and Women's Hospital and Department of Epidemiology, Harvard School of Public Health, Harvard University, Boston, USA
[5]Department of Infection, Centre for Clinical Microbiology, University College London, London, UK

Acknowledgements The authors would like to acknowledge Ross Harris from the Statistics Unit at Public Health England for his advice on metaregression.

Contributors JM drafted the article and PS, IA, TDM and TC revised it critically for important intellectual content. All authors made substantial contributions to the conception and design of the review, and the analysis and interpretation of data. All authors approved the final version for publication.

Funding Public Health England and University College London Impact Studentship. JM is funded through a Public Health England and University College London Impact Studentship. IA is funded through a NIHR Senior Research Fellowship

Competing interests None.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement No additional data are available.

REFERENCES

1. Lambregts-van Weezenbeek CSB, Sebek MMGG, van Gerven PJHJ, *et al*. Tuberculosis contact investigation and DNA fingerprint surveillance in The Netherlands: 6 years' experience with nation-wide cluster feedback and cluster monitoring. *Int J Tuberc Lung Dis* 2003;7:S463–470.
2. Borgdorff MW, van den Hof S, Kremer K, *et al*. Progress towards tuberculosis elimination: secular trend, immigration and transmission. *Eur Respir J* 2010;36:339–47.
3. Kik SV, Verver S, Van Soolingen D, *et al*. Tuberculosis outbreaks predicted by characteristics of first patients in a DNA fingerprint cluster. *Am J Respir Crit Care Med* 2008;178:96–104.
4. Small PM, McClenny NB, Singh SP, *et al*. Molecular strain typing of Mycobacterium tuberculosis to confirm cross-contamination in the mycobacteriology laboratory and modification of procedures to minimize occurrence of false-positive cultures. *J Clin Microbiol* 1993;31:1677–82.
5. De Vries G, van Hest RAH, Richardus JH. Impact of mobile radiographic screening on tuberculosis among drug users and homeless persons. *Am J Respir Crit Care Med* 2007;176:201–7.
6. Kamerbeek J, Schouls L, Kolk A, *et al*. Simultaneous detection and strain differentiation of Mycobacterium tuberculosis for diagnosis and epidemiology. *J Clin Microbiol* 1997;35:907–14.
7. Van Embden JD, Cave MD, Crawford JT, *et al*. Strain identification of Mycobacterium tuberculosis by DNA fingerprinting: recommendations for a standardized methodology. *J Clin Microbiol* 1993;31:406–9.
8. Supply P, Allix C, Lesjean S, *et al*. Proposal for standardization of optimized mycobacterial interspersed repetitive unit-variable-number tandem repeat typing of Mycobacterium tuberculosis. *J Clin Microbiol* 2006;44:4498–510.
9. Schürch AC, van Soolingen D. DNA fingerprinting of Mycobacterium tuberculosis: from phage typing to whole-genome sequencing. *Infect Genet Evol* 2012;12:602–9. http://www.ncbi.nlm.nih.gov/pubmed/22067515 (accessed 13 Mar 2012).
10. Gardy JL, Johnston JC, Ho Sui SJ, *et al*. Whole-genome sequencing and social-network analysis of a tuberculosis outbreak. *N Engl J Med* 2011;364:730–9.
11. Walker TM, Ip CL, Harrell RH, *et al*. Whole-genome sequencing to delineate Mycobacterium tuberculosis outbreaks: a retrospective observational study. *Lancet Infect Dis* 2013;13:137–46.
12. Houben RMGJ, Glynn JR. A systematic review and meta-analysis of molecular epidemiological studies of tuberculosis: development of a new tool to aid interpretation. *Trop Med Int Health* 2009;14:892–909.
13. Fok A, Numata Y, Schulzer M, *et al*. Risk factors for clustering of tuberculosis cases: a systematic review of population-based molecular epidemiology studies. *Int J Tuberc Lung Dis* 2008;12:480–92.
14. Borgdorff MW, Van Den Hof S, Kalisvaart N, *et al*. Influence of sampling on clustering and associations with risk factors in the molecular epidemiology of tuberculosis. *Am J Epidemiol* 2011;174:243–51. http://aje.oxfordjournals.org/content/early/2011/05/23/aje.kwr061 (accessed 29 Mar 2012).
15. Glynn JR, Bauer J, de Boer AS, *et al*. Interpreting DNA fingerprint clusters of Mycobacterium tuberculosis. European concerted action on molecular epidemiology and control of tuberculosis. *Int J Tuberc Lung Dis* 1999;3:1055–60.
16. Glynn JR, Crampin AC, Yates MD, *et al*. The importance of recent infection with Mycobacterium tuberculosis in an area with high HIV prevalence: a long-term molecular epidemiological study in Northern Malawi. *J Infect Dis* 2005;192:480–7.
17. De Beer JL, Kremer K, Ködmön C, *et al*. First worldwide proficiency study on variable-number tandem-repeat typing of Mycobacterium tuberculosis complex strains. *J Clin Microbiol* 2012;50:662–9.
18. Maes M, Kremer K, van Soolingen D, *et al*. 24-locus MIRU-VNTR genotyping is a useful tool to study the molecular epidemiology of tuberculosis among Warao Amerindians in Venezuela. *Tuberculosis (Edinb)* 2008;88:490–4.
19. Sougakoff W. Molecular epidemiology of multidrug-resistant strains of Mycobacterium tuberculosis. *Clin Microbiol Infect* 2011;17:800–5.
20. Weniger T, Krawczyk J, Supply P, *et al*. MIRU-VNTRplus: a web tool for polyphasic genotyping of Mycobacterium tuberculosis complex bacteria. *Nucleic Acids Res* 2010;38:W326–31.
21. Supply P. MIRU-VNTR typing: the new international standard for TB molecular epidemiology Symposium of the Institut Pasteur de Tunisia, 2010.
22. Van Soolingen D, Borgdorff MW, de Haas PE, *et al*. Molecular epidemiology of tuberculosis in the Netherlands: a nationwide study from 1993 through 1997. *J Infect Dis* 1999;180:726–36.
23. Cowan LS, Diem L, Monson T, *et al*. Evaluation of a two-step approach for large-scale, prospective genotyping of Mycobacterium tuberculosis isolates in the United States. *J Clin Microbiol* 2005;43:688–95.
24. Centers for Disease Control and Prevention. New CDC program for rapid genotyping of Mycobacterium tuberculosis isolates. *JAMA* 2005;293:2086–2086.
25. Bauer J, Kok-Jensen A, Faurschou P, *et al*. A prospective evaluation of the clinical value of nation-wide DNA fingerprinting of tuberculosis isolates in Denmark. *Int J Tuberc Lung Dis* 2000;4:295–9.
26. Bauer J, Yang Z, Poulsen S, *et al*. Results from 5 years of nationwide DNA fingerprinting of Mycobacterium tuberculosis complex isolates in a country with a low incidence of M. tuberculosis infection. *J Clin Microbiol* 1998;36:305–8.
27. Zolnir-Dovc M, Poljak M, Erzen D, *et al*. Molecular epidemiology of tuberculosis in Slovenia: results of a one-year (2001) nation-wide study. *Scand J Infect Dis* 2003;35:863–8.
28. Hanekom M, van der Spuy GD, Gey van Pittius NC, *et al*. Discordance between mycobacterial interspersed repetitive-unit-variable-number tandem-repeat typing and IS6110 restriction fragment length polymorphism genotyping for analysis of Mycobacterium tuberculosis Beijing strains in a setting of high incidence of tuberculosis. *J Clin Microbiol* 2008;46:3338–45.
29. Supply P, Lesjean S, Savine E, *et al*. Automated high-throughput genotyping for study of global epidemiology of Mycobacterium tuberculosis based on mycobacterial interspersed repetitive units. *J Clin Microbiol* 2001;39:3563–71.
30. Gopaul KK, Brown TJ, Gibson AL, *et al*. Progression toward an improved DNA amplification-based typing technique in the study of Mycobacterium tuberculosis epidemiology. *J Clin Microbiol* 2006;44:2492–8.

