Modelling Tuberculosis in Ethiopia: Spatiotemporal Transmission Dynamics and Effects of Public Health Interventions

Debebe Shaweno Adewo

ORCID ID - http://orcid.org/0000-0001-9596-5443

Submitted in total fulfilment of the requirements of the degree of

Doctor of Philosophy

September 2019

Department of Medicine

Faculty of Medicine, Dentistry and Health Sciences

The University of Melbourne
Abstract

Tuberculosis (TB) is now the world’s leading infectious killer with an estimated 10 million cases and 1.6 million deaths in 2016. A small number of countries bear the majority of the burden of disease, with two-thirds of cases occurring in only seven countries. TB transmission occurs in both households and the local community, leading to focal disease hotspots which perpetuate TB spread within and across community groups. Integrating spatial analysis with mathematical transmission dynamic models can help in evaluating the role of these hotspots in the spread of TB and understanding the potential impact of geographically targeted interventions.

The first part of this thesis evaluates whether TB exhibits spatial heterogeneity in rural and remote regions of Ethiopia using data from all TB patients treated in a remote administrative region of the country. This study demonstrated considerable spatial heterogeneity in TB distribution in this resource-limited setting. However, most of these heterogeneities were accounted for by health facility availability, implying differential case detection between areas with better and poorer access to health care. Thus, this Chapter cautions that spatial analysis of TB and the identification of geographical hotspots using programmatic data alone can be misleading, as it may be strongly influenced by undetected cases, which is in turn dependent on local programmatic performance.

Building on the findings outlined above, Chapter three presents a systematic review of methods used in published spatial analyses of TB. From this review, this Chapter elucidates limitations in the current approaches to spatial analysis of TB. Of particular importance is the consistent failure to account for unreported or undetected cases, despite notification data being used in 95 percent of the reviewed studies. The Chapter also describes methodological flaws to many of the studies, in particular the use of conventional regression analysis to draw spatial conclusions. In addition, most spatial analyses of TB distribution used residential information to define the location of patients, which potentially understates the importance of other community settings, despite more than 80% of all transmission events occurring outside households.

The study in Chapter 4 proposes a method to address the limitations outlined in the previous chapters – particularly the lack of methods to account for undetected cases. The model estimates both incidence and case detection rates simultaneously across space and time, providing a useful platform for regularly tracking spatial patterns and temporal trends. In
addition, this technique is general and can be applied to any disease in any setting. Applied to the Ethiopian setting, this model identifies previously unrecognised areas of high TB burden in locations with no available health care facilities.

With the aim of quantifying the role of TB hotspots in community transmission as well as evaluating the potential impact of targeting spatial hotspots, Chapter 5 utilises incidence data generated by the novel method described above to identify spatial TB hotspots. At this point, the thesis constructs spatially structured mathematical models and quantifies the extent to which these hotspots account for the spatial spread of TB. Findings from this work suggest that TB transmission in the same study region in rural Ethiopia is localised and the role of spatial hotspots in the spatial spread is limited, although their impact is considerable in adjacent locations due to very high relative incidence in the hotspot compared to the other regions. Finally, Chapter 6 uses the same model introduced in Chapter 5 to evaluate the impact of various TB intervention strategies before concluding the thesis.

Overall, this thesis advances current approaches to spatial analysis and provides a means to account for the problem of undetected cases. It also provides a platform to estimate both incidence and case detection rate simultaneously, and hence could provide as an alternative approach to the spatial interpretation of TB epidemiology. This is of particular importance to high endemic settings, where considerable number of TB cases are missed, and case notification is biased to areas with better access to health care. Importantly, the study also concludes that the impact of spatial TB hotspots on the spatial spread of disease in remote regions of Ethiopia is limited and transmission is predominantly locally driven. Hence, interventions strategies that are spatially targeted may not achieve anticipated outcomes, although the overall effect of these interventions remains considerable due to extremely high incidence in the hotspot regions.
Declaration

This is to certify that:

▪ the thesis comprises only my original work toward the Doctor of Philosophy except where indicated in the preface;
▪ due acknowledgement has been made in the text to all other materials used;
▪ the thesis is fewer than 100,000 words in length, exclusive of tables, figures, maps, references and appendices; and
▪ the thesis complies with the stipulations set out for the degree of Doctor of Philosophy by the University of Melbourne.

Signed,

Debebe Shaweno Adewo  September 5, 2019
Preface

Chapters of this thesis from 2 to 5 consist of reprints of published articles in which I was the primary author responsible primarily for the planning, execution and manuscript preparation for publication. Work presented in Chapter 6 is in preparation for submission to a journal. Below, I provide statements regarding the contribution of authors as reported in the published articles.

Chapter 2 - DS (Debebe Shaweno) conceived the study, performed the analysis, drafted the manuscript and revised it based on feedback from co-authors and reviewers. TS (Tamrat Shaweno) collected data; EM (Emma McBryde), JD (Justin Denholm) and JT (James Trauer) helped to design the conceptual framework of the study and all provided inputs into revisions.

Chapter 3- DS (Debebe Shaweno) and EM (Emma McBryde) conceived the study, JD (Justin Denholm) and JT (James Trauer) refined it further. DS (Debebe Shaweno) developed data extraction checklist, and DS (Debebe Shaweno), MK (Malancha Karmakar) and KA (Kefyalew Addis) extracted the data. DS (Debebe Shaweno) drafted the manuscript, and all authors provided inputs into revisions and approved the final draft for submission.

Chapter 4 - DS (Debebe Shaweno) and EM (Emma McBryde) wrote the code in WinBUGS, DS (Debebe Shaweno) performed model simulation and drafted the initial study concept and EM (Emma McBryde, JD (Justin Denholm) & JT (James Trauer) refined this further. DS (Debebe Shaweno) wrote the initial manuscript draft, and all authors provided input into revisions and approved the final draft and submission for publication.

Chapter 5 - D.S (Debebe Shaweno) wrote the initial code and performed model simulation, J.M.T (James M Trauer) and E.S.B(Emma S McBryde) added additional lines of code, D.S (Debebe Shaweno) drafted the initial study concept and E.S.M (Emma S McBryde, J.T.D (Justin T Denholm) & J.M.T (James M Trauer) refined this further. D.S (Debebe Shaweno) wrote the initial draft of the manuscript, and all authors provided input into revisions and approved the final draft and submission for publication.

Chapter 6 – Debebe Shaweno wrote the model code and performed model simulation, Emma McBryde contributed additional lines of code. Emma McBryde, James Trauer and Justin Denholm provided advice in developing economic models, Tan Doan contributed to sensitivity analysis and cost-coverage analysis.
Acknowledgements

A number of individuals and institutions have supported me in producing this thesis. Massive thanks go to my doctoral supervisors: Prof Emma McBryde, A/Prof Justin Denholm and Dr James Trauer. It was a real privilege and honour for me to share your wealth of expertise, insights and I thank you for allowing me to study in your very collegial environment.

Support from my principal supervisor, Prof Emma McBryde started before I joined the University of Melbourne during the time of application. She lodged my PhD application in person to the University of Melbourne admission office, arranged accommodation for my early days and continued to provide technical assistance throughout the course of my PhD.

I thank A/Prof Justin Denholm and Dr James Trauer for being ‘advanced colleagues’ on the social side of life.

My ability to learn infectious disease modelling was greatly enhanced by the generous provision of transmission dynamic model codes in MATLAB by Dr James Trauer and he was just a click away in responding to emails and reviewing draft manuscripts. Thank you for all.

Dr Tan Doan and Mr Romain Ragonnet have been so helpful during the course of my PhD and the guidance from Dr Tan Doan on coding Latin hypercube sampling (LHS) was extraordinary and he was always available at my convenience.

I owe my thanks to my brother Tamrat Shaweno, and colleagues Ellesa Moges and Sewunet Dasho for facilitating the collection of data from all health facilities at no cost. Tamrat not only coordinated the collection of data but also entered into EpiData software. The contribution of my uncle, Wedo Atto was so incredible. He paid repeated visits to higher officials of the Ethiopian Central Statistical Agency to make sure that I have got the shape files (maps) for the study Zone. Without your immense support, this PhD project would not have been a reality.

I am very thankful to my wife, Mrs Emebet Tekletsadik, for her exceptional devotion and commitment in shouldering family responsibilities during the period of my PhD study, while also doing undertaking her own PhD studies. My daughter, Soliana Shaweno has been a source of energy during the course of this PhD. Thank you Sola.

This doctoral research is funded by the University of Melbourne through a Melbourne International Research Scholarship and Melbourne International Fee Remission Scholarship. I
would also like to acknowledge the Centre for Research Excellence in Policy Relevant Infectious diseases Simulation and Mathematical Modelling (PRISM²) for funding my attendance at conferences.
Publications in Peer-Reviewed Journals

Primary Publications


Secondary Publications


Articles Submitted for Publication

Shaweno D, Trauer JM, Denholm JT, McBryde ES. Geospatial clustering and modelling provide policy guidance to distribute funding for active TB case finding in Ethiopia
Presentations in International Conferences

Oral presentations


Poster presentations

Shaweno D, Trauer JM, Denholm JT, McBryde ES. The role of geospatial hotspots in the spatial spread of tuberculosis: A mathematical model. *PRISM² International Conference, August 20-24, Palm Cove, Australia*
# Table of Contents

Abstract ................................................................................................................................. ii

Declaration ............................................................................................................................ iv

Preface ................................................................................................................................... v

Acknowledgements .............................................................................................................. vi

Publications in Peer-Reviewed Journals ........................................................................... viii
  Primary Publications ........................................................................................................ viii
  Secondary Publications .................................................................................................... viii

Articles Submitted for Publication ...................................................................................... ix

Presentations in International Conferences ......................................................................... x
  Oral presentations ........................................................................................................... x
  Poster presentations ....................................................................................................... x

Table of Contents ................................................................................................................ xi

List of Tables ....................................................................................................................... xv

List of Figures ...................................................................................................................... xvi

List of Supplementary Figures .......................................................................................... xvi

Abbreviations ...................................................................................................................... xvii

Chapter 1 ............................................................................................................................... 1

## 1 Literature Review ........................................................................................................... 1

  1.1 Introduction to the Epidemiology of Tuberculosis ....................................................... 2
    1.1.1 Global Epidemiology ............................................................................................ 2
    1.1.2 Current State of Molecular TB Diagnostics ....................................................... 5
    1.1.3 Methods to estimate tuberculosis incidence ....................................................... 6
    1.1.4 Epidemiology of tuberculosis in Ethiopia ........................................................... 8

  1.2 Introduction to spatiotemporal analysis of tuberculosis ........................................... 10
    1.2.1 Data visualisation ............................................................................................... 10
    1.2.2 Overview of spatial cluster identification methods ........................................ 11
    1.2.3 Spatial statistical modelling .............................................................................. 14
    1.2.4 Spatial analysis of tuberculosis ................................................................. 17
1.3  Introduction to mathematical modelling of tuberculosis......................................................18
1.3.1  Issues with standard TB transmission models ..............................................................20
1.3.2  Approaches to spatially structured transmission dynamic models .................................20
1.3.3  Spatial metapopulation models .......................................................................................21
1.3.4  Spatial metapopulation TB transmission models .............................................................22
1.4  Rationale for this doctoral research ....................................................................................23
1.5  Research aims .....................................................................................................................25
1.6  Thesis scope .......................................................................................................................25
1.7  Overview of methods and materials ..................................................................................27
1.7.1  Study setting and data sources .......................................................................................27
1.7.2  Data collection and management ...................................................................................27
1.7.3  Data analysis ..................................................................................................................28
1.8  Organisation of the thesis ....................................................................................................30

Chapter 2 ..................................................................................................................................32

2  Heterogeneity in the distribution of tuberculosis in Sheka Zone, Ethiopia: drivers and temporal trends ..................................................................................................................32

Chapter 3 ..................................................................................................................................41

3  Methods used in the spatial analysis of tuberculosis epidemiology: A systematic review ............................................................................................................................................41
  3.1  Supplementary Materials .................................................................................................60
         3.1.1  Trends in the spatial analysis of TB ...........................................................................60

Chapter 4 ..................................................................................................................................61

4  A novel Bayesian geospatial method for estimating tuberculosis incidence reveals many missed TB cases in Ethiopia ........................................................................................................61
  4.1  Supplementary Materials .................................................................................................70
         4.1.1  Model code ................................................................................................................70
         4.1.2  Outputs from Candidate Models ..............................................................................72

Chapter 5 ..................................................................................................................................74

5  The role of geospatial TB hotspots in the spatial spread of TB in Ethiopia: a mathematical model ............................................................................................................................74
  5.1  Supplementary Materials .................................................................................................86
         5.1.1  Spatial TB clusters in Sheka Zone, Ethiopia ...............................................................86
         5.1.2  Deviance information criterion ..................................................................................86
Chapter 6  ........................................................................................................................................... 92

6  Impact of geographically targeted interventions for TB control in Ethiopia: A mathematical model................................................................. 92

6.1  Abstract .................................................................................................................................. 93
6.2  Background ............................................................................................................................. 94
6.3  Methods ................................................................................................................................ 95
6.3.1 Overview of past work .......................................................................................................... 95
6.3.2 Intervention strategies ......................................................................................................... 95
6.3.3 Cost-coverage analysis and cost-effectiveness analysis ..................................................... 97
6.3.4 Sensitivity analysis ............................................................................................................... 98
6.4  Results .................................................................................................................................. 99
6.4.1 Impact of intervention strategies on TB epidemiology ....................................................... 99
6.4.2 Sensitivity analysis ............................................................................................................... 100
6.4.3 Cost-coverage curves .......................................................................................................... 101
6.4.4 Cost-effectiveness analysis ................................................................................................. 102
6.4.5 What proportion of TB budget should go to hotspots for a maximum impact? .. 102
6.5  Discussion ............................................................................................................................. 104
6.6  Conclusions ........................................................................................................................... 106
6.7  Authors contributions ........................................................................................................... 106
6.8  Funding.................................................................................................................................. 106
6.9  Competing interests............................................................................................................... 106
6.10 References ............................................................................................................................ 107
6.11 Supplementary materials ....................................................................................................... 110
6.11.1 Cost-coverage curves ........................................................................................................ 110

Chapter 7 ....................................................................................................................................... 113

7  Conclusions and future directions............................................................................................ 113

7.1 Lessons learned ....................................................................................................................... 114
7.1.1 Limitations in approaches to spatial analysis of TB ............................................................ 114
7.1.2 Data driven spatial TB heterogeneity ................................................................................. 115
7.1.3 A novel approach to deal with data driven heterogeneity .............................................. 115

xiii
7.1.4 Importance of mathematical modelling to capture spatial TB spread ...............116
7.2 Future directions and limitations .....................................................................118

8 References ..............................................................................................................121
List of Tables

Table S 1. Incidence and case detection estimates from candidate models .......................... 72
Table S 2. Regression coefficients and predicted case detection rates from the candidate models ........................................................................................................................................................................ 73
Table S 3. Parameter and their ranges for sensitivity analysis ................................................. 110
List of Figures

Figure 1. Spatial regression decision process ................................................................. 16
Figure 2. Proportion of entire population screened and reduction in TB incidence for the entire study region, hotspot regions and the two non-hotspot regions ........................................... 99
Figure 3. Sensitivity of predicted overall TB incidence to variations in selected model parameters .................................................................................................................. 100
Figure 4. Impact of program spending .............................................................................. 101
Figure 5. Impact of geographic funding allocation on population TB incidence under four budget envelope scenarios .............................................................. 103

List of Supplementary Figures

Figure S 1. Trends in the spatial analysis of TB ................................................................. 60
Figure S 2. Spatial distribution of TB in Sheka Zone, 2010-2014 ........................................ 86
Figure S 3. Posterior distribution of fitted parameters under the assumption of mixing with different contact rates in each region ...................................................................... 88
Figure S 4. Trace plots of estimated parameters and the likelihood from the best fitting model ......................................................................................................................... 89
Figure S 5. Posterior distribution of fitted parameters ...................................................... 89
Figure S 6. Posterior distribution of fitted parameters ...................................................... 90
Figure S 7. Correlation between pairs of parameters of model D ..................................... 90
Figure S 8. Posterior distribution of parameters, model D ............................................... 91
Figure S 9. Sensitivity of cost-coverage curves to a unit cost parameter ......................... 111
Figure S10. Sensitivity analysis of intervention impact unit cost per person screened ........ 112
Figure S 11. Sensitivity of predicted TB incidence to changes in selected model parameters ...................................................................................................................... 112
## Abbreviations

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>ART</td>
<td>Antiretroviral Treatment</td>
</tr>
<tr>
<td>ARTI</td>
<td>Annual Risk of TB infection</td>
</tr>
<tr>
<td>BCG</td>
<td>Bacillus Calmette–Guérin</td>
</tr>
<tr>
<td>CAR</td>
<td>Conditional Autoregressive Models</td>
</tr>
<tr>
<td>CHWs</td>
<td>Community Health Workers</td>
</tr>
<tr>
<td>DOTS</td>
<td>Directly Observed Treatment Short-course</td>
</tr>
<tr>
<td>GIS</td>
<td>Geographic Information System</td>
</tr>
<tr>
<td>GWR</td>
<td>Geographically Weighted Regression</td>
</tr>
<tr>
<td>HBCs</td>
<td>High burden countries</td>
</tr>
<tr>
<td>HEP</td>
<td>Health Extension Program</td>
</tr>
<tr>
<td>HIV</td>
<td>Human Immunodeficiency Virus</td>
</tr>
<tr>
<td>IHME</td>
<td>Institute of Health Metrics and Evaluation</td>
</tr>
<tr>
<td>k-NN</td>
<td>k-nearest neighbourhood</td>
</tr>
<tr>
<td>LISA</td>
<td>Local Indicators of Spatial Autocorrelation</td>
</tr>
<tr>
<td>LTBI</td>
<td>Latent TB Infection</td>
</tr>
<tr>
<td>MAUP</td>
<td>Modifiable Areal Unit Problem</td>
</tr>
<tr>
<td>MCMC</td>
<td>Markov-Chain Monte Carlo</td>
</tr>
<tr>
<td>MDR</td>
<td>Multidrug-resistant</td>
</tr>
<tr>
<td>Mtb</td>
<td>Mycobacterium tuberculosis</td>
</tr>
<tr>
<td>NNI</td>
<td>Nearest Neighbourhood Index</td>
</tr>
<tr>
<td>PLHIV</td>
<td>People living with HIV</td>
</tr>
<tr>
<td>SMR</td>
<td>Standardised Morbidity/Mortality Ratio</td>
</tr>
<tr>
<td>SNNPR</td>
<td>Southern Nations, Nationalities and Peoples Region</td>
</tr>
<tr>
<td>TB</td>
<td>Tuberculosis</td>
</tr>
<tr>
<td>WGS</td>
<td>Whole Genome Sequencing</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organization</td>
</tr>
</tbody>
</table>
Chapter 1

1 Literature Review

CHAPTER OVERVIEW

Understanding the current status of TB epidemiology and approaches to spatial and mathematical TB models are key to designing appropriate interventions. In this regard, before describing the main themes of the doctoral research, this chapter sets the scene by outlining four major themes. First, it presents the current epidemiological situation of TB, along with the current research needs - including the paucity of suitable methods to estimate TB incidence. Second, it presents the main spatial analysis methods, their limitations and how they can potentially be used to optimise TB interventions. Third, it describes the role of mathematical models and particularly of spatially structured mathematical models in describing spatial TB dynamics and how they can be used to optimise spatially targeted interventions. Fourth, the thesis summarises the background to this doctoral research, research questions, and the overall aim of the doctoral research. Finally, the chapter provides a brief description of methods used – particularly the data sources and data management approaches.
1.1 Introduction to the Epidemiology of Tuberculosis

1.1.1 Global Epidemiology

Tuberculosis (TB) persists as a global public health problem of serious magnitude and severity, despite having been curable since combination chemotherapy was first established in the 1970s [1]. Recognising that the global TB epidemic was out of control in many parts of the world, the World Health Organization (WHO) declared TB a global emergency in 1993 and introduced programmatic TB management protocols known as directly observed treatment short-course (DOTS) [2, 3]. Today, after more than two decades of DOTS implementation, an estimated 23 percent of the world’s population remains latently infected, and 10 million people developed some form of the disease in 2017 with 1.6 million attributable deaths [4]. Although the estimated number of incident TB episodes each year is slowly decreasing, TB remains the number one cause of death from a single infectious agent [1, 5].