31. Hunter PR, Gaston MA. Numerical index of the discriminatory ability of typing systems: an application of Simpson's index of diversity. *J Clin Microbiol* 1988;26:2465–6.
32. WHOITB data (n.d.). WHO. http://www.who.int/tb/country/en/index.html (accessed 12 Dec 2012).
33. Kruijshaar ME, Pimpin L, Abubakar I, *et al*. The burden of TB-HIV in the EU: how much do we know? A survey of surveillance practices and results. *Eur Respir J* 2011;38:1374–81.
34. World Health Organization (n.d.) WHO Tuberculosis Country Profiles. http://www.who.int/tb/country/data/profiles/en/
35. Field N, Cohen T, Struelens MJ, *et al*. Strengthening the Reporting of Molecular Epidemiology for Infectious Diseases (STROME-ID): an extension of the STROBE statement. *Lancet Infect Dis* 2014;14:341–52.
36. Freeman MF, Tukey JW. Transformations related to the angular and the square root. *Ann Math Statist* 1950;21:607–11.
37. Guang-ming D, Zhi-guo Z, Peng-ju D, *et al*. Differences in the population of genetics of Mycobacterium tuberculosis between urban migrants and local residents in Beijing, China. *Chin Med J* 2013;126:4066–71.
38. Zmak L, Obrovac M, Katalinic Jankovic V. First insights into the molecular epidemiology of tuberculosis in Croatia during a three-year period, 2009 to 2011. *Scand J Infect Dis* 2014;46:123–9.
39. Mandal S, Bradshaw L, Anderson LF, *et al*. Investigating transmission of Mycobacterium bovis in the United Kingdom in 2005 to 2008. *J Clin Microbiol* 2011;49:1943–50.
40. Asgharzadeh M, Kafil HS, Roudsary AA, *et al*. Tuberculosis transmission in Northwest of Iran: using MIRU-VNTR, ETR-VNTR and IS6110-RFLP methods. *Infect Genet Evol* 2011;11:124–31.
41. Dymova MA, Liashenko OO, Poteiko PI, *et al*. Genetic variation of Mycobacterium tuberculosis circulating in Kharkiv Oblast, Ukraine. *BMC Infect Dis* 2011;11:77.
42. Sails AD, Barrett A, Sarginson S, *et al*. Molecular epidemiology of Mycobacterium tuberculosis in East Lancashire 2001-2009. *Thorax* 2011;66:709–13.
43. Evans J. *Analysis of prevalent Mycobacterium tuberculosis strains in the United Kingdom: detection, distribution and expansion of MIRU-VNTR profiles containing high numbers of isolates*. Vienna, Austria: European Society of Clinical Microbiology and Infectious Diseases, 2010.
44. Nikolayevskyy VV, Brown TJ, Bazhora YI, *et al*. Molecular epidemiology and prevalence of mutations conferring rifampicin and isoniazid resistance in Mycobacterium tuberculosis strains from the southern Ukraine. *Clin Microbiol Infect* 2007;13:129–38.
45. Roetzer A, Schuback S, Diel R, *et al*. Evaluation of Mycobacterium tuberculosis typing methods in a 4-year study in Schleswig-Holstein, Northern Germany. *J Clin Microbiol* 2011;49:4173–8.
46. Dymova MA, Kinsht VN, Cherednichenko AG, *et al*. Highest prevalence of the Mycobacterium tuberculosis Beijing genotype isolates in patients newly diagnosed with tuberculosis in the Novosibirsk oblast, Russian Federation. *J Med Microbiol* 2011;60:1003–9.
47. Alonso-Rodriguez N, Martínez-Lirola M, Sánchez ML, *et al*. Prospective universal application of mycobacterial interspersed repetitive-unit-variable-number tandem-repeat genotyping to characterize Mycobacterium tuberculosis isolates for fast identification of clustered and orphan cases. *J Clin Microbiol* 2009;47:2026–32.
48. Glynn JR, Vyonycky E, Fine PEM. Influence of sampling on estimates of clustering and recent transmission of Mycobacterium tuberculosis derived from DNA fingerprinting techniques. *Am J Epidemiol* 1999;149:366–71.
49. TB Strain Typing Project Board HPA. TB Strain Typing Cluster Investigation Handbook for Health Protection Units 1st Edition. 2011. https://hpaintranet.hpa.org.uk/Content/ProgrammesProjects/HPAProgrammes/HPAKeyHealthProtectionProgrammes/Respiratory/TB/StrainTyping/ (accessed 30 Nov 2011).
50. Walker TM, Monk P, Grace Smith E, *et al*. Contact investigations for outbreaks of Mycobacterium tuberculosis: advances through whole genome sequencing. *Clin Microbiol Infect* 2013;19:796–802.
51. Gurjav U, Jelfs P, McCallum N, *et al*. Temporal dynamics of Mycobacterium tuberculosis genotypes in New South Wales, Australia. *BMC Infect Dis* 2014;14:455–5.
52. Allix-Béguec C, Fauville-Dufaux M, Supply P. Three-year population-based evaluation of standardized mycobacterial interspersed repetitive-unit-variable-number tandem-repeat typing of Mycobacterium tuberculosis. *J Clin Microbiol* 2008;46:1398–406.
53. Allix-Béguec C, Supply P, Wanlin M, *et al*. Standardised PCR-based molecular epidemiology of tuberculosis. *Eur Respir J* 2008;31:1077–84.
54. Tuite AR, Guthrie JL, Alexander DC, *et al*. Epidemiological evaluation of spatiotemporal and genotypic clustering of mycobacterium tuberculosis in Ontario, Canada. *Int J Tuberc Lung Dis* 2013;17:1322–7.
55. Tessema B, Beer J, Merker M, *et al*. Molecular epidemiology and transmission dynamics of Mycobacterium tuberculosis in Northwest Ethiopia: new phylogenetic lineages found in Northwest Ethiopia. *BMC Infect Dis* 2013;13:131.
56. Smit PW, Haanpera M, Rantala P, *et al*. Molecular Epidemiology of Tuberculosis in Finland, 2008-2011. *PLoS ONE* 2013;8:e85027.
57. Oelemann MC, Diel R, Vatin V, *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;45:691–7.
58. Ojo OO, Sheehan S, Corcoran DG, *et al*. Molecular epidemiology of Mycobacterium tuberculosis clinical isolates in Southwest Ireland. *Infect Genet Evol* 2010;10:1110–16.
59. Aleksic E, Merker M, Cox H, *et al*. First Molecular Epidemiology Study of Mycobacterium tuberculosis in Kiribati. *PLoS ONE* 2013;8:e55423. http://www.scopus.com/inward/record.url?eid=2-s2.0-84873163328&partnerID=40&md5=3994b8e5638129b621abc4d7d6d5e3b8 http://dx.doi.org/10.1371/journal.pone.0055423
60. De Beer JL, van Ingen J, de Vries G, *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;51:1193–8.
61. Varghese B, Supply P, Shoukri M, *et al*. Tuberculosis transmission among immigrants and autochthonous populations of the Eastern Province of Saudi Arabia. *PLoS ONE* 2013;8:e77635. http://www.scopus.com/inward/record.url?eid=2-s2.0-84885784886&partnerID=40&md5=4fdbf4015a999a9fcd1a1c31207a75a2
62. Lim LK-Y, Sng LH, Win W, *et al*. Molecular epidemiology of Mycobacterium tuberculosis complex in Singapore, 2006-2012. *PLoS ONE* 2013;8:e84487.
63. Bidovec-Stojkovic U, Zolnir-Dovc M, Supply P. One year nationwide evaluation of 24-locus MIRU-VNTR genotyping on Slovenian Mycobacterium tuberculosis isolates. *Respir Med* 2011;105(Suppl 1):S67–73.
64. Jonsson J, Hoffner S, Berggren I, *et al*. Comparison between RFLP and MIRU-VNTR genotyping of mycobacterium tuberculosis strains isolated in stockholm 2009 to 2011. *PLoS ONE* 2014;9:e95159. http://www.plosone.org/article/fetchObject.action?uri=info%3Adoi%2F10.1371%2Fjournal.pone.0095159&representation=PDF
65. Muwonge A, Malama S, Johansen TB, *et al*. Molecular epidemiology, drug susceptibility and economic aspects of tuberculosis in Mubende District, Uganda. *PLoS ONE* 2013;8:e64745. http://www.scopus.com/inward/record.url?eid=2-s2.0-84878608813&partnerID=40&md5=babbd6d006ca64e327fb19e01b6bc697
66. Hamblion EL, Wynne-Edwards E, Anderson C, *et al*. A summary of strain typing and clustering of TB in London in 2010 and an analysis of the associated risk factors. *Thorax* 2011;66:A88–9.
67. Hang NTL, Maeda S, Lien LT, *et al*. Primary drug-resistant tuberculosis in Hanoi, Vietnam: present status and risk factors. *PLoS ONE* 2013;8:e71867. UNSP.

BMJ Open

Effect of study design and setting on tuberculosis clustering estimates using Mycobacterial Interspersed Repetitive Units-Variable Number Tandem Repeats (MIRU-VNTR): a systematic review

Jessica Mears, Ibrahim Abubakar, Theodore Cohen, Timothy D McHugh and Pam Sonnenberg

BMJ Open 2015 5:
doi: 10.1136/bmjopen-2014-005636

BMC Public Health

RESEARCH ARTICLE Open Access

Prospective evaluation of a complex public health intervention: lessons from an initial and follow-up cross-sectional survey of the tuberculosis strain typing service in England

Jessica Mears[1], Ibrahim Abubakar[1,2], Debbie Crisp[3], Helen Maguire[4,5], John A Innes[6], Mike Lilley[7], Joanne Lord[8], Ted Cohen[9,10], Martien W Borgdorff[11,12], Emilia Vynnycky[2,13], Timothy D McHugh[14] and Pam Sonnenberg[1*]

Abstract

Background: The national tuberculosis strain typing service (TB-STS) was introduced in England in 2010. The TB-STS involves MIRU-VNTR typing of isolates from all TB patients for the prospective identification, reporting and investigation of TB strain typing clusters. As part of a mixed-method evaluation, we report on a repeated cross-sectional survey to illustrate the challenges surrounding the evaluation of a complex national public health intervention.

Methods: An online initial and follow-up questionnaire survey assessed the knowledge, attitudes and practices of public health staff, physicians and nurses working in TB control in November 2010 and March 2012. It included questions on the implementation, experience and uptake of the TB-STS. Participants that responded to both surveys were included in the analysis.

Results: 248 participants responded to the initial survey and 137 of these responded to the follow-up survey (56% retention).
Knowledge: A significant increase in knowledge was observed, including a rise in the proportion of respondents who had received training (28.6% to 67.9%, $p = 0.003$), and the self-rated knowledge of how to use strain typing had improved ('no knowledge' decreased from 43.2% to 27.4%).
Attitudes: The majority of respondents found strain typing useful; the proportion that reported strain typing to be useful was similar across the two surveys (95.7% to 94.7%, $p = 0.67$).
Practices: There were significant increases between the initial and follow-up surveys in the number of respondents who reported using strain typing (57.0% to 80.5%, $p < 0.001$) and the proportion of time health protection staff spent on investigating TB (2.74% to 7.08%, $p = 0.04$).

Conclusions: Evaluation of a complex public health intervention is challenging. In this example, the immediate national roll-out of the TB-STS meant that a controlled survey design was not possible. This study informs the future development of the TB-STS by identifying the need for training to reach wider professional groups, and argues for its continuation based on service users' perception that it is useful. By highlighting the importance of a well-defined sampling frame, collecting baseline information, and including all stakeholders, it provides lessons for the implementation of similar services in other countries and future evaluations of public health interventions.

Keywords: Tuberculosis, Strain typing, MIRU-VNTR, Complex intervention, Service evaluation

* Correspondence: P.Sonnenberg@ucl.ac.uk
[1]Department of Infection and Population Health, University College London, London, UK
Full list of author information is available at the end of the article

Background

Complex public health interventions – interventions involving multiple interacting components – when applied at a national level, are often implemented in a way that makes evaluating them with rigorously designed trials difficult [1]. Instead, they require a more pragmatic approach using the available data [2].

Molecular typing of *Mycobacterium tuberculosis* is a tool for TB surveillance and control. It has been used in combination with epidemiological information to identify outbreaks [3], identify new routes of transmission [4], refute suspected transmission [5,6], evaluate TB control programmes [7,8] and detect laboratory cross contamination [9,10].

The National Tuberculosis Strain Typing Service (TB-STS) is a complex public health intervention involving laboratory, public health and clinical services across England and was introduced in January 2010 [11]. A mixed-method prospective evaluation of the acceptability, implementation, effectiveness and cost-effectiveness of the service was undertaken [12]. Here we report in detail on one component of the evaluation: a cross-sectional initial and follow-up survey of those delivering and using the TB-STS to assess their knowledge, and to understand the impact of the service on changes in attitudes and practices associated with strain typing.

Methods

Intervention

A full description of the TB-STS, with laboratory guidelines for MIRU-VNTR strain typing and reporting [13] and a handbook for public health actions relating to cluster investigations (TB strain typing and cluster investigation handbook [14]) can be found on the Health Protection Agency website [11]. Briefly, the TB-STS involves prospectively typing the first *M. tuberculosis* isolate from every culture-confirmed tuberculosis (TB) patient using 24 locus Mycobacterial Interspersed Repetitive Units-Variable Number Tandem Repeats (MIRU-VNTR), a standardised molecular typing method [15]. Based on the strain type result, patients are grouped into 'clusters' [13,14] which are reported to the Health Protection Units (HPUs). If a cluster meets a certain threshold, as outlined in the TB strain typing and cluster investigation handbook, [14] then a cluster investigation is launched to try to establish epidemiological links between the clustered patients, thereby identifying the transmission setting and/or an outbreak. As part of a cluster investigation the HPU may decide to carry out enhanced contact tracing or screening around the patients in the cluster or the identified transmission setting. By combining patients' strain type with epidemiological information the TB-STS aims to inform public health decision-making at the local level.

The various components of the TB-STS were implemented at different times (but always on a national scale): prospective strain typing was introduced across England in January 2010; one cluster investigator was appointed in January 2010 and the remaining two were appointed in January 2011; the training programme for health protection staff working in HPUs was carried out between January 2011 and February 2012, consisting of a seminar at the national Health Protection Conference, an online seminar, a workshop conducted at each HPU, the publication of the handbook [14] and a Q&A sheet [11] (in December 2010); and the software linking patients' electronic TB record and strain typing result with information from clusters investigations was not developed during this study period.

Study design

An initial survey was conducted in November 2010 and a follow-up survey in March 2012 using a web-based survey questionnaire (www.objectplanet.com/opinio). The target population were all public health staff, chest/respiratory physicians and TB nurses working in TB control in England. Questions were asked about the knowledge (awareness of the service, training, resources and self-reported knowledge), attitudes (perceived usefulness of the service) and practice (if and how strain typing is accessed and used, and its associated workload). All questions and possible responses are available in the appendix (Additional file 1). The survey was piloted with a nurse, a physician and a public health specialist. The initial survey was emailed to all users of the TB notification system [16] and to staff responsible for TB control in HPUs who were asked to pass it on to their local TB teams; the sampling frame could not be enumerated. The follow-up survey was emailed to respondents to the initial survey.

Analysis

Participants that responded to both surveys were included in the statistical analysis. Responses from people working at national, regional or PCT-level, including cluster investigators, and people working in Wales were excluded from this analysis. We compared the knowledge, attitudes and practices of public health and clinical staff working on TB control in the initial and follow-up surveys by calculating and comparing medians and inter-quartile ranges (IQR), and means and standard deviations (SD), and using two-sample t-tests, chi^2 tests and logistic regression, where appropriate. Calculations exclude item non-responses. Analyses were conducted overall, by professional category and the TB incidence of the HPU area in which respondents worked (low, medium and high incidence defined as an annual notification rate of <10/100,000 population, 10 to 19/100,000 and ≥20/100,000 respectively).

Ethical considerations

The study was classified as a service evaluation by University College London Hospital Foundation Trust therefore specific ethical approval was not required.

Results

Survey responses

There were 248 responses to the initial survey, 137 responses to the follow-up survey (55% retention), and for 124 we have responses to both the initial and follow-up surveys (Figure 1). Respondents to the initial survey who did not respond to the follow-up survey were not significantly different to those that responded to both: no particular profession, full-time/part-time position or those working in areas with different TB incidences was more (or less) likely to respond to the follow-up survey, and there was no significant difference between the proportion of people who had heard of the TB-STS or had access to strain typing at the time of the initial survey (Table 1). Respondents were from all nine regions of England and covered 24 (of 26) HPUs.

Knowledge

Between the initial and follow-up surveys there were increases in the proportion of respondents who had heard of the TB-STS, had access to strain typing results, had received training, and had access to training resources (Table 2). The self-rated knowledge of how to use strain typing also increased over time (Figure 2). Nurses reported lower average knowledge in both surveys compared to physicians and health protection staff.

Attitudes

69 people (69/124 = 56%) from the initial survey and 95 people (95/124 = 77%) from the follow-up survey reported that they used strain typing. Opinions of the usefulness of TB strain typing was high amongst all respondents and did not change between the surveys (95.7% to 94.7%, p = 0.667; Table 3). A greater proportion of respondents from low TB incidence areas found strain typing useful, compared to those working in high TB incidence areas (97.4% compared to 89.3% in the follow-up survey, respectively), though this result was not statistically significant (OR = 0.13, 95% CI 0.014-1.128, p = 0.075).

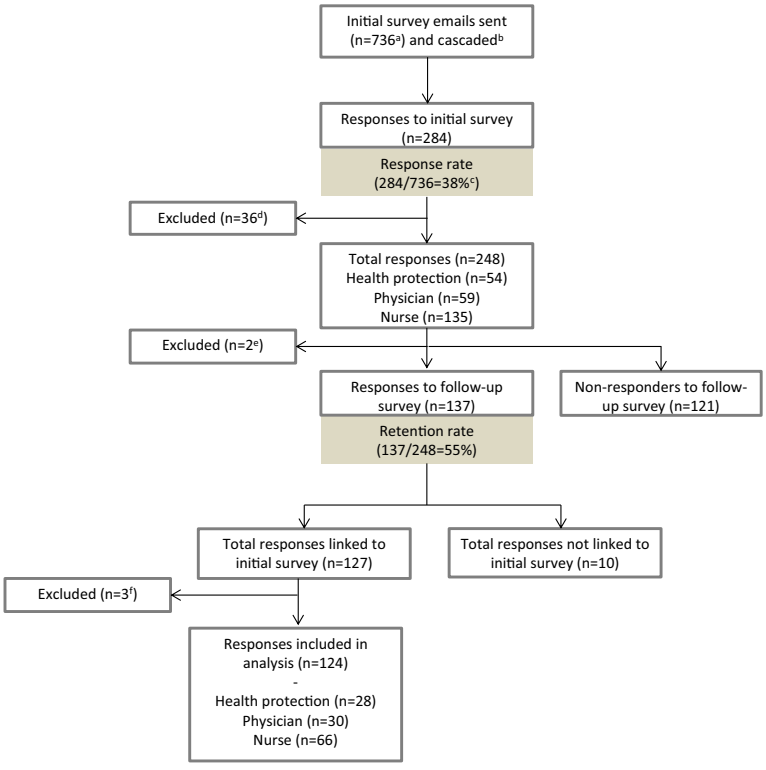


Figure 1 Responses to the initial and follow-up surveys. [a]The email was sent to all users of the Enhanced TB Surveillance database. This included all administrative staff as well as well as staff working at national, regional and Primary Care Trust level, for whom the survey may not be relevant. [b]It is not known how many people received the email via through the HPU cascade. [c]This response rate is an underestimation because of the denominator used. [d]Respondents working at national, regional or PCT-level (n = 27) and those from Wales (n = 9) were excluded from this analysis. [e]Email addresses not available from the initial survey (n = 2). [f]In some cases it was not possible to link the follow-up response to the initial response (n = 10). Respondents working at national, regional or PCT-level and those from Wales (n = 4) were excluded from this analysis.

Table 1 Characteristics of responders and non-responders to the follow-up survey

		Initial and follow-up responses[a]		Non-responders to the follow-up survey	
		N	%	n	%
Total		124		121	
Profession	HPU	28	22.6	23	19.0
	Physician	30	24.2	29	24.0
	Nurse	66	53.2	69	57.0
TB incidence[b]	Low	56	45.2	50	42.0
	Medium	33	26.6	32	26.9
	High	35	28.2	37	31.1
Work time	Full-time	95	79.2	87	77.0
	Part-time	25	20.2	26	21.5
Heard of the TB-STS		105	85.4	100	84.7
Access to strain typing		90	72.6	99	81.8

[a]Using the information reported in the initial survey.
[b]Area where respondents worked is defined as low, medium and high TB incidence: <10/100,000, 10-19/100,000, ≥20/100,000 population, respectively.
There were no significant differences between characteristics of non-responders and responders, including access to strain typing (81.8 % vs. 72.6 %, chi^2 test p = 0.085).

Practices

Figure 3 shows a significant increase in the number of respondents that reported using strain typing between the initial and follow-up surveys. There was an increase in the number of respondents who reported using strain typing to identify links between cases (65.3% to 78.2%, p = 0.02; the most common use), disprove links between cases (46.8% to 58.9%, p = 0.06) and to justify stopping contact tracing (20.2% to 30.7%, p = 0.06) (Table 4).

Table 5 shows workload reported by nurses and health protection staff. For the nurses, no significant changes in contact tracing workload were reported.

Health protection staff reported a significant increase in the mean number of investigations initiated because of

Table 2 Knowledge: Awareness to the TB-STS and access to strain typing data and resources

		Initial survey		Follow-up survey		
		n	%	n	%	p-value[d]
Heard of the TB-STS[a]	Total	105	85.4	123	99.2	<0.001
Profession	Health protection	28	100	28	100	.
	Physician	20	66.7	30	100	0.001
	Nurse	57	86.4	65	98.5	0.015
TB incidence	Low	49	87.5	56	100	0.006
	Medium	24	72.7	32	97.0	0.010
	High	32	91.4	35	100	0.077
Access to strain typing data[b]	Total	90	72.6	108	87.1	0.004
Profession	Health protection	26	92.9	27	96.4	0.553
	Physician	21	70.0	23	76.7	0.559
	Nurse	43	65.2	58	87.9	0.002
TB incidence[c]	Low	38	67.9	47	83.9	0.047
	Medium	24	72.7	28	84.9	0.228
	High	28	80.0	33	94.3	0.074
Access to training	(health protection staff)	8	28.6	19	67.9	0.003
Access to resources	(health protection staff)	16	57.1	23	82.1	0.042

[a]Have you heard of the TB-STS (apart from in this survey)? *(Yes / No)*.
[b]Do you have access to strain typing data? *(Yes / No)*.
[c]Area where respondents worked is defined as low, medium and high TB incidence: <10/100,000, 10-19/100,000, ≥20/100,000 population, respectively.
[d]chi^2 test of significance comparing responses from the initial and follow-up surveys.