Tuberculosis occurs throughout the world, but low and middle-income countries are responsible for the greatest share of the global TB burden [6], driven mainly by sustained community transmission [5, 7]. In order to direct the limited resources available to countries with the highest TB burden and maximise both national and global impact, the WHO identified 22 TB high burden countries (HBCs) in 1998 [8], which were subsequently revised to 30 HBCs in 2015 [1]. These countries are together responsible for about 90 percent of the global TB burden and the subsequent strengthening of DOTS has been associated with the current reduction in global TB incidence. However, the rate of reduction has remained very slow, at around 1.5%-2% annually [1, 9, 10] and is even slower in African region than in other regions [10]. Compared to low-income countries, high-income countries have seen a more rapid decline in incidence, which is linked to high health expenditure, implementation of a comprehensive package of effective interventions, low rates of HIV and higher human development index [11, 12].

Although country level estimates in high-incidence settings show an overall declining trend, different local micro-epidemics exist, where the tuberculosis epidemic continues to grow and the risk of TB transmission and death are unevenly distributed across population[13, 14]. In these settings, even spatially adjacent neighbourhoods could have substantially different TB burden depending on the social, economic and environmental backgrounds of population that enhance TB transmission, which may include living in areas with overcrowded population and
To interrupt transmission in these high-risk population groups, scaling-up active-case detection strategies coupled with immediate treatment are required. However, the implementation of active case-detection strategies has remained inefficient and populations at increased TB risk have not been systematically targeted.

The health systems in high TB burden countries have relied on passive case detection and thus more than 50 percent of prevalent TB cases have remained undetected in many settings [17], leading to unnecessary death and to unrecognized transmission of tubercle bacilli [18]. Without early case detection and treatment, majority of these missing cases will die with up to 70 percent of people with sputum smear-positive pulmonary TB and about 40 percent of people with all clinical forms of TB dying within 10 years of disease onset [5]. Many of the missing cases are among high-risk population groups that have difficulty accessing public health services such as people living with HIV, migrants, refugees, and people living in geographic areas with high TB prevalence [19].

On the other hand, these untreated missing individuals with active TB can infect 10–15 people over the course of a year before they die from TB [20]. Finding these missing cases and breaking the cycle of transmission is a major priority for national TB control programs, although this needs a strong health care system with sensitive and quick diagnostic facilities; and innovative strategies to reach high-risk population.

High risk population groups especially in high burden settings might warrant targeted interventions to efficiently use limited resources, as TB programs in these settings have remained severely unfunded [21]. Moreover, the need to adopt innovative strategies to target high risk population becomes necessary because the tools (vaccines, drugs) and practices we currently use have inadequacies in effectively curbing TB transmission and death. For example, the bacillus Calmette–Guérin vaccine, which was developed more than a century ago, is about 80% effective at protecting children from severe forms of TB, particularly tuberculous meningitis, and miliary TB, but it does not provide significant protection against the more common pulmonary manifestation that is the most transmissible form and is responsible for most deaths [22]. The protection provided by BCG has not been consistent across all forms of TB and age groups [23], its effects appear to wane over time and the live vaccine itself can cause TB disease in children living with HIV [24, 25]. Currently there remains no alternative vaccine that is effective in protecting TB [22], although a number of candidates have entered different phases of clinical trials [1, 26].
Similarly, although drugs that could effectively cure TB have been available since the 1940s to the 1960s [27, 28], lack of combination therapy and of careful patient supervision during the course of treatment has triggered the proliferation of multidrug-resistant TB (MDR-TB) [29, 30]. Today, MDR-TB (TB resistant to isoniazid and rifampicin – the two most important first-line drugs) is a major threat, with over 490,000 new cases occurring each year [1]. Most MDR-TB cases now arise through person-to-person transmission of drug-resistant strains that had originally emerged following suboptimal use of these drugs decades earlier [26]. Drug-resistant tuberculosis is now predicted to be the single largest cause of mortality from antimicrobial resistance by 2050 if urgent action is not taken [31].

Currently, more than 95 percent of people with DS-TB can be successfully treated in 6 months with a four-drug regimen when treatment adherence is optimum [30, 32], although the success rate is usually lower in practice (83 percent in 2016 globally) [1]. To promote adherence, treatment durations substantially shorter than the current 6 months are desirable [26]. However, the efficacy of shorter treatment regimens from recent clinical trials was found to be inferior compared to the standard 6 month treatment approach [33, 34] and regimens that are superior to the those traditionally used are urgently needed[26].

Unfortunately, success rates for MDR-TB remain less than optimum (54 percent) globally [35] for various reasons, including the long treatment duration, high pill burden and significant drug side effects [30]. Currently, a major shift is underway in reducing the treatment duration of MDR-TB from a minimum of 20 months to about 9 months [36], although the current duration is still too high. Furthermore, the side effects from many MDR-TB and XDR-TB drugs are serious and severe [32], such that new drugs and treatment strategies are needed to reduce pill burden and treatment duration [30].

The emergence of the HIV pandemic has also had a major impact on TB incidence rates over the past two decades [37]. This pandemic has led to striking increases in case load in developing countries, especially in Sub-Saharan Africa in the 1990s [29]. HIV affects TB epidemiology by altering the natural course of infection and increasing the risk of latent TB (LTBI) reactivation [38]. Unique challenges remain in the treatment of TB among individuals coinfected with HIV, including serious drug–drug interactions between TB drugs and antiretroviral drugs [30, 35]. In addition, diagnosis of both active TB and LTBI are more
challenging, with TST having a lower sensitivity in diagnosing LTBI in HIV-infected individuals than in other populations [38].

Because of the inherent limitations in the current tools and approaches in TB control described above, prioritizing population at increased TB risk for diagnosis and treatment with strong treatment adherence support may yield maximum gains.

1.1.2 Current State of Molecular TB Diagnostics

Until recently, TB has had no effective point-of-care test that can quickly and accurately detect mycobacterium tuberculosis (Mtb). Smear microscopy remains the standard method of testing even today, although it is neither sensitive nor able to detect drug resistance [39]. Its results depend heavily on the experience of the technician performing the test and misses considerable numbers of cases in people infected with HIV, leading to reliance on clinical assessments and other less specific diagnostics such as chest X-ray [40]. More recently, rapid developments have been made in molecular TB diagnostics, such that the diagnosis of TB and its drug resistance profile has entered a new era of faster and more cost-effective molecular detection [41]. Endorsed by WHO in 2008, the Xpert MTB/RIF assay can simultaneously detect both M. tuberculosis and determine rifampin-resistance mutations, considerably cutting the time to diagnosis to less than 2 hours [1]. Generally, automated molecular diagnostic tests such as Xpert reduce diagnostic delays without requiring high-level technical skills [39, 42].

Importantly, some molecular methods allow differentiation of M. tuberculosis strains and have greatly increased accuracy, to the point that the movement of specific strains through both time and space can be followed [20]. However, the resolution of standard molecular methods such as MIRU-VNTR is too low to distinguish transmission, as geographical regions may have limited strain diversity [43, 44]. In addition, currently available rapid molecular tests assess only a limited number of TB drug resistance mutations [1] and cannot fully replace phenotypic tests in guiding the treatment of MDR/XDR-TB [26]. Therefore, scaling up the currently limited availability of tests such as whole genome sequencing (WGS) that simultaneously report molecular susceptibility results for multiple drugs and isolates would be particularly useful developments.

Globally, the use of molecular tests is increasing, and many countries are replacing smear microscopy with molecular methods. Although universal drug susceptibility testing is a key component of the End TB strategy, many low-income countries continue to rely on
conventional methods, with about 44 percent of notified TB cases in 2017 not confirmed bacteriologically [4]. Access to appropriate TB diagnostics and drug susceptibility tests enables accurate identification of TB cases and initiation of treatment, thereby reducing mortality and curbing transmission in the community [39]. It is estimated that 36 percent of all TB cases globally in 2017 were missed, partly due to a lack of diagnostic capacity in low income settings[4].

1.1.3 Methods to estimate tuberculosis incidence

In an attempt to rapidly curtail the TB epidemic, the WHO has developed a global strategy that targets a 95% reduction in TB deaths and a 90% reduction in TB incidence by 2035 compared with 2015 levels [1]. To understand the epidemiologic situation and track progress towards TB elimination goals, accurate measurement of epidemiological indicators is crucial. The number of new TB cases per capita each year (incidence rate) has been the central measure to track progress during both the Millennium Development Goal (MDG) and Sustainable Development Goal (SDG) periods [10]. It also serves as the denominator for calculating both case fatality rate and case detection rate.

Direct measurement of TB incidence has been impractical because it requires the follow up of a very large group of disease-free individuals over considerable time periods of years to decades [4]. As a result, national level TB incidence has never been directly measured, and reported rates are invariably estimated [45, 46]. Because TB incidence is estimated but not measured, it is inherently uncertain and consequently all other indicators derived from it (e.g. case detection rate and case fatality rate) also reflect this uncertainty [47]. During the pre-chemotherapy era, one of the most prominent approaches to estimating TB incidence was the Styblo rule, which proposed a fixed mathematical relationship between incidence, prevalence and annual risk of infection, although its assumptions fail to hold since the advent of chemotherapy [48, 49]. Styblo proposed that, in the absence of control, an annual risk of infection (ARTI) of 1000 infections per 100,000 population would correspond with an incidence of new cases of smear-positive TB of approximately 50 per 100,000 population per year. This implies that, in the absence of control measures, an average patient with a new episode of smear-positive TB would generate approximately 20 infections over their infectious lifetime [47]. However, the annual risk for infection measured through tuberculin surveys can be imprecise due to false positives (e.g. from cross reactivity with BCG vaccination) and false negatives (e.g. from lack of Mtb-specific immunity, especially in those with poor immune
status such as those with HIV infection) which in turn can lead to inaccurate measurements of incidence.

WHO has adopted various approaches to estimating incidence based on country-specific TB-related program quality indicators, including completeness of case notification data, the presence of prevalence surveys and the presence of vital registration systems [46, 50]. In some countries with recent and well-conducted prevalence surveys, incidence is calculated from a combination of these survey findings and estimates of the duration of disease, with the latter derived from pre-chemotherapy era natural history observations or from mathematical models [4]. Because of uncertainties around the duration of untreated disease, incidence estimates from these methods differ and optimal methods remain elusive [51].