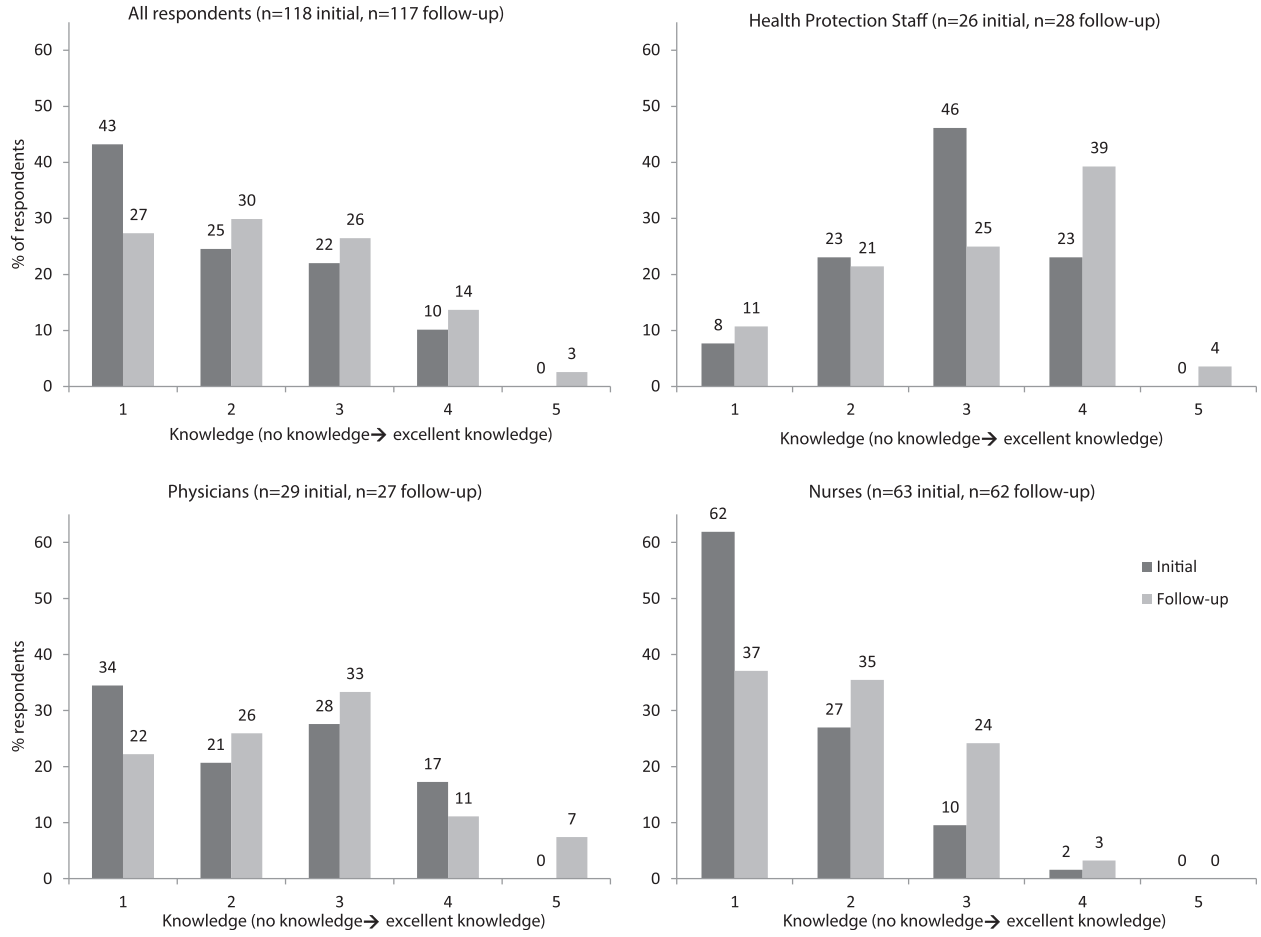


Figure 2 Self-reported knowledge of strain typing. Self-reported knowledge about how to use strain typing was scored on a scale of 1 to 5, where 1 represented 'no knowledge' and 5 represented 'excellent knowledge'. Dark bars represent responses to the initial survey and light bars represent responses to the follow-up survey.

Table 3 Attitudes: Number and proportion of respondents that reported strain typing to be useful[a]

		Initial survey				Follow-up survey[b]				
		Useful		Not useful		Useful		Not useful		
		n	%	n	%	n	%	n	%	P[d]
Total respondents that reported using strain typing		66	95.7	3	4.3	89	94.7	5	5.3	0.667
Profession	Health protection	22	95.7	1	4.3	24	96.0	1	4.0	0.952
	Physician	16	100	0	0.0	20	95.2	1	4.8	0.464
	Nurse	28	93.3	2	6.7	45	93.8	3	6.3	0.942
TB incidence[c]	Low	31	100	0	0.0	38	97.4	1	2.6	0.450
	Medium	16	94.1	1	5.9	26	96.3	1	3.7	0.736
	High	19	90.5	2	9.5	25	89.3	3	10.7	0.892

[a]The following question was asked to respondents who reported that they used strain typing data for TB control (Figure 3): Do you find the strain typing information useful? *(Very useful / Quite useful / Not very useful / Useless)* 'Very useful' and 'Quite useful' are grouped into 'useful', and 'Not very useful' is presented as 'Not useful'. No one reported finding the strain typing 'useless' in either survey.
[b]One response was missing from the follow-up survey.
[c]Area where respondents worked is defined as low, medium and high TB incidence: <10/100,000, 10-19/100,000, ≥20/100,000 population, respectively.
[d]chi^2 test for significance comparing responses from the initial and follow-up surveys, missing items were excluded.

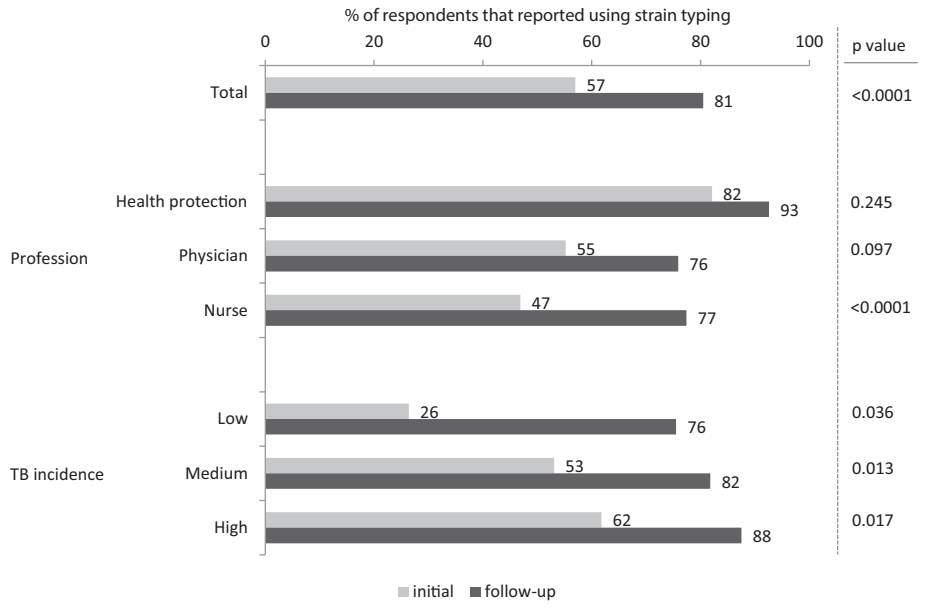


Figure 3 Practices: Respondents that use strain typing for TB control[a]. The proportion respondents that reported using strain typing for TB control. Dark bars represent responses to the initial survey and light bars represent responses to the follow-up survey. P-values from chi^2 tests for significance comparing initial and follow-up proportions are shown. [a]How often do you use strain typing data in your case management of outbreak investigation? *Never / For few cases / For about half of cases / For many cases / For every case.* 'For few cases', 'for about half of cases', 'for many cases' and 'for every case' were grouped to show the proportion of respondents that use strain typing.

epidemiological links between patients over a three month period (mean 0.5 to 2.8, p = 0.04) and the mean number of these investigations for which strain typing was used to provide more information (0.6 to 1.8, p = 0.03), but there was no change in the number that were influenced by the strain typing (1.2 to 0.4, p = 0.17). There was no reported difference in the number of clusters investigated because of their strain type (in high incidence areas a large, but non-significant, decrease was reported) and the number of strain typing investigations that identified epidemiological links between cases remained low (Table 5). Overall, the proportion of time health protection staff spent on cluster investigations increased significantly (from 2.7% to 7.2%, p = 0.04).

There was no reported change over time in the frequency at which physicians were called to incident meetings (a meeting, often multi-disciplinary, held to discuss a TB patient, group or cluster of cases that are of particular concern) (p = 0.503; most reported once every three months or less (65.5% at in the initial survey and 67.9% at follow-up)) and there was no change in the number of physicians who reported strain typing as being relevant to an incident meeting (57.8% to 55.6%, p = 0.875).

Discussion

Main findings

We present results from an initial and follow-up survey assessing the knowledge, attitudes and practices of those implementing and using the TB-STS. There were 124 responses to both the surveys, representing health protection staff and clinic-based physicians and nurses from 24 (of 26) HPUs across England. Strain typing was used by more people, and an increase in knowledge of the TB-STS was reported at the follow-up survey. A change in attitude was not measured as the majority of respondents found strain typing useful to them at both time points. With respect to workload associated with the TB-STS, there was no change over time in the contact tracing activities of nurses or the frequency of incident meetings attended by physicians; however the proportion of time health protection staff spent on investigating TB transmission increased significantly. Despite strain typing being used to provide more information to public health staff at follow-up, there was no increase in epidemiological links identified.

How this relates to previous studies

National TB strain typing services have been introduced in other countries [17-20], but the knowledge, attitudes and practices of users have not been evaluated. However, the impact of strain typing on contact tracing activities in the Netherlands has been assessed [5]. Consistent with this study, we found no change in the workload associated with strain typing for nurses and physicians, even though strain typing was used by more people at the follow-up survey (indicating the successful roll-out

Table 4 Practices: How respondents use strain typing data[a]

		Initial survey		Follow-up survey		
		n	%	n	%	P[c]
Identify clusters and links between cases		81	65.3	97	78.2	0.024
Profession	Health protection	22	78.6	25	89.3	0.275
	Physician	18	60.0	21	70.0	0.417
	Nurse	41	62.1	51	77.3	0.058
TB incidence[b]	Low	34	60.7	41	73.2	0.160
	Medium	20	60.6	28	84.8	0.027
	High	27	81.8	28	84.8	0.771
Disprove clusters and links between cases		58	46.8	73	58.9	0.056
Profession	Health protection	21	75.0	24	85.7	0.313
	Physician	13	43.3	15	50.0	0.605
	Nurse	24	36.4	34	51.5	0.079
TB incidence[b]	Low	27	48.2	33	58.9	0.256
	Medium	15	45.5	22	66.7	0.083
	High	16	48.5	18	54.5	0.632
Justify extended contact tracing		51	41.1	60	48.4	0.250
Profession	Health protection	16	57.1	19	67.9	0.408
	Physician	11	36.7	10	33.3	0.787
	Nurse	24	36.4	31	47.0	0.217
TB incidence[b]	Low	20	35.7	25	44.6	0.335
	Medium	13	39.4	19	57.6	0.139
	High	18	54.5	16	48.5	0.632
Justify stopping contact tracing		25	20.2	38	30.6	0.058
Profession	Health protection	13	46.4	13	46.4	1
	Physician	3	10.0	5	16.7	0.448
	Nurse	9	13.6	20	30.3	0.021
TB incidence[b]	Low	9	16.1	18	32.1	0.047
	Medium	8	24.2	13	39.4	0.186
	High	8	24.2	7	21.2	0.771
To provide more information		34	27.4	44	35.5	0.171
Profession	Health protection	13	46.4	10	35.7	0.415
	Physician	5	16.7	6	20.0	0.739
	Nurse	16	24.2	28	42.4	0.027
TB incidence[b]	Low	15	26.8	19	33.9	0.411
	Medium	8	24.2	12	36.4	0.284
	High	11	33.3	13	39.4	0.615

[a]What do you use strain typing for? (multiple selections possible) (*Don't know / Identify clusters and links between cases / Disprove clusters and links between cases / Justify extended contact tracing / Justify stopping contact tracing / To provide more information / Other (please specify)).*
[b]Area where respondents worked is defined as low, medium and high TB incidence: <10/100,000, 10-19/100,000, ≥20/100,000 population, respectively.
[c]chi^2 test for significance comparing responses from the initial and follow-up surveys.

of the service). This may be because it is difficult to measure marginal changes in workload associated with a particular service where the workforce is already working at full capacity. Health protection staff, however, spent a greater proportion of time on cluster investigations. Given that the Handbook had not been published and all the cluster investigation coordinators were not in position at the time of the initial survey, this is not surprising and suggests that the TB-STS had been integrated into the TB control activities of the HPUs.

Based on evidence from the USA one would expect more possible transmission links to be identified when

Table 5 Practices: the workload associated with the TB-STS for nurses and health protection staff

		TB incidence[a]	Survey	n[b]	median	(IQR)	mean	(SD)	p-value[c]
Nurses	No. contacts screened in the last month	Total	Initial	57	21	(11–36)	37.1	(53.5)	
			Follow-up	55	20	(8–40)	33.9	(45.1)	0.37
		Low	Initial	26	16	(6–35)	23.8	(24.8)	
			Follow-up	23	15	(6–25)	17.2	(13.8)	0.13
		Medium	Initial	17	25	(14–30)	30.2	(26.2)	
			Follow-up	18	23	(15–42)	43.7	(43.7)	0.18
		High	Initial	14	32.5	(14–100)	70.2	(93.3)	
			Follow-up	14	16.5	(10–80)	48.6	(58.9)	0.24
	No. hours spent on contact tracing in the last month	Total	Initial	55	8	(4–16)	12.0	(10.8)	
			Follow-up	52	7.5	(3.5-15.5)	16.1	(41.7)	0.24
		Low	Initial	25	8	(3–14)	10.1	(10.5)	
			Follow-up	21	6	(3–15)	11.5	(14.7)	0.35
		Medium	Initial	16	12	(4–23)	14.4	(11.4)	
			Follow-up	18	7.5	(4–12)	10.2	(7.8)	0.10
		High	Initial	14	9	(6–15)	12.5	(10.8)	
			Follow-up	13	8	(3–16)	31.9	(81.1)	0.19
	% time spent on contact tracing	Total	Initial	57	20	(10–30)	24.2	(16.5)	
			Follow-up	54	20	(10–25)	21.7	(17.6)	0.22
		Low	Initial	26	20	(10–25)	21.2	(16.1)	
			Follow-up	23	20	(6–25)	21.8	(19.5)	0.45
		Medium	Initial	17	20	(20–30)	24.1	(13.8)	
			Follow-up	17	20	(10–25)	19.4	(10.4)	0.14
		High	Initial	14	30	(15–40)	30.0	(19.7)	
			Follow-up	14	20	(10–40)	24.4	(21.7)	0.24
Health protection staff	Investigations initiated because of epidemiological links	Total	Initial	23	0	(0–1)	0.5	(0.8)	
			Follow-up	21	1	(0–2)	2.8	(6.1)	0.04
		Low	Initial	15	0	(0–1)	0.3	(0.62)	
			Follow-up	14	0.5	(0–1)	1.5	(2.3)	0.04
		Medium	Initial	3	1	(0–1)	0.7	(0.7)	
			Follow-up	3	1	(1–4)	2.0	(1.7)	0.14
		High	Initial	5	0	(0–1)	0.8	(1.3)	
			Follow-up	4	1.5	(0.5-15)	7.8	(13.5)	0.14

Table 5 Practices: the workload associated with the TB-STS for nurses and health protection staff *(Continued)*

Strain typing used to provide more information in epidemiological investigation	Total	Initial	22	0	(0–1)	0.6	(1)	
		Follow-up	22	1	(0–2)	1.8	(2.5)	0.03
	Low	Initial	14	0	(0–1)	0.4	(0.8)	
		Follow-up	14	0.5	(0–2)	1.4	(2)	0.05
	Medium	Initial	4	0.5	(0–2)	1.0	(1.4)	
		Follow-up	3	1	(0–2)	1.0	(1)	0.50
	High	Initial	4	0.5	(0–2)	1.0	(1.4)	
		Follow-up	5	2	(1–3)	3.2	(4)	0.17
Strain typing influences an epidemiological investigation	Total	Initial	23	0	(0–1)	0.8	(1.1)	
		Follow-up	14	0.5	(0–2)	1.2	(1.6)	0.17
	Low	Initial	14	0	(0–1)	0.4	(0.8)	
		Follow-up	8	0	(0–0.5)	0.6	(1.4)	0.34
	Medium	Initial	4	0.5	(0–2)	1.0	(1.4)	
		Follow-up	2	1	(0–2)	1.0	(1.4)	0.50
	High	Initial	5	1	(1–3)	1.6	(1.3)	
		Follow-up	4	2	(1.5-3.5)	2.5	(1.7)	0.20
Investigation initiated because of strain typing	Total	Initial	23	0	(0–2)	2.2	(6.3)	
		Follow-up	22	0	(0–1)	1.1	(2.3)	0.79
	Low	Initial	14	0	(0–1)	0.4	(0.8)	
		Follow-up	14	0	(0–0)	0.5	(1.3)	0.43
	Medium	Initial	4	0.5	(0–1)	0.5	(0.6)	
		Follow-up	4	0.5	(0–1.5)	0.8	(1)	0.34
	High	Initial	5	4	(3–6)	8.6	(12.2)	
		Follow-up	4	1	(1–5.5)	3.3	(4.5)	0.78
Epidemiological links identified in strain typing cluster	Total	Initial	22	0	(0–0)	0.4	(0.8)	
		Follow-up	20	0	(0–0)	0.4	(0.8)	0.52
	Low	Initial	13	0	(0–0)	0.2	(0.6)	
		Follow-up	13	0	(0–0)	0.3	(0.9)	0.30
	Medium	Initial	3	0	(0–1)	0.3	(0.6)	
		Follow-up	3	0	(0–0)	0.0	(0)	0.81
	High	Initial	6	0.5	(0–1)	0.8	(1.2)	
		Follow-up	4	0.5	(0–1.5)	0.8	(1)	0.55

Table 5 Practices: the workload associated with the TB-STS for nurses and health protection staff *(Continued)*

% time spent on investigations	Total	Initial	23	1	(0–5)	2.7	(3.2)	
		Follow-up	25	5	(0–5)	7.2	(11.1)	0.04
	Low	Initial	15	0	(0–5)	2.1	(3.1)	
		Follow-up	15	5	(0–12)	8.3	(13.1)	0.04
	Medium	Initial	3	5	(0–5)	3.3	(2.9)	
		Follow-up	4	5	(2.5-5)	3.8	(2.5)	0.42
	High	Initial	5	5	(1–5)	4.4	(3.7)	
		Follow-up	6	3.5	(0–5)	6.2	(9.5)	0.35

[a]Area where respondents worked is defined as low, medium and high TB incidence: <10/100,000, 10-19/100,000, ≥20/100,000 population, respectively.
[b]n is number of people who answered the question.
[c]Paired t-test comparing initial and follow-up responses.

strain typing informs contact tracing activities [6,8]. However, in this study we found the proportion of time health protection staff spent on cluster investigations increased and the number of investigations that used strain typing increased, but there was no increase in the reported number of possible transmission links found between clustered cases. This discordance between the findings on subjective report of utility and the public health outcomes reported could be because the current methods used by public health staff to identify epidemiological links may be inappropriate or ineffective, or there may have been an increase in suspected (but not established) transmission because of the strain typing information. For the TB-STS to have a public health impact and reduce TB transmission, cluster investigations would have to lead to the detection of previously unidentified latently infected and active TB cases.

Limitations

The way the TB-STS was implemented and the survey design have resulted in a number of limitations that provide important lessons for the TB-STS, the evaluation of future services and other complex interventions.

Firstly, the survey was developed after the initiation of the TB-STS so baseline information could not be collected and we are likely to have underestimated the difference between the surveys. However, the initial survey was conducted before the roll-out of any training for the TB-STS and prior to the employment of all national staff to coordinate cluster investigations. An alternative study design, which would have had a control group, would have been possible if the TB-STS was rolled out in a step-wise process across the country, rather than nationally.

Secondly, the target population for the survey was all public health staff, physicians and nurses working in TB control in England. It was not possible to enumerate the sampling frame because no formal or informal register of clinical and health protection staff working in TB could be identified. As a result, we could not calculate a response rate.

Finally, the 50% retention rate between the surveys is quite low and we may have lost the opinions and experiences of a particular group of people. However, non-responders to the follow-up survey did not differ significantly to those that responded to both surveys based profession or burden of TB in their geographical area. Because the study was conducted as part of a programmatic service implementation, results must be interpreted accordingly.

Recommendations

The findings of this survey inform the development of the TB-STS and the design of future evaluations. Despite a significant increase in the number of health protection staff who had received training, there remained some that had not received any, suggesting the need for an ongoing training programme that also takes into account turnover of staff. Self-reported knowledge of how to use the strain typing information was lower for nurses compared with physicians and health protection staff, possibly representing a gap in the training strategy, which did not include nurses. The finding that physicians had the highest self-reported knowledge across the two surveys, even though they were not included in the training strategy, might be because they have had access to information on strain typing from other sources and, relative to nurses and public health staff, might self-rate their knowledge higher.

The perception of usefulness did not change over time as most people found strain typing to be useful in both surveys. This suggests that any changes in practice are due to increasing knowledge and access to strain typing, rather than attitudes towards strain typing. Therefore, to improve use and impact of the TB-STS, there should be a focus on improving training and making strain typing data easily accessible so that it can become better integrated into the TB service.