In the majority of cases, incidence is estimated using data acquired from routine programmatic notification data which is then adjusted by a number of methods, including expert opinion regarding the magnitude of unreported or undiagnosed cases [52]. Underdiagnoses occurs either as a result of failure to detect disease in patients presenting for care or because care is never sought, for example due to geographical or financial barriers [5]. Thus, because of uncertainty around the proportion of undetected or unreported cases, the incidence estimate is also uncertain [4]. Results from prevalence surveys have often revealed considerable underestimation of TB burden calculated using methods dependent on expert opinion [1, 52].

The other set of TB incidence estimates also come from the global TB disease burden estimates from the Institute of Health Metrics and Evaluation (IHME). This method uses a Bayesian meta-regression model to produce incidence estimates by using expert provided case detection rates (CDR) as a covariate [53]. Estimates from this method and the WHO methods have provided widely varying estimates and the thus the need for method refinement and additional data sources has been underscored [54].

For the purpose of tracking progress towards TB elimination, population-level TB prevalence surveys are resource-intensive, mainly because there are usually few cases of active TB in a community (between 0.1% and 1% in highly endemic settings) at any given time [37]. The ideal alternative is to improve surveillance coverage, so that notifications become complete enough to truly reflect incidence, although this remains a distant goal such that alternative approaches to estimating TB incidence entirely independent of expert opinion are vital.
1.1.4 Epidemiology of tuberculosis in Ethiopia

Efforts to control tuberculosis in Ethiopia started in the early 1960s with the establishment of TB centres and sanatoriums in major urban areas, [55, 56] long after the recognition of the disease as a major threat to public health. The distribution of TB has shown considerable spatial variation with a predominance in urban areas since the 1940s [57]. The first national TB infection survey conducted in 1953-1955 estimated a 3.0 percent ARTI, which declined to 1.4% during the second such survey undertaken in 1988-1989 [56].

Ethiopia piloted a DOTS programme in 1991, which was fully implemented nationwide by 1997 [55, 58]. To improve access to TB care, the country started to involve the private sector in the delivery of TB services in 2006 [59, 60]. Similarly, the Ethiopian community-based Health Extension Program (HEP) has improved access by helping people with TB symptoms to access TB diagnostic facilities [61]. Through these interventions, Ethiopia has made significant progress in TB prevention and control over the past two decades, achieving all three MDG targets for tuberculosis control by 2015. The rate of decline in national TB incidence is now among the highest in the world at 6.9 percent per year, although less is known about local dynamics. However, the reported rapid overall national decline might not reflect the dynamics at different subnational levels because there could still be some transmission hotspots locally.

Despite the above achievements, today Ethiopia still remains one of the world’s 30 high TB, high TB/HIV, and high MDR-TB burden countries [8] with about 29,000 deaths and 172,000 new incident cases in 2017 [4]. These figures mean that the country is now the third highest contributor to the WHO Africa regional TB incidence, after Nigeria and South Africa [62]. In 2014, TB was the 7th top cause of mortality in the country [63].

Drug-resistant TB and HIV-associated TB remain the main challenges for TB control in the country. In 2017, 2.7 percent of all new TB cases and 14 percent of previously treated cases of TB were MDR [4]. The incidence of MDR/RR-TB in the country is estimated at 5.2 per 100,000 population. Currently, MDR-TB treatment coverage (as a percentage of the estimated incidence of MDR/RR-TB) is very low, with less than 15 percent of estimated MDR-TB patients with initiated on effective treatment. Treatment success among MDR cases in Ethiopia is 75% which is much higher than the global average, although the number of MDR cases enrolled into treatment is very low. HIV infection presents another immense challenge for TB control, with about 7 percent of TB cases coinfected with HIV [4, 64]. The national TB program
tests approximately 86 percent of notified TB cases for HIV and 92 percent of newly diagnosed HIV-positive TB patients were started on ART [4].

The End TB Strategy calls for universal access to drug susceptibility testing (DST); that is, DST for at least rifampicin for all TB cases, plus DST for at least fluoroquinolones and second-line injectable agents among all TB cases with rifampicin resistance. However, only about 42% of new Ethiopian TB cases are tested for rifampicin resistance, leading to inappropriate treatment of initial MDR/RR-TB cases as drug-susceptible TB [65]. This inappropriate treatment can fuel community transmission and lead to poor health outcomes.

Although TB control efforts in Ethiopia have been in place for decades, current TB incidence is high and the wide gap between high burden districts and low burden districts has remained unchanged [66, 67] mainly due to limited resources. In 2017 the TB funding shortfall was estimated at 56 percent of the total budget required for TB control and, as a result of this funding gap, about 42 percent of pulmonary cases were not bacteriologically confirmed in 2016 [1]. Thus, optimally designing efficient and effective strategies is critical for resource allocation and continued progress towards TB elimination.

Engagement of communities is one of the core components of the End TB Strategy. Community-based TB activities, including a wide range of activities that contribute to the detection, referral and treatment of TB are carried out by community health workers (CHWs) and volunteers [1]. Ethiopia has adopted community-based practices for more than a decade, and the involvement of CHW has improved TB control substantially [68]. The HEP recruits and trains female health care workers from each kebele (the smallest administrative areal unit in Ethiopia) and forms an integral part of Ethiopian health care system. More than 40,000 health extension workers are currently deployed, and their TB-related duties include helping individuals with TB symptoms access diagnostic testing and supporting TB patients during treatment.
1.2 Introduction to spatiotemporal analysis of tuberculosis

According to the “epidemiologic triad” (the most widely accepted disease causation model), infectious disease is the result of an interplay between the infectious agent, the host and the environment. Individual-level studies (case-control and cohort) are most often conducted in a limited number of geographical areas, leading to the recruitment of cohorts of individuals with similar environmental exposures or societal backgrounds [69]. As a result, these studies may fail to identify the effects of environmental, cultural or societal factors on health outcomes [69]. However, spatial studies provide a platform to examine the links between disease incidence and environmental, social and behavioural factors [70].

Spatial analysis is used to detect disease clusters and predict spatial disease risk [71, 72]. It also provides aetiological clues by highlighting sources of heterogeneity underlying spatial patterns in health outcomes [73, 74] and has now become a key tool for priority setting and resource allocation [10]. The objectives of spatial data analysis are the description of spatial patterns (visualisation), the identification of disease clusters (exploration) and the explanation or prediction of disease risk (modelling). The following section provides a description of these three objectives of spatial analysis.

1.2.1 Data visualisation

Geographic data visualisation involves presentation of disease patterns using maps, which may assist in the formulation of aetiological hypotheses [75]. Maps provide an engaging and easy-to-understand means to visualise spatial data and can be used to describe patterns, identify outliers and communicate findings. Dependent on data availability and purpose, maps can represent point locations of cases or cases aggregated into administrative areas, either as rates or counts [76]. Often exact spatial coordinates are unavailable for individual cases, and thus rates or counts by administrative areas must be mapped. Point maps generally present clearer messages if the number of points is small and the density of clustering is not excessive [77]. If points are too densely clustered, interpretation of the map can be facilitated either by generating estimates aggregated at an administrative level or by applying smoothing methods [78].

Maps based on aggregated cases may involve plotting case counts, rates or excess risk (e.g. standardised morbidity ratios, SMRs) by administrative unit [75, 78]. Although such maps are easy to interpret, they can introduce bias because the size of the regions and the locations of their boundaries are typically a reflection of arbitrary administrative divisions rather than the
spatial distribution of epidemiological factors [77]. Thus, patterns observed across regions may depend on the spatial scale chosen, an effect known as the modifiable areal unit problem (MAUP) [79].

When administrative areas differ in risk and population size, the rates of disease for those areas may have different degrees of variability. In particular, rates for areas with small populations are often highly stochastically variable, giving a false impression of extremes in disease burden [80]. Therefore, spatial tools in common use do not typically provide reliable estimates [73, 81]. Recently, advanced spatial analysis tools that deal with numerical problems have been introduced to facilitate visualisation of spatial patterns of disease risk [81]. Two commonly applied techniques are empirical Bayes smoothing and spatial empirical Bayes smoothing.

**Empirical Bayes smoothing:** In empirical Bayes estimation, rates are adjusted upward or downward towards the regional rate, according to the size of the population on which they are based [80]. This method estimates the spatially-varying disease risk (posterior) from a weighted combination of the observed local risk (likelihood) and the regional risk (prior) [78]. The relative weights of the two components depend on the local population size, such that areas with a large population size receive a strong weight and so are less smoothed. In contrast, if the local population is relatively small, the derived estimates will be smoothed towards the overall disease rate to a greater extent (i.e., greater weighting on the prior) [78, 82].

**The Spatial Empirical Bayes Smoothing:** Smoothing towards an overall disease rate as in empirical Bayes smoothing can lead to loss of information that could occur as a result of spatial autocorrelation. Under the spatial empirical Bayes smoothing approach, disease rates are smoothed towards the average disease in the adjacent regions to preserve spatial dependence [80, 83].

1.2.2 Overview of spatial cluster identification methods

Although the term clustering can be used to refer to a wide range of geospatial, molecular and statistical techniques, in the following discussion, I focus on the geospatial meaning only. A disease cluster refers to any area of significantly higher risk within the study region [84]. Spatial clustering methods are tools which can be used to answer various fundamental public health questions related to the presence of disease clusters in a given area [80].

Many cluster exploration methods rely on known probability distributions, such as the Poisson distribution, or may utilise Monte Carlo simulation methods, which involve generating a large
of random possible outcomes [80]. Known or simulated distributions can then be compared with observed clustering to quantify deviation from expected patterns. These methods generally assess presence of global clustering or local clustering [85]. Global tests assess evidence of clustering without determining specific locations of clusters, whereas local tests determine the presence of clusters around specific locations with particular exposure level [80, 86].

Many of the methods for exploring spatial clustering rely on defining spatial weights in order to describe spatial autocorrelation. Spatial weights describe the proximity of spatial units and are often defined based on adjacency (also termed contiguity, i.e. whether or not areas share a common boundary). In this case, if areas $i$ and $j$ border one another, they are considered adjacent and are assigned a spatial weight of one ($\omega_{ij} = 1$), with all other such weights being set to zero [80]. Spatial autocorrelation can then be identified where disease risks at nearby locations are more closely related than disease risks in distant locations [87].