The findings of this survey argue for the continuation of the TB-STS. A majority of people reported the TB-STS to be useful and health protection staff reported an increase in the number of investigations for which strain typing was used to provide more information, although there was no increase in the number of investigations that were influenced by strain typing. This discordance between the findings on subjective report of utility and the investigation outcomes reported may signify the high value placed on information.

When implementing a public health intervention and planning an evaluation it is essential to have a well-defined sampling frame and a baseline that can be measured before the start of the service implementation. Where possible, the evaluation of a service should start prior to its implementation in order to capture the baseline and to design the evaluation based on the planned service implementation. This survey is an example of where this was not possible and highlights the importance of acknowledging the context in which the service was implemented, both for assessing its success and understanding the limitations of the evaluation design.

The variation in knowledge, attitudes and practices across the professions illustrates the importance of including all the service stakeholders in the evaluation. For example, in the TB-STS, nurse respondents reported lower knowledge, suggesting that they could benefit from being included in the training strategy.

This survey is the first component of the evaluation of the TB-STS. To better understand the public health utility

and evaluate the impact of such a service, a comprehensive mixed-methods evaluation is underway [12]. This includes modelling of the effectiveness and cost-effectiveness and qualitative studies.

Conclusions

Evaluating a complex public health intervention requires a pragmatic approach, taking into account how the service has been implemented. In these initial and follow-up surveys, public health staff, physicians and TB nurses found the TB-STS useful and increased the amount they used it in the first two years of the service, arguing for the continuation of the service. Despite this, the impact of the TB-STS on cluster investigations remained unclear. We recommend continuing the service but with ongoing and more thorough training of service users and focussing on improving knowledge and making data more accessible. Future evaluations of complex interventions should be initiated prior to the implementation of the service, and would benefit from an enumerable sampling frame and a measurable baseline.

Additional file

Additional file 1: Survey Questions.

Abbreviations

HPU: Health Protection Unit; MIRU-VNTR: Mycobacterial interspersed repetitive units-variable number tandem repeats; NHS: National Health Service; PHE: Public Health England; TB: Tuberculosis; TB-STS: Tuberculosis strain typing service.

Competing interests

JM, IA, HM, ML and EV have been employed by Public Health England in the last 5 years. Public Health England played no role in the study design, data interpretation, data analysis, writing of the report or the decision to submit for publication. There are no other potential conflicts of interest.

Authors' contributions

All authors contributed to the original conception, design and editing of the survey, and were involved in the interpretation of the survey results. JM drafted and conducted the survey, and analysed the data. IA and PS contributed to the survey analysis. The paper was drafted by JM, with further comments from all other authors. All authors have given final approval of the manuscript. The authors were all part of the TB-STS Evaluation Group.

Acknowledgements

The authors would like to acknowledge Public Health England's TB Strain Typing Project Board for their support for this evaluation and all those within PHE and the NHS who cascaded and responded to the survey. JM is funded by University College London and Public Health England Impact Studentship. IA is funded by an NIHR Senior Research Fellowship.

Author details

[1]Department of Infection and Population Health, University College London, London, UK. [2]Centre for Infectious Disease Surveillance and Control, Public Health England, London, UK. [3]George Elliot NHS Trust, Warwickshire, UK. [4]Field Epidemiology Services Public Health England, London, UK. [5]European Programme for Intervention Epidemiology, European Centre for Disease Control, Stockholm, Sweden. [6]Department of Infection and Tropical Medicine, Birmingham Heartlands Hospital, Heart of England NHS Foundation Trust, Birmingham, UK. [7]South Midlands and Hertfordshire Public Health England Centre, Hertfordshire, UK. [8]Health Economics Research Group, Brunel University, London, UK. [9]Division of Global Health Equity, Brigham and Women's Hospital, Boston, USA. [10]Department of Epidemiology, Harvard School of Public Health, Boston, Massachusetts, USA. [11]Department of Infectious Diseases, Public Health Service (GGD) Amsterdam, Amsterdam, The Netherlands. [12]Department of Clinical Epidemiology, Academic Medical Centre, University of Amsterdam, Amsterdam, The Netherlands. [13]Faculty of Epidemiology and Population Health, London School of Hygiene & Tropical Medicine, London, UK. [14]Centre for Clinical Microbiology, Research Department of Infection, University College London, London, UK.

Received: 3 September 2014 Accepted: 22 September 2014
Published: 2 October 2014

References

1. Campbell M, Fitzpatrick R, Haines A, Kinmonth AL, Sandercock P, Spiegelhalter D, Tyrer P: **Framework for design and evaluation of complex interventions to improve health.** *BMJ* 2000, **321:**694–696.
2. Petticrew M: **When are complex interventions "complex"? When are simple interventions "simple"?** *Eur J Public Health* 2011, **21:**397–398.
3. Ruddy MC, Davies AP, Yates MD, Yates S, Balasegaram S, Drabu Y, Patel B, Lozewicz S, Sen S, Bahl M, James E, Lipman M, Duckworth G, Watson JM, Piper M, Drobniewski FA, Maguire H: **Outbreak of isoniazid resistant tuberculosis in north London.** *Thorax* 2004, **59:**279–285.
4. Malakmadze N, González IM, Oemig T, Isiadinso I, Rembert D, McCauley MM, Wand P, Diem L, Cowan L, Palumbo GJ, Fraser M, Ijaz K: **Unsuspected Recent Transmission of Tuberculosis among High-Risk Groups: Implications of Universal Tuberculosis Genotyping in Its Detection.** *Clin Infect Dis* 2005, **40:**366–373.
5. Lambregts-van Weezenbeek CSB, Sebek MMGG, van Gerven PJHJ, de Vries G, Verver S, Kalisvaart NA, van Soolingen D: **Tuberculosis contact investigation and DNA fingerprint surveillance in The Netherlands: 6 years' experience with nation-wide cluster feedback and cluster monitoring.** *Int J Tuberc Lung Dis* 2003, **7**(12 Suppl 3):S463–S470.
6. McNabb SJN, Kammerer JS, Hickey AC, Braden CR, Shang N, Rosenblum LS, Navin TR: **Added epidemiologic value to tuberculosis prevention and control of the investigation of clustered genotypes of Mycobacterium tuberculosis isolates.** *Am J Epidemiol* 2004, **160:**589–597.
7. De Vries G, van Hest RAH, Richardus JH: **Impact of mobile radiographic screening on tuberculosis among drug users and homeless persons.** *Am J Respir Crit Care Med* 2007, **176:**201–207.
8. Clark CM, Driver CR, Munsiff SS, Driscoll JR, Kreiswirth BN, Zhao B, Ebrahimzadeh A, Salfinger M, Piatek AS, Abdelwahab J: **Universal genotyping in tuberculosis control program, New York City, 2001–2003.** *Emerg Infect Dis* 2006, **12:**719–724.
9. Ruddy M, McHugh TD, Dale JW, Banerjee D, Maguire H, Wilson P, Drobniewski F, Butcher P, Gillespie SH: **Estimation of the Rate of Unrecognized Cross-Contamination with Mycobacterium tuberculosis in London Microbiology Laboratories.** *J Clin Microbiol* 2002, **40:**4100–4104.
10. Jasmer RM, Roemer M, Hamilton J, Bunter J, Braden CR, Shinnick TM, Desmond EP: **A Prospective, Multicenter Study of Laboratory Cross-Contamination of Mycobacterium tuberculosis Cultures.** *Emerg Infect Dis* 2002, **8:**1260–1263.
11. Public Health England: **National Tuberculosis Strain Typing Service website.** http://webarchive.nationalarchives.gov.uk/20140629102627/http://www.hpa.org.uk/Topics/InfectiousDiseases/InfectionsAZ/Tuberculosis/NationalTuberculosisStrainTypingService/.
12. Mears J, Vynnycky E, Lord J, Borgdorff M, Cohen T, Abubakar I, Sonnenberg P: **Evaluation of the Tuberculosis Strain Typing Service (TB-STS) in England.** *Lancet* 2013, **382**(Supplement 3):S73 [Public Health Science].
13. Brown T, Evans JT, Sails AD TBDB, Anderson LF, Stone M, Jebbari H, Nikolayevsky V: *HPA Mycobacterium Tuberculosis Strain Typing: a Guide to Data Production and Distribution.* 2012.
14. Public Health England: *TB Strain Typing and Cluster Investigation Handbook.* 3rd edition. 2014.
15. Supply P, Allix C, Lesjean S, Cardoso-Oelemann M, Rüsch-Gerdes S, Willery E, Savine E, de Haas P, van Deutekom H, Roring S, Bifani P, Kurepina N, Kreiswirth B, Sola C, Rastogi N, Vatin V, Gutierrez MC, Fauville M, Niemann S, Skuce R, Kremer K, Locht C, van Soolingen D: **Proposal for standardization of optimized mycobacterial interspersed repetitive unit-variable-number tandem repeat typing of Mycobacterium tuberculosis.** *J Clin Microbiol* 2006, **44:**4498–4510.

16. Van Buynder P: **Enhanced surveillance of tuberculosis in England and Wales: circling the wagons?** *Commun Dis Public Health* 1998, **1:**219–220.
17. Bauer J, Kok-Jensen A, Faurschou P, Thuesen J, Taudorf E, Andersen AB: **A prospective evaluation of the clinical value of nation-wide DNA fingerprinting of tuberculosis isolates in Denmark.** *Int J Tuberc Lung Dis* 2000, **4:**295–299.
18. Bidovec-Stojkovic U, Zolnir-Dovc M, Supply P: **One year nationwide evaluation of 24-locus MIRU-VNTR genotyping on Slovenian Mycobacterium tuberculosis isolates.** *Respir Med* 2011, **105**(Suppl 1):S67–S73.
19. Van Soolingen D, Borgdorff MW, de Haas PE, Sebek MM, Veen J, Dessens M, Kremer K, van Embden JD: **Molecular epidemiology of tuberculosis in the Netherlands: a nationwide study from 1993 through 1997.** *J Infect Dis* 1999, **180:**726–736.
20. Centers for Disease Control and Prevention: **New CDC Program for Rapid Genotyping of Mycobacterium Tuberculosis Isolates.** *JAMA* 2005, **293:**2086.

doi:10.1186/1471-2458-14-1023
Cite this article as: Mears *et al.*: Prospective evaluation of a complex public health intervention: lessons from an initial and follow-up cross-sectional survey of the tuberculosis strain typing service in England. *BMC Public Health* 2014 **14**:1023.

Evaluation of the Tuberculosis Strain Typing Service in England

J Mears[1,2], E Vynnycky[1,3], J Lord[4], M Borgdorff[5], T Cohen[6], I Abubakar[1,2], P Sonnenberg[4]
on behalf of the National TB strain-typing evaluation group

1. Public Health England, UK; 2. University College London, UK 3. London School of Hygiene & Tropical Medicine, UK;
4. Brunel University, UK ; 5. Public Health Service (GGD), The Netherlands; 6. Harvard School of Public Health, USA

Background

- The national TB Strain-Typing Service (TB-STS), introduced in England in 2010, aims to type the first *Mycobacterium tuberculosis* isolate from all TB cases using 24-locus MIRU-VNTR.
- A TB-STS with better-targeted cluster investigations could reduce the TB burden by increasing the detection of latent infections and accelerating diagnosis of active cases.
- We describe the first service-wide evaluation of a national TB-STS to estimate effects on TB incidence and cost-effectiveness.

Methods

Data analysis

We estimated the:

- Proportion of *M. tuberculosis* false-positive isolates in 2010-12, which were identified only after strain-typing results were known.
- Effect of cluster investigations on diagnostic delay (symptom onset to diagnosis) for TB cases diagnosed in 2010-2011.
- Contact tracing yield (mean no. of infections or active cases identified per case) in NC London and Leicester in 2011, according to membership of investigated clusters.

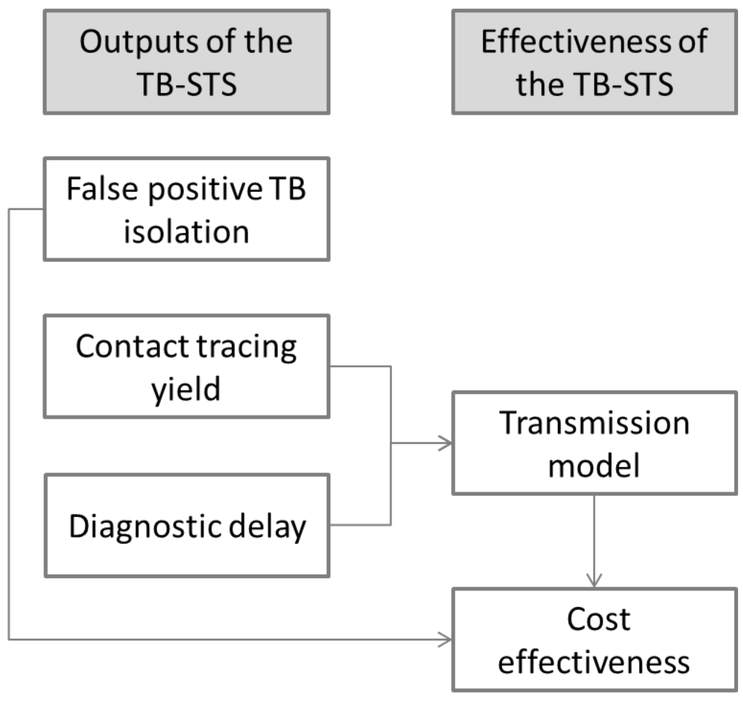


Fig 1: Components of the mixed methods evaluation of the outputs and effectiveness of the TB-STS.

Modelling & cost-effectiveness

- Using an age-structured deterministic transmission model we estimated the effect of the TB-STS in high, medium, low incidence groups, with plausible assumptions for in/out-migration, the proportion of infections detected, treatment delay & preventive treatment.
- Cost-effectiveness used model & service evaluation output, establishment & running costs & QALY (Quality Adjusted Life Year) impacts of infection & disease.

Results

- 17/30 (57%) incidents of false-positive isolation escaped identification before strain-typing results were known.
- Cluster investigations had no significant impact on diagnostic delay or contact tracing yield(Fig 2).
- According to mathematical modelling:
  - The TB-STS had the greatest impact in high incidence groups (Fig. 3)
  - Larger impacts were predicted with plausible ↓ in treatment delay than with ↑ in detection of infections
  - The impact depended on assumed TB prevalence among immigrants, in and out-migration, uptake and retention of PT

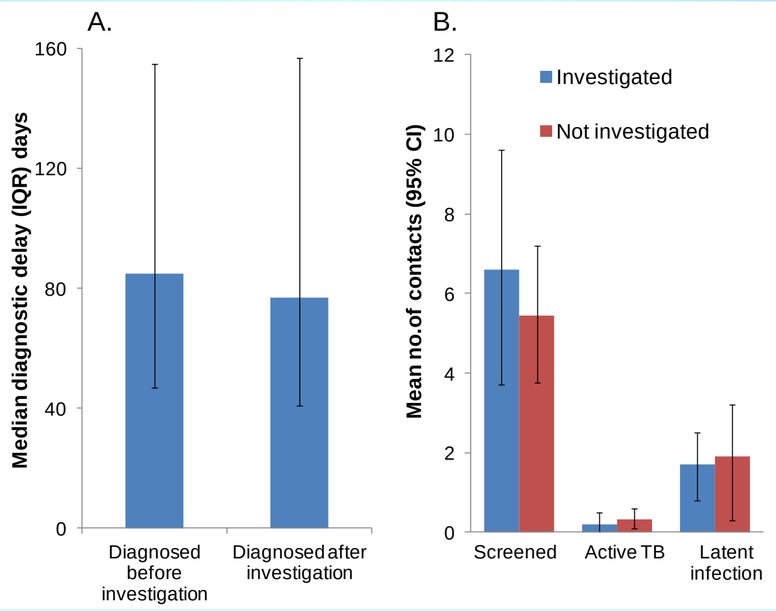


Fig 2: A. Diagnostic delay (& IQ range) and B. Contact tracing yield (& 95% CI) in cases diagnosed before or after launching a cluster investigation

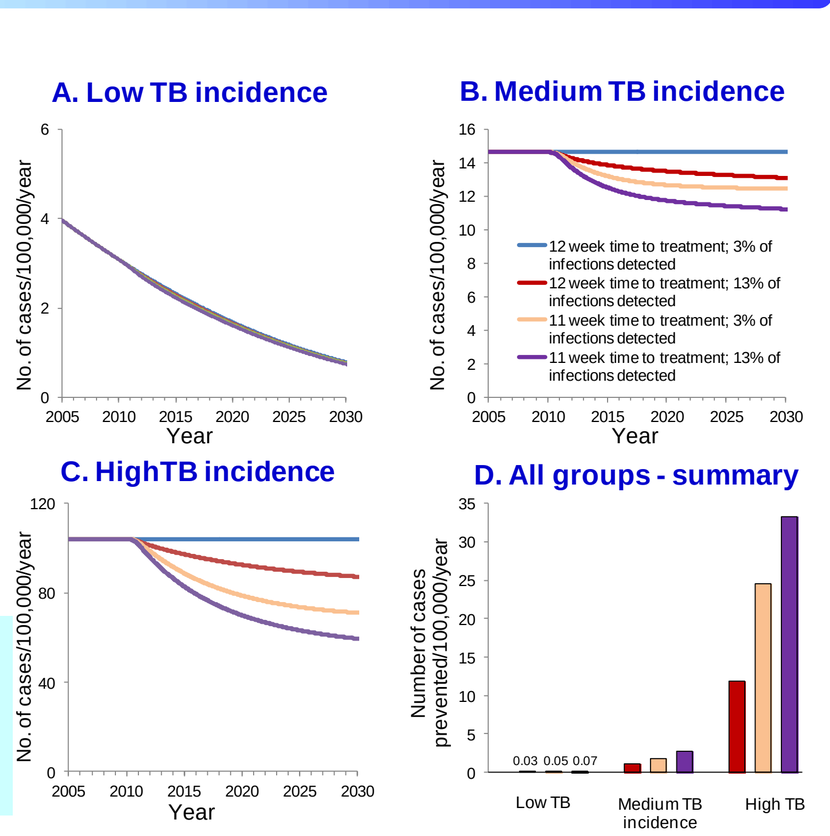


Fig 3: Predicted impact of the TB-STS using mathematical modelling (see 3B for legend)

- For base case assumptions, the TB-STS was not estimated to be cost effective over 20-years (£97,311/QALY).

Conclusions

- Insufficient baseline and post-implementation data limit direct estimation of public health outcomes and interpretation of the model.
- Current models suggest TB-STS is not cost-effective.
- However, the TB-STS provides an invaluable resource for epidemiological and microbiological research & surveillance, informing national & local TB control efforts & targeting interventions.

Health Protection Conference 2011, Warwick

Evaluation of the National TB Strain Typing Service: Results from a Baseline Survey

J Mears, on behalf of the TB Strain Typing Evaluation Group

Abstract

This survey is part of the National Tuberculosis Strain Typing Service (TBSTS) Evaluation. It aimed to establish the utilisation and perception of the STS by those working in TB control, before the TBSTS is fully implemented.

An online self-completion questionnaire was developed by a multi-disciplinary team. The survey-link was emailed to nurses, physicians and health protection (HP) staff in England using the Enhanced Tuberculosis Surveillance user-list, and cascaded through Local and Regional Services. It assessed current awareness and knowledge of TB strain typing (STing), current use and perceived usefulness of the TBSTS.

284 clinic and HP staff within all 26 Health Protection Units (HPUs) in England responded. 85% had heard of the TBSTS. 15% used STing in at least half of cases, 62% for a few cases, and 23% had never used it. 96% found the ST quite or very useful when utilised. 8% of all respondents had received training. STing was most commonly used to identify clusters and/or links between cases. HPUs reported an average of 1.3 (range0-30) cluster investigations initiated during a 3-month period because of the ST, compared to 0.8 (0-4) because of known epidemiological links.

The TBSTS is not utilised by the entire TB control service; however, when used it is perceived to be useful. The number of cluster investigations initiated by HPUs may be small due to a lack of training. They are likely to increase as the TBSTS is implemented and training delivered. A repeat survey in 2011 will measure changes in service use.

Health Protection 2011

Evaluation of the National TB Strain Typing Service: Results from a Baseline Survey

Jessica Mears on behalf on the TB Strain Typing Evaluation Group[1]

Background

This survey is part of the National Tuberculosis Strain Typing Service (TB STS) Evaluation. It aimed to address elements of the processes of the TB STS at baseline and provide inputs for transmission and cost effectiveness models.