The following sections provide descriptions of commonly used cluster identification methods and methods used or reviewed in the spatial analysis of TB in this research. Other cluster detection methods used in the spatial analysis of TB are presented in Chapter 3.

**1.2.2.1 Gi and Gi*(d) Getis and Ord’s local Gi(d) statistic**

The local G-Statistic identifies the location and degree of clustering of disease events in space [78]. The distance-based $Gi$ and $Gi^*$ statistics of Getis and Ord can be computed for each location in the data set as the ratio of the sum of the values in neighbouring locations (defined to be within a given distance band $(d)$ or order of contiguity) to the sum over all the values in the study area. The two statistics differ in that the value observed at location $i$ is included in the calculation of $Gi^*$ but not for $Gi$. Considering an area divided into multiple subregions such that $y_j$ refers to the observed disease burden (e.g. incidence or prevalence) for a region that neighbours region $i$, and $\omega_{ij}(d)$ is a distance-based binary spatial weights matrix defining the nearness of region $i$ to region $j$, the $Gi$ statistic is [88, 89]:

$$Gi(d) = \frac{\sum_j \omega_{ij}(d)(y_j)}{\sum_j y_j}$$

$Gi^*(d)$ is the standardised form of $Gi(d)$ statistic, which is the ratio of the deviation of an attribute value from average attribute value to the standard deviation and is given as follows:
where $x_j$ is the health event for feature $j$, $\omega_{ij}$ is the spatial weight between feature $i$ and $j$, $n$ is the total number of features and:

$$
\tilde{y} = \frac{\sum_{j=1}^{n} y_j}{n}, \quad \text{and} \quad s = \sqrt{\frac{\sum_{j=1}^{n} y_j^2}{n} - \left(\tilde{y}\right)^2}
$$

### 1.2.2.2 Local Indicators of Spatial Autocorrelation (LISA)

Local Moran’s I is a special case of Pearson’s correlation coefficient that accounts for spatial dependence and measures the correlation between values of the same variable in different locations [77]. The local Moran’s I test detects local spatial autocorrelation by decomposing the Global Moran’s I statistic into contributions for each area within a study region as follows [80]:

$$
I_i = z_i \sum_{i \neq j} \omega_{ij} z_j, \quad \text{where} \quad z_i = \frac{x_i - \bar{x}}{s}
$$

where $z_i$ and $z_j$ represent the standardised observed values of the health indicator of interest in area $i$ and its neighbours, and $\omega_{ij}$ is a spatial weight measuring the nearness of subareas $i$ and $j$ [78]. The value of local Moran’s I ranges between minus -1 and 1, with a score of zero indicating no clustering and a positive score indicating a spatial concentration of similar values [85, 87]. These may be either clusters of high values that represent hotspots or clusters of low values that represent cold spots. The local G-Statistic described above assesses for hotspots and cold spots only. By contrast, LISA statistics provide information not just on hotspots and cold spots, but also spatial outliers [78, 80], and hence is a focus of this PhD thesis (Chapter 5).

LISA can be used to test the null hypothesis that the health outcome of interest has no local spatial association by Monte Carlo methods as proposed by Anselin [90]. The method first determines a reference distribution of the LISA statistic under the null hypothesis of no spatial autocorrelation. To determine the reference distribution, the observed data values are randomly redistributed among subareas. This random reassignment is performed a large number of times and the LISA statistic is computed for each subregion for each randomisation. The reference distribution describes the LISA values that would be expected under the null hypothesis of no spatial autocorrelation. The observed LISA statistic for subarea $i$ is then compared with the reference distribution to determine the statistical significance of local spatial clustering [80].
1.2.2.3 Kulldorff’s Spatial Scan Statistic

Kulldorff’s spatial scan statistic is a widely used cluster detection method in spatial epidemiology implemented in SaTScan software [84]. The method scans the study area by superimposing a large number of circular or cylindrical windows centred on the centroids of each subregion with a varying scan radius from zero up to a predefined maximum [91, 92] to identify the most likely spatial clusters [93-95]. The method randomly distributes events among the set of all locations and determines the most likely cluster (which by definition is a false cluster under the null hypothesis) and determines its associated likelihood ratio statistic. The algorithm then repeats this exercise many times, retaining the maximum likelihood ratio statistic for each random allocation. The statistical significance of the most likely cluster detected in observed dataset is estimated by determining the proportion of simulated maximised likelihood ratio test statistics exceeding that of the observed data [96]. The method can account for confounding factors, and can be used to detect clusters across both space and time which might be missed as a result of data averaging [77].

Determining the maximum radius of the circle is subjective but important, with different maximum radii leading to different clustering patterns. Clusters resulting from too large a maximum size could contain several epidemiologically unrelated locations, in contrast to clusters from smaller radii that could miss epidemiologically relevant clusters [97]. Thus, the difficulty in choosing optimal input parameters makes the use of this method problematic [84]. Further, the assumption of circular or elliptical cluster shape of the spatial scan statistic is not conducive to the identification of often irregular high-risk administrative areas and is thus unable to provide disease risk related to each natural areal unit [98, 99], posing additional challenges when determining optimal resource allocation. [100, 101].

1.2.3 Spatial statistical modelling

Epidemiological modelling of spatial data aims to explain or predict the occurrence of disease [77]. Standard regression models, which assume that disease rates at particular locations are independent of rates at neighbouring locations and that the residuals are normally distributed, are unsuitable for the analysis of spatial data that typically display spatial autocorrelation [86]. A number of factors may induce spatial dependence between regions, including unmeasured confounding, neighbourhood effects (subjects’ behaviour is influenced by that of subjects in the neighbouring region), and grouping effects (where subjects choose to be close to similar
subjects) [102]. Failure to account for spatial dependence may lead to inflated regression coefficients and spurious associations [103, 104].

Principles of spatial modelling have been well described elsewhere in the literature. In particular, the standard procedure is to start with classical ordinary least squares (OLS) regression models without spatial structure and then look for spatial dependence in the residuals, which would imply the need for a spatially explicit regression model [105]. Assessment of spatial lag and residuals is carried out for evidence of spatial autocorrelation often using Moran’s I test or the Lagrange Multiplier (LM-Error and LM-Lag test statistics). If there is no evidence of spatial dependence, results from the OLS may be reported [105]. In the presence of spatial autocorrelation, a regression model is selected statistically based on the attribute that shows spatial autocorrelation. For example, a spatial lag model, which incorporates an autoregressive term for the dependent variable, is used when there is spatial autocorrelation in the dependent variable [106], reflecting spill over or neighbourhood effect. However, autocorrelation in the residuals may also indicate that important variables have been omitted, for which the solution is to use a spatially lagged error model. The general spatial regression model selection algorithm [105] is presented in Figure 1 below.
Some spatial autocorrelation may be modelled by incorporating known risk factors into a regression model, but spatial correlation may remain in the model residuals after accounting for these covariate effects [106]. Currently, there are several spatial regression techniques, including Bayesian methods, that are suitable for incorporating random effects that may be either spatially structured or unstructured (spatially independent) [77].

1.2.3.1 Bayesian geospatial approaches

Recently, Bayesian methods have become popular in spatial epidemiological analyses [73, 93]. In spatial Bayesian modelling, the spatial pattern of disease risk is estimated by including covariates and a set of random effects into a model [81, 97, 101, 107]. Although various choices exist, conditional autoregressive (CAR) models are commonly used to account for spatial dependence in disease risk not captured by predictors [78, 101].

The CAR models capture spatial dependence by specifying the distribution of the random effect for a region as being dependent on the collection of random effects for all neighbouring regions.
The simplest CAR prior is the intrinsic conditional autoregressive (ICAR) model where the conditional expectation of \( \nu_i \) is equal to the mean of the random effects in neighbouring areas, while the conditional variance is inversely proportional to the number of neighbours \( n_i \). This variance structure recognises the fact that in the presence of strong spatial correlation, the more neighbours an area has the more information there is in the data about the value of its random effect [108]. The model is specified as:

\[
v_i | v_j \sim \text{Normal} \left( \frac{\sum_{j \neq i} v_j}{n_i}, \frac{\sigma^2}{n_i} \right)
\]

where \( j \) indexes neighbours for region \( i \), \( n_i \) is the number of regions that neighbour region \( i \), and \( \sigma^2 \) is the variance.

Apart from accounting for the residual error, CAR Bayesian regression models allow information sharing across regions - thereby producing robust estimates of disease risk in areas with limited information (sparse data) [74]. In spatial epidemiology, regions closer to each other are expected to be more similar, such that information from neighbouring sites can be used to improve the estimates of disease burden in a given area. Hence, spatially correlated random effects models are fitted by considering the variation at a particular site as normally distributed relative to the mean of its neighbours, producing smoothed rates across the study region. These Bayesian smoothing approaches produce robust and reliable estimation of health outcomes of interest in small areas, by using information from neighbouring areas [109].

The models described above provide global regression coefficients for entire regions, yielding a single statistic. In contrast, geographically weighted regression (GWR) is a local regression technique that measures the strength of association between predictors and response variables (disease rates) over space, providing estimates of local regression coefficients at each subregion studied [80, 106].

1.2.4 Spatial analysis of tuberculosis

Descriptions of spatial analyses of TB are provided in Chapters 2 to 5, with detailed descriptions of methodological applications presented in Chapter 3. In particular, methodological approaches to spatial TB cluster detection, spatial TB modelling and application of these methods to inform public health interventions of TB were presented in detail. Use of genotypic methods in improving our understanding of spatial TB clusters is also described in detail.
1.3 Introduction to mathematical modelling of tuberculosis

Mathematical models are tools to understand the natural history of infectious diseases, their mechanism of spread, the likely outcomes of various proposed control strategies and their economics [110]. In mathematical modelling, biological problems are described in mathematical concepts from which biological conclusions are drawn.

To determine the progression of infectious disease in a population, infectious disease modelling approaches can encompass both deterministic and stochastic frameworks, with deterministic compartmental models being particularly widely used. In compartmental models, the population is divided into different, mutually exclusive (disjoint) groups (termed compartments), each representing a specific stage of the disease, with the transitions from one compartment of the model to another described mathematically. These transition relationships, which define the dynamics of the disease within the simulated population are then formally written as equations (typically ordinary differential equations) that express the change in the number of individuals in the compartments over time [111].