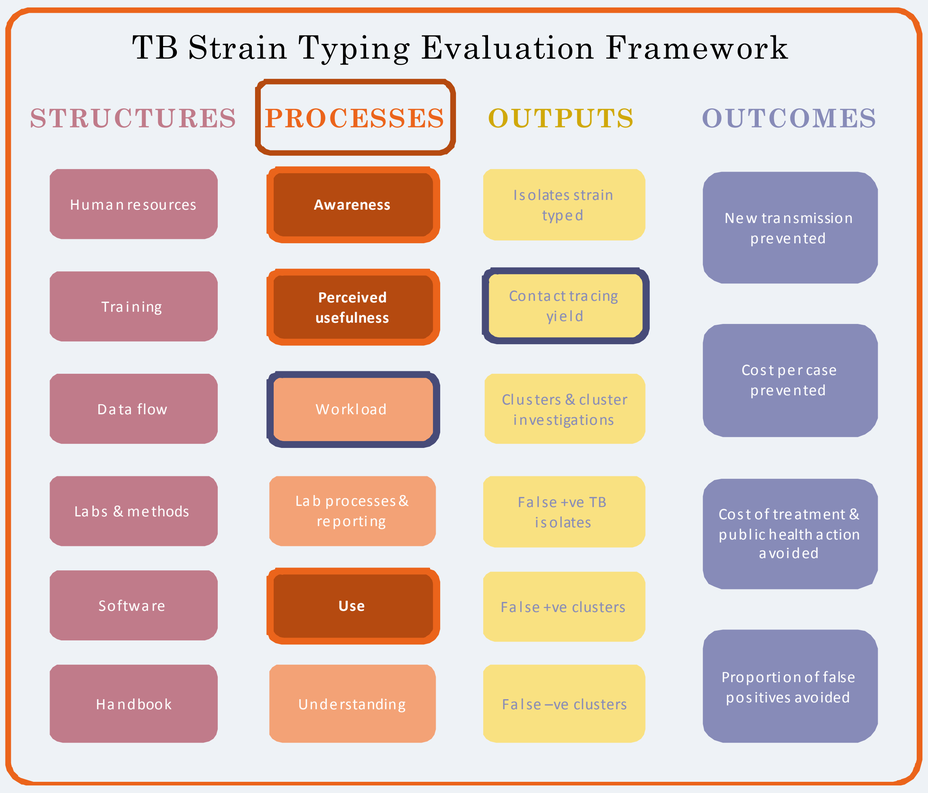


Figure 1: The TB Strain Typing Evaluation is using a structures, processes, outputs and outcomes framework. The components of this framework that are addressed in this survey are outlined in blue and orange; those in orange are presented here.

Objectives

This poster reports on three of the objectives of the first survey, carried out in December 2010. These three objectives of the baseline survey were to establish, within the HPA and TB Services in England and Wales, baseline:

1. Awareness and knowledge of the STS
2. Use of the STS
3. Perceived usefulness of the STS

Methods

Design

- Two cross-sectional surveys: at baseline and once the STS has been implemented
- A link to a web-based self-completion questionnaire was cascaded through local and regional services and through the Enhanced TB Surveillance user list

Target Population

- Nurses and physicians working in clinics in England and Wales
- Health protection staff working in HPUs in England and Wales

Results

- There were 296 respondents
- 42 were excluded
- A total of 254 respondents from within all 26 HPUs were included in the analyses

		N	%
Workplace setting			
Clinical		**198**	**79.2**
	Nurse	139	55.6
	Physician	59	23.6
Health Protection		**52**	**20.8**
	Nurse	18	7.2
	Scientist/admin	7	2.8
	Consultant	27	10.8
missing		4	
TB Incidence/100,000			
Low	<10	115	45.63
Medium	10-20	62	24.60
High	>20	75	29.76
missing		2	

Table 1: Survey respondents' workplace by setting and incidence

Results 1: Awareness of the STS

	N	%
Heard of the Strain Typing Service		
Yes	211	84.4
No	39	15.6
missing	4	
Received any training on using strain typing data		
Yes	21	8.5
No	225	91.5
missing	8	

Table 2: Numbers and proportions of respondents who had heard of the STS and received training

Results 2: Use of the STS

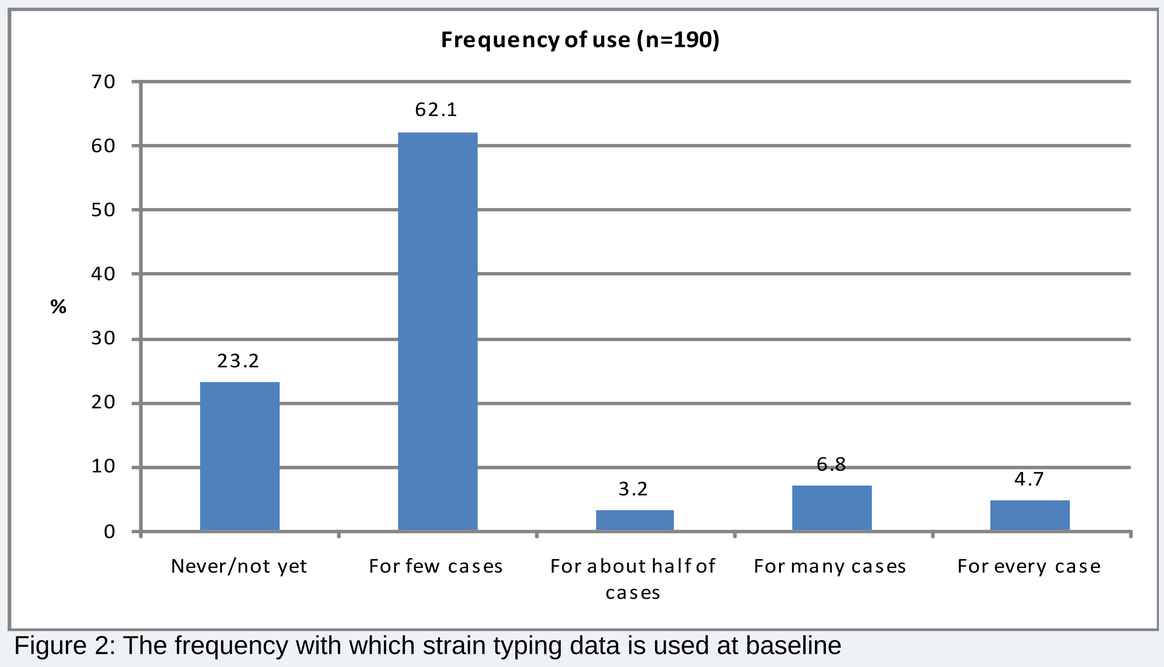


Figure 2: The frequency with which strain typing data is used at baseline

	N	%
Do you access strain typing data? (n=254)		
Yes	197	77.6
No	57	22.4
If yes, how do you use strain typing data?* (n=197)		
Don't know	5	2.5
Identify clusters and links between cases	173	87.8
Disprove clusters and links between cases	105	53.3
Justify extended contact tracing	95	48.2
Justify stopping contact tracing	43	21.8
To provide more information	78	39.6
Other	12	6.1

Table 3: How the strain typing data is used once it is accessed. *Participants could select >1 option

Results 3: Perceived usefulness of the STS

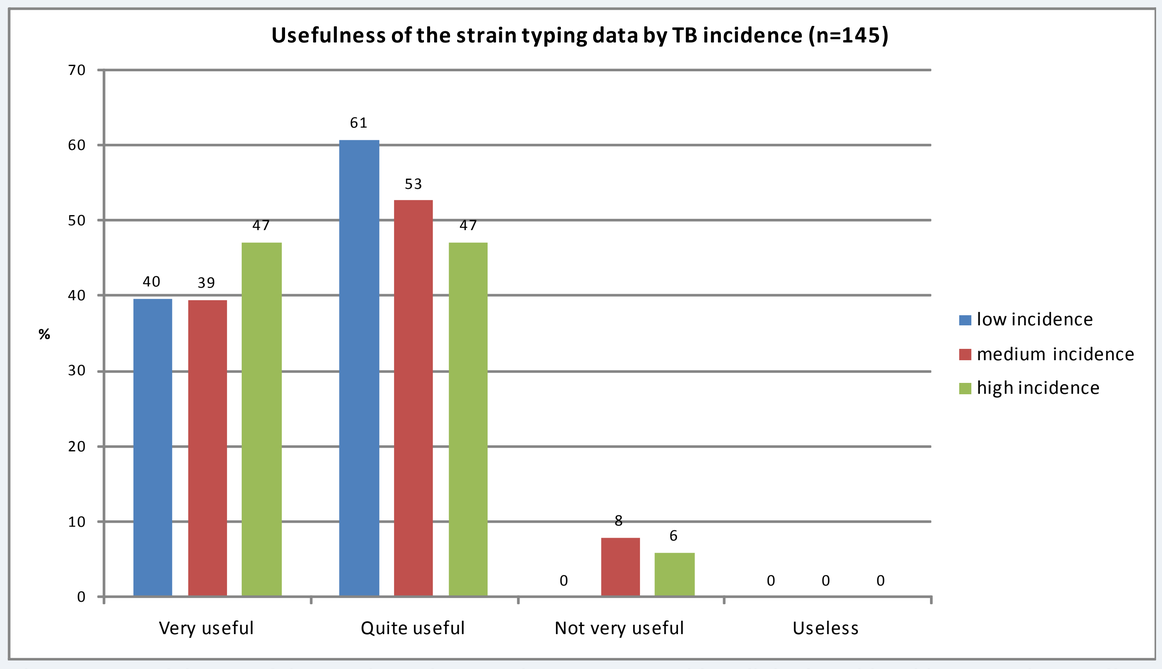


Figure 3: Usefulness of the strain typing service as reported by clinical and health protection staff working in areas of low (n=56), medium (n=38) and high incidence (n=51)

Limitations

- This was a convenience sample as a sampling frame was not available. However, there were a large number of responses from across the target area so the survey can be used to measure changes over time
- We aimed to capture the baseline situation, before the implementation of the STS. However, because there was a large variation in ad hoc use of strain typing across the country (including universal strain typing in the West Midlands) and the STS was implemented at different rates around the country, the definition of 'baseline' varied across geographical area

Conclusions

- The survey was answered by 257 clinical and health protection professionals working in areas covered by all 26 HPUs across England and in Wales, but we cannot be sure of the representativeness of the respondents
- Although many respondents had heard of the TB STS, very few had received any training on using the data
- Most respondents had accessed strain typing data but only used it for a few cases; however, when strain typing was used it was perceived to be useful
- A repeat survey in 2012 will measure changes in use of the TB STS

[1]The TB Strain Typing Evaluation Group consists of Ibrahim Abubakar, Martien Borgdorff, Ted Cohen, Debbie Crisp, Chris Griffiths, John Hayward, John Innes, Michelle Kruijshaar, Mike Lilley, Jo Lord, Helen Maguire, Tim McHugh, Pam Sonnenberg (Chair) and Emilia Vynnycky.