Within the class of compartmental models, the Susceptible-Infectious-Recovered (SIR) framework is an archetypal model, although there are many extensions to this model that are used to increase realism [112]. In the SIR modelling framework, the population is partitioned into three compartments as Susceptible to the pathogen, Infected (and infectious), and Recovered (and thus no longer infectious), which are joined by flows representing the rate of change between each state [46]. The model operates by calculating the proportion of the population in each of these compartments at a series of time points, by determining the compartment size at each subsequent time step from the rates of transition and the compartment size at the preceding time step.

The potential transmissibility of an infectious agent is often measured using a concept known as the force of infection. Force of infection is the per capita rate at which susceptible individuals contract infection and is proportional to the prevalence of infection in the population, with the proportionality constant often estimated by fitting to data [113]. The transmissibility of the pathogen depends both on characteristics of the pathogen (e.g. the mechanism of transmission and how prolific the pathogen is within the host) and characteristics of the population (such as the number of contacts each individual makes per unit time). The prevalence of infection
determines the likelihood that any contact of a susceptible person is with an infectious individual [112].

Transmission of an infectious disease is often described by a quantity called reproductive number (\(R_0\)). The reproductive number is the average number of new primary infectious produced by a single infectious agent over the duration of time they are infectious when introduced into a totally susceptible population. Depending on the research setting, the reproductive number can be expressed as a product of infection rate and duration of infectiousness, making it useful in determining the future dynamics of disease in the population. An \(R_0 > 1\) indicates that the disease will continue to spread while an \(R_0 < 1\) indicates that the disease will eventually die out [114]. A recent systematic review reported TB reproductive number between 0.24 to 4.3, although the study lacks data from many high burden settings [115].

While SIR models are often suitable for diseases with short latent periods, models that capture latency are required for infections with very long latency periods such as TB. Therefore, mathematical TB models almost invariably extend upon the basic SIR models by adding greater complexity. Currently, there are dozens of TB model structures, which highlights the significant variation in several aspects of TB biology (especially disease progression following infection). The level of complexity ranges from models with no latency compartment to models with a sequence of latency compartments, with transition to active TB occurring from the final compartment only [116] or from all latency compartments [117].

Two recent systematic reviews that compared modelled outputs with empirical data demonstrated that the majority of TB models failed to reproduce the data, highlighting the need to include important features of TB into the models [116, 118]. Used in about half of TB modelling studies, a model with a single latency compartment is dominant, although it poorly fits empirical data from various settings. This model allows a proportion of susceptible individuals to bypass the latency compartment upon infection to immediately join the active TB compartment. However, TB models that incorporate two latency compartments (that are positioned either serially or in parallel) demonstrate good fit to data from various settings [116, 118]. A recent work by Trauer et al is an example of a two-latent compartment model positioned serially, where individuals enter an early latent compartment following infection and may progress rapidly to active disease or enter the late latent compartment from which
progression occurs more slowly [117]. This model structure is adapted in this thesis and is described in detail in Chapter 5.

Mathematical models of TB have been useful tools to test hypotheses that would be impractical to test using other methods, such as prospective or experimental studies [119]. In this regard, among other contributions, pioneering TB models have proven to be alternative methods to explain why TB incidence declined before the availability of effective treatments [119] and have quantified the impact of BCG vaccination and chemotherapy [120]. More recently, TB models have been used to provide insights into the types of programmatic interventions likely to maximise impact on the population level. These areas of application include assessment of the likely impact shorter MDR-TB regimens [121], TB diagnostics [122], various TB control programmes [123], and new vaccines and mass preventive treatment [124, 125], although bottlenecks to the uptake of such technologies persist.

1.3.1 Issues with standard TB transmission models

The standard deterministic TB transmission models assume no spatial variation in TB risk [120, 124], implying that the population is well-mixed, and that the probability of transmission is equal in different spatial subdivisions. In reality, populations are rarely well-mixed and thus contacts are not made at random in the population.

TB transmission often occurs within a household or small community because prolonged duration of contact is typically required for infection to occur [126], creating the potential for localised clusters to develop [127]. Spatial studies from both high endemic settings and low endemic settings have indicated that TB occurs in spatial clusters implying geographical pockets of intense transmission [94, 128]. The spatial spread of TB from these transmission hotspots is facilitated by population movement [112].

1.3.2 Approaches to spatially structured transmission dynamic models

In order to adequately incorporate spatial structure into transmission dynamic models, characterising how an infectious disease spreads across space (or spatial diffusion) is crucial. Spatial spread of an infectious disease depends on the predominant geographical patterns of interactions between infected and susceptible hosts.

Three types of spatial diffusion of infectious diseases have been described: contagious, hierarchical, and network [129]. Contagious diffusion reflects the gradual spread outward from
a point of origin to nearby locations and indicates localised human spatial interaction, such that people are more likely to interact with their neighbours than with those located farther away. This often occurs in regions characterised by factors that limit long range population mobility, such as low income or poor access to transportation. In hierarchical diffusion, disease starts from large cities, spreads to medium-sized cities and thence to smaller cities following transport corridors over long distances [80, 114]. Network diffusion, describes spread through social networks [130], which may lead to clusters in social gathering areas [131], although the role of spatial interactions in network diffusion varies according to the disease type and the geographical and social settings [80, 114].

Spatial heterogeneity can be incorporated into both population-based and agent-based models. Two population-based approaches: metapopulation models and spatially continuous models have been described for inclusion of spatial heterogeneities into mathematical models. Spatially continuous models assume a continuous population and are often used to model diffusion of contagious epidemics across space. By contrast, metapopulation models consider cross-infection between infectives and susceptible individuals in different spatial subdivisions and are computationally less intensive and more frequently used [132].

### 1.3.3 Spatial metapopulation models

A metapopulation model divides the population into a number of spatially discrete subgroups, with individuals in each spatial group assumed to mix homogeneously and with each spatial group coupled to one another in some way. Coupling terms are used to represent the way that infection spreads between groups [114]. Metapopulation models can be either mobility models or cross-coupled models depending on how the coupling between spatial regions is captured, with the latter being commonly used in the geographic spread of infectious diseases studies [129].

In spatially structured models, the rate at which individuals in different population groups become infected depends on the rate at which they interact with others in their subgroup and other subgroups [133]. The cross-coupled models do not incorporate explicit mobility among spatial groups, but rather describe transmissions between them by defining a coupling matrix, often abbreviated “WAIFW” (who acquires infection from whom). The WAIFW matrix describes the relative degree of transmission within a particular spatial group compared to transmission between groups, and can be described as:
Terms $\omega$ and $v$ are scaling factors that describe the relative amount of between-group coupling relative to within group coupling and their values are typically between 0 and 1 because between group interactions are generally less frequent than within group interactions. In other words, $\omega$ is the ratio of between-patch to within-patch effective contact rates [134]. The within and between group rates of transmission are found by multiplying the appropriate element of the coupling matrix by $\beta$, the transmission parameter. In the above model, the within transmission parameters are given by $\beta \times 1 = \beta$, while the between group transmission parameters are equal to $\beta \omega$ or $\beta v$ [114].

1.3.4 Spatial metapopulation TB transmission models

Although the application of spatially structured metapopulation models to TB is limited, some such models have provided great insights into the potential impact of TB interventions of TB. Spatially structured models have been used to explore the impact of population mobility using a two patch metapopulation framework (high vs low incidence settings) in the spatial spread of TB. This study documented the central role of cross region population mobility in the spatial spread of TB [135, 136], highlighting the potential impact of movement restriction on TB epidemiology (although impractical in reality). Therefore, failure to incorporate spatial heterogeneity of TB into the transmission dynamic models could lead to wrong predictions and parameter estimates. A similar study that modelled the impact of case detection demonstrated that targeting spatial TB hotspots could significantly reduce TB incidence compared to the same intervention applied across the general population [137]. In contrast, results from a study that evaluated the impact of hypothetical vaccine targeting spatial hotspots in India showed modest benefits [138], implying that spatial targeting could have various impacts in different settings. These studies were limited to overcrowded urban areas [137] and specific cross-border settings [136] and there was no consensus on how to define a spatial hotspot for spatial targeting.
1.4 Rationale for this doctoral research

In 2014, the World Health Assembly approved the ‘End TB Strategy’, which sets ambitious targets to achieve a 95% reduction in TB death and 90% reduction in TB incidence rate by 2035. To achieve these goals, new tools such as drugs and vaccines will be required [30, 32]. However, the development of new tools is challenging, with only a few tools progressing from the discovery stage to successfully entering TB clinical practice over recent years [32]. Therefore, until new tools become more widely available, the existing tools for diagnosis and treatment of patients with TB must be implemented effectively [30, 32, 139]. One such approach is targeting interventions to locations at increased risk of TB transmission [10, 140], although the impact of such spatial targeting needs to be predicted and measured as it is likely to be variable.

Evidence-based spatial targeting has the potential to greatly increase both the impact and cost-effectiveness of control measures, ensuring optimal use of the limited resources available for TB control. The local dynamics of TB vary across settings and depends on the disease prevalence and the degree of contact between individuals. Hence, the role of spatial hotspots in the spatial spread of TB and their role in prevention and control strategies needs to be systematically addressed before resources are allocated. In this regard, spatially structured metapopulation models are suitable to study the impact of spatially targeted interventions [137, 141], although there remains no consensus on how to define spatial hotspots.

Accurate detection of spatial TB hotspots can be challenging due to reasons described in Chapter 3 of this thesis and include uncertainty about the locations of infection and incomplete incidence data. TB also differs from other infectious diseases in several ways that are likely to influence apparent spatial clustering, including its long latency and prolonged infectious period that allow for significant population mobility between sequential cases [142]. Thus, clustering of cases may not necessarily indicate intense transmission, but could rather reflect aggregation of population groups at higher risk of disease, such as migrants [143]. In addition, use of surveillance data, which suffer from a considerable number of missing cases, may be biased due to differential case detection efforts resulting from inequitable access to care [144]. Thus, methods that account for undetected cases are particularly important in the identification of spatial TB clusters.
In Ethiopia (one of the top 30 high TB burden countries in the world), little has been documented about the spatial clustering of TB in rural and remote regions. In addition, the extent to which such spatial hotspots are responsible for the spatial spread of tuberculosis has not been systematically evaluated, such that the epidemiological gains of targeted public health interventions remain poorly understood.
1.5 Research aims

The overall aim of this thesis was to integrate spatiotemporal analysis methods with mathematical TB transmission models to evaluate spatial targeting strategies for TB prevention and control in a high TB endemic setting, Ethiopia.

Thus, in order to evaluate the benefits of spatial targeting in Ethiopian context, this thesis aims to answer the following research questions:

1. How are TB cases spatially distributed in a remote region of Ethiopia and what is the effect of health care access on the spatial distribution of TB?
2. Does the observed spatial pattern of tuberculosis in highly endemic settings reflect the true underlying incidence?
3. How should spatial TB hotspots be defined in the presence of a considerable number of undetected TB cases?
4. What is the role of spatial TB hotspots in propagating the spatial spread of TB in rural and remote regions of Ethiopia?
5. What is the likely impact of public health intervention strategies targeting spatial TB hotspots?

1.6 Thesis scope

The thesis includes two broad categories of studies: spatial analysis and mathematical modelling

Theme 1: Using spatial analysis to understand the distribution of TB and its drivers
- Describe the spatial distribution of TB and the factors driving the observed patterns using clinical records of TB patients.
- Conduct a survey of the application of the spatial methods to TB to understand methods that account for the effect of undetected cases.
- Develop a novel spatial method that accounts for undetected cases.

Theme 2: Using mathematical modelling to understand the role spatial hotspots in the spatial spread of TB and inform the development of spatially targeted case finding strategies.
• Develop a spatially structured mathematical model of TB transmission in high incidence settings to quantify the extent to which spatial hotspots contribute to TB transmission.

• Develop a spatially structured mathematical model of TB transmission in high incidence settings to investigate the impact of spatially targeted public health interventions for TB.

Figure 1. Scope of research
1.7 Overview of methods and materials

This section presents the data used in this thesis to answer the research questions described above. I used data from various sources including health facility records of TB patients, census results, online databases and shapefiles.

1.7.1 Study setting and data sources

Data on TB patients used in this thesis came from Sheka Zone in Southern Nations, Nationalities and Peoples Region (SNNPR) of Ethiopia. Ethiopia is located in the Horn of Africa and is bordered by Kenya, Somalia, Sudan, South Sudan, Eritrea, and Djibouti. The projected total population of Ethiopia was 92 million in 2016 [145]. The country is divided into nine regional states, which are further divided into zones, woredas and kebeles [45].

Sheka Zone is located 695 km from the national capital, Addis Ababa, and 970 km from the regional capital, Hawassa. The projected Zonal population was 247,815 in 2014 [146]. The Zone is divided into three administrative districts which are further divided into 66 kebeles. Kebeles are the smallest geographical administrative units and are typically used as the address of all individuals residing therein, with populations ranging from hundreds to thousands and the median kebele population in the Zone being 2200. Sheka Zone is served by 13 health centres, although only seven had a functional laboratory for TB sputum smear microscopy and none had access to x-ray or culture facilities at the time of data collection. TB diagnosis and treatment is based on National TB Treatment Guidelines [55]. Patients diagnosed with TB are registered on a Unit TB Register for DOTS in their presenting health facility, with recorded information including name, kebele of residence, age, sex, weight, sputum smear result, TB type, TB category, HIV status, use of chemoprophylactic therapy (cotrimoxazole and isoniazid), antiretroviral treatment (ART) status, TB treatment regimen, treatment outcome and dates of treatment initiation and of treatment outcome.

1.7.2 Data collection and management

Data on all TB patients treated between 2010 and 2014 in all health centres in Sheka Zone were collected from Unit TB registers. Data were entered into EpiData 3.2 (The EpiData Association, Odense Denmark) and exported to Stata 13.1 (College Station, Texas 77845 USA) and ArcMap 10.2.2 (ESRI, Redlands, California, USA) for analysis.
Only TB cases diagnosed in residents of the Zone were included in all spatial analyses presented in this doctoral thesis to obtain the actual spatial distribution in the Zone. TB cases from all health facilities were linked to their place of residence and thus, all spatial analyses in this thesis are based on patients’ place of residence and not place of presentation at the time of diagnosis. Attributes of kebeles including overall TB case counts, total population, population density and TB incidence were also calculated.

Polygon shape files for kebeles were obtained from the Central Statistical Agency of Ethiopia. The study was approved by the University of Melbourne Health Sciences Human Ethics subcommittee and the Zonal Health Department, Sheka Zone, Ethiopia.

Data were collected after ethical approval from the University of Melbourne, Health Sciences Human Ethics Subcommittee, Melbourne, VIC, Australia (Ethics Id – 1544898) and permission from the Zonal Health Department, Sheka Zone, Masha, Ethiopia.

### 1.7.3 Data analysis

In this thesis three broad analyses were conducted: spatial analysis, systematic review and transmission dynamic modelling. In the spatial analysis, a range of analyses was done, including the use of spatial cluster detection methods to identify statistically significant clusters. The spatial analysis also included generalised multivariate linear regression models and Bayesian conditional autoregressive models to determine drivers of spatial pattern and to account for missing cases respectively (Chapters 2 and 3). All spatial analyses in this thesis used queen contiguity spatial weights matrix that defines spatial relationship between spatial features as 1 if features shared a border or 0 otherwise.

The systematic review of the application of spatial methods to TB was performed to document the use and misuse of methods, with the results presented according to various methodological themes along with the proportion of times each method was reported.

Finally, transmission dynamic models were constructed to quantify the impact of spatial hotspots in the spatial spread of TB and to determine the potential impact of spatially targeted interventions. The impact of hotspot regions on the spatial spread of TB was determined by fitting the model to TB notification data using a Metropolis-Hastings algorithm (Chapters 5 and 6).
Spatial analyses in this thesis were performed using ArcGIS, GeoDa, WinBUGS and R, while transmission dynamic models were constructed in MATLAB. Stata was used for generalised linear regression models.
1.8 Organisation of the thesis

This thesis consists of various interrelated components, structured into the following seven major chapters:

Chapter 1: Introduction - The first chapter provides context to this PhD project by describing the current epidemiological status of TB (both global and local), spatial analysis, mathematical modelling, challenges in the current approaches to the spatial analysis of TB and how spatial analysis can be combined with mathematical models to inform public health interventions of TB. This chapter also presents the data sources used and approaches to data management.

Chapter 2: Spatial distribution of TB in rural and remote Ethiopia - This chapter explores the spatial distribution of TB in a remote region of Ethiopia and the drivers of the observed spatial pattern, including the effects of access. The chapter highlights how notification data can be biased towards areas with relatively better access to TB diagnostic facilities.

Chapter 3: Application of spatial methods to tuberculosis - This chapter presents a state-of-the-art review of the application of spatial analysis methods to TB epidemiology and provides a detailed description of the use and misuse of these methods, especially the impact of undetected or missing TB cases on the spatial distribution of TB. The chapter also highlights the role of genotypic methods in unpacking the driver of spatial TB hotspots.

Chapter 4: A novel Bayesian geospatial analysis - This chapter develops a novel methodological approach to improve spatial analysis of TB, providing an accessible platform to estimate TB incidence and case detection rates simultaneously. The chapter provides a detailed description of how a Bayesian geospatial binomial mixture model can be used to estimate TB incidence and case detection rate from routinely available notification data.

Chapter 5: The role of geospatial hotspots in the spatial analysis of TB - This chapter utilises data generated using the novel geospatial method presented in Chapter 4 to construct a spatially structured meta-population model to quantify the extent to which spatial hotspots account for the geographic spread of TB.

Chapter 6: Effect of spatially targeted interventions on TB epidemiology - This chapter uses a mathematical model and parameter estimates from Chapter 5 to evaluate the impact of public health intervention strategies on TB epidemiology.
Chapter 7: Discussion and conclusions: This chapter presents the main findings from thesis, their implications and proposes future research directions.
Chapter 2

2 Heterogeneity in the distribution of tuberculosis in Sheka Zone, Ethiopia: drivers and temporal trends

CHAPTER SUMMARY

Having established the rationale and objectives of the overall PhD thesis, this Chapter presents the spatial distribution of TB and its drivers in a remote region of Ethiopia. Using spatial analysis techniques including a spatial autocorrelation test, the Chapter examines whether a rural and remote region of Ethiopia where the population is scattered over a wide geographical area also has spatially heterogeneous TB distribution. Given that the spatial analysis presented in this Chapter is based on surveillance data and only a few local areas have functional health facilities, the Chapter quantifies the potential role of several drivers (in particular the role of access to health care) on the observed spatial TB pattern.

Importantly, the Chapter documents heterogeneous spatial TB distribution in this region of Ethiopia that was significantly correlated with local health facility availability. Thus, the Chapter highlights the potential bias arising from undetected cases, especially in areas with poor health care access, and calls for novel approaches to spatial analysis.
Heterogeneity of distribution of tuberculosis in Sheka Zone, Ethiopia: drivers and temporal trends

D. Shaweno,* T. Shaweno,† J. M. Trauer,*‡ J. T Denholm,§ E. S. McBryde**

*Department of Medicine, University of Melbourne, Melbourne, Victoria, Australia; †Department of Epidemiology, College of Health Sciences, Jimma University, Jimma, Ethiopia; ‡School of Public Health and Preventive Medicine, Monash University, Melbourne, §Victorian Tuberculosis Program at the Peter Doherty Institute for Infection and Immunity, Melbourne, ¶Department of Microbiology and Immunology, University of Melbourne, Melbourne, Victoria, #Australian Institute of Tropical Health & Medicine, James Cook University, Townsville City, Queensland, Australia

SUMMARY

OBJECTIVE: To describe the distribution of tuberculosis (TB) and its drivers in Sheka Zone, a geographically remote region of Ethiopia.

METHODS: We collected data on TB patients treated from 2010 to 2014 in the Sheka Zone. Predictors of TB incidence were determined using a multivariate generalised linear regression model.

RESULTS: We found significant spatial autocorrelation of TB incidence by kebele (the smallest administrative geographical subdivision in Ethiopia) (Moran’s I = 0.3, P < 0.001). The average TB incidence per kebele ranged from 0 to 453 per 100 000 population per year, and was significantly associated with average TB incidence across adjacent kebeles, TB incidence in the same kebele in the previous year and health facility availability. Each increment in TB incidence by 10/100 000/year in adjacent kebeles or in a previous year was associated with an increase in TB incidence of respectively 3.0 and 5.5/100 000/year. Availability of a health centre was associated with an increase in TB incidence of 84.3/100 000.

CONCLUSIONS: TB incidence in rural Ethiopia is highly heterogeneous, showing significant spatial autocorrelation. Both local transmission and access to health care are likely contributors to this pattern. Identification of local hotspots may assist in developing and optimising effective prevention and control strategies.

KEY WORDS: epidemiology; cluster analysis; transmission; tuberculosis

DESPITE CHANGES to public health policy and the adoption of the DOTS strategy some decades ago, Ethiopia has remained a high tuberculosis (TB) burden country.1 Although about 85% of TB cases in Ethiopia occur in rural and pastoralist areas, the epidemiology of TB in geographically remote regions remains incompletely described.2 Previously published epidemiological studies on TB in Ethiopia often focus on areas close to major zonal, regional or national capital cities.3–5 However, a comprehensive understanding of the TB epidemiology in both centrally located and less accessible areas is critical in guiding effective TB control efforts.

Past studies have shown heterogeneity in the spatio-temporal distribution of TB,6–8 while increased TB rates have been linked to population density,6,7,9 poverty,10,11 human immunodeficiency virus (HIV),12–14 urban residence15 and social gatherings.7 Available studies have documented both presence and persistence of high TB rates in some geographical regions of Ethiopia,6,8 although factors leading to the establishment and perpetuation of these heterogeneities are not fully understood.

In the present study, we aimed to understand the geographical and temporal distribution of TB and explore likely drivers using case records from a cohort of TB patients in a geographically inaccessible zone of Ethiopia.

METHODS

Setting
We collected data on TB patients registered for TB treatment from 2010 to 2014 in Sheka Zone, Southern Nations, Nationalities and Peoples Region (SNNPR), Ethiopia. Sheka Zone is geographically inaccessible (695 km away from the national capital, Addis Ababa, and 970 km from the regional capital, Hawassa) and is thus lacking basic facilities and infrastructure. Census data indicate a zonal population of 199 671 in 2007, projected to have increased to 247 815 in 2014.16 Sheka Zone is divided into
three administrative districts, which are further divided into 66 kebeles, the smallest geographical administrative units in Ethiopia, typically used as the address of all individuals residing there, with populations ranging from hundreds to thousands; the median kebele population in Sheka Zone is 2200. Although Sheka Zone is geographically remote, all kebeles in the administrative capitals of each district were treated as urban centres. In urban areas, kebeles represent neighbourhoods, while in rural areas they include both habitable and uninhabitable areas (farm lands and forests). The population density of the three districts varies widely: respectively 27.4, 62.9 and 263.0 persons per km² in Masha (northern), Andraccha (central) and Yeki (southern).  

Sheka Zone is served by 13 health centres, although during the study period only seven had functional laboratories for TB sputum smear microscopy and none had access to X-ray or culture facilities. TB diagnosis and treatment is based on Ethiopia’s national TB treatment guidelines, according to which patients with symptoms suggestive of pulmonary TB are considered smear-positive if at least two of three sputum samples are smear-positive. Patients with negative sputum smears who fail to respond to treatment with broad-spectrum antibiotics are considered to have smear-negative pulmonary TB, although the diagnosis of smear-negative and extrapulmonary cases also incorporates clinical judgement. Patients diagnosed with TB are registered in a TB Unit Register for DOTS at their presenting health facility; information on name, kebele of residence, age, sex, weight, sputum smear result, TB type, TB category, HIV status, use of chemoprophylactic therapy (cotrimoxazole and isoniazid), antiretroviral treatment (ART) status, anti-tuberculosis treatment regimen, treatment outcome and dates of treatment initiation and treatment outcome are recorded.

Data collection
Data on all TB patients diagnosed and/or treated from 2010 to 2014 at all health centres in the Zone were collected from TB Unit registers by TS, a faculty member of Jimma University, Ethiopia. As patients are entered sequentially in Unit registers, where treatment commencement dates were incomplete we used the midpoint between the start dates of the adjacent registered patients. Data from health centres with power supply were directly entered into EpiData 3.2 (EpiData Association, Odense, Denmark); data from health centres that lacked power supply were transferred to a standardised data entry pro forma and then transferred to EpiData. Where kebeles of residence on TB Unit register did not match the names in a shape file, official names known to the Central Statistical Agency (CSA) were used after confirmation from local authorities. Data were analysed using Stata 13.1 (College Station, TX, USA) and ArcMap 10.2.2 (Environmental Systems Research Institute, Redlands, CA, USA).

Polygon shape files for kebeles were obtained from the CSA of Ethiopia and used for mapping and area calculations, with aggregate patient-level information then linked to these polygons. The annual projected population based on the 2007 census was used for population density calculation and as the denominator for incidence rate calculations.

Approach to analysis
To obtain an overview of the true geographical distribution of the disease in Sheka Zone, we included only cases of TB diagnosed in residents of the Zone. We calculated attributes of kebeles, including overall TB case counts, total population, population density and TB incidence. We evaluated whether the average TB incidence rate per kebele over the 5-year period was spatially autocorrelated using global Moran’s I. Generalised multivariate linear regression (Gaussian family with identity link) was performed to determine predictors of TB incidence. The outcome considered was TB incidence in each kebele in each year from 2011 to 2014, with the following predictor variables: health facility availability, population density, average TB incidence rate in adjacent kebeles in the same year, TB incidence rate in the previous year in the same kebele (including 2010), and proportion of presenting TB cases co-infected with HIV. Analogous models were developed for the outcome of smear-positive TB by year and kebele, with the same set of exposure variables. A multivariate model in which all covariates were initially included, but then excluded by backward elimination, is presented. All predictors not significant at \( P < 0.05 \) were eliminated from the final model, with the variable with the greatest \( P \) value dropped first. As we considered that treatment outcomes were less likely to be spatially correlated, we used Pearson’s \( \chi^2 \) to compare among kebeles.

The study was approved by the Melbourne University Health Sciences Human Ethics Subcommittee, Melbourne, VIC, Australia, and the Zonal Health Department, Sheka Zone, Masha, Ethiopia.

RESULTS
Patient characteristics
Among 1732 TB patients diagnosed and treated from 2010 to 2014 in all health facilities, 1683 were resident in 55 kebeles in Sheka Zone and were included in the final analysis. Virtually all the remaining TB patients (\( n = 49, 2.8\% \)) were from the neighbouring province (Gambella), and were excluded from further analysis.

The majority of the patients included were male (57.8%), young to middle-aged adults from urban kebeles. Pulmonary TB comprised around three
HIV status, TB category, Type of TB, Age, years, Residence, Sex

The higher incidence rates observed in visually demonstrates the spatial autocorrelation and average TB incidence and child TB incidence. This of population, health facilities (overall and

Figure 1 shows choropleth maps of distribution by kebele of population, health facilities (overall and with TB laboratory facilities), population density, average TB incidence and child TB incidence. This visually demonstrates the spatial autocorrelation and the higher incidence rates observed in kebeles with functional health centres. The distribution of both paediatric TB and overall TB followed similar patterns in the Zone. Average TB incidence rate per kebele ranged from 0 to 453/100 000/year (Figure 1). The percentage of kebeles with a TB incidence rate of ≥200/100 000 over the 5-year period was substantial, ranging from 7.6% in 2010 to 18.2% in 2014 (Figure 2).

Overall, TB incidence in Sheka Zone increased from 111 to 151/100 000 over the study period. In this study, both the overall TB incidence and child TB incidence exhibited significant spatial autocorrelation (Moran’s I = 0.28 overall, 0.34 child, two-sided P < 0.001 overall, <0.001 child). Both overall TB incidence and sputum smear-positive TB incidence exhibited significant spatial autocorrelation by kebele for each study year except for 2011, with Moran’s I ranging from 0.23 to 0.30.

Treatment success (cure plus completion) showed significant difference by kebele (P= 0.001, Pearson’s χ² based on Monte Carlo simulation), although an analogous analysis showed no statistically significant difference in treatment outcome by access to diagnosis and treatment (P = 0.9, Pearson’s χ²).

Regression analysis
Availability of a health centre, average TB incidence in adjacent kebeles and TB incidence in the previous year significantly predicted overall TB incidence in the index kebele on both univariate and multivariate regression analysis (Table 2). Although population density appeared significant on univariate analysis, this variable was eliminated from the multivariate model. In the analogous analysis for smear-positive TB incidence, population density as well as all other variables significantly associated with overall TB incidence emerged as significant predictors (results not shown).

Availability of a functional health centre in a kebele was associated with an increase of 84/100 000 in TB incidence on multivariate regression, while an increase of 10/100 000 in average TB incidence in adjacent kebeles predicted a 3.0/100 000 increment in TB incidence in an index kebele. Similarly, an increase of 10 cases/100 000 in a given year predicted 5.5 cases/100 000 in the subsequent year in the same kebele. Population density and proportion of TB cases with HIV were not related to TB incidence in the multivariate model. Overall, the model explained 50.6% (adjusted R² = 0.506) of the variation in the average TB incidence rate.

**DISCUSSION**
We found that TB distribution in Sheka Zone, a remote region of Ethiopia, is highly heterogeneous, with incidence ranging widely from 0 to 453/100 000/ year, and demonstrating significant spatial autocorrelation. TB incidence was associated with TB incidence in adjacent kebeles, TB incidence in the previous year and proximity of health facilities, suggesting roles for intense transmission within ‘hotspots’ and care access as important potential drivers. While broad demographic and epidemiological parameters reported here, such as proportion of new and retreatment TB, are similar to those observed in previous studies in other Ethiopian contexts, our study presents detailed consideration of geospatial drivers of incidence heterogeneity.3,21
As in previous studies from Ethiopia,\textsuperscript{6} we included only TB cases resident within the study zone, such that our overall incidence estimates are somewhat lower than official notification rates.\textsuperscript{21} Linking patients to their place of residence rather than their place of presentation,\textsuperscript{21} as in our study, is likely to provide a more accurate picture of the spread of the epidemic, as TB transmission occurs largely within households and communities.\textsuperscript{22} Although untraced TB patients treated in the neighbouring zones may lead to underestimation of the true TB incidence rate in the zone, the fact that less than 3% of patients registered at health centres from Sheka Zone reported residence outside of the zone suggests that cross-border notifications are a relatively minor consideration.

TB incidence observed in a given kebele was correlated with the average TB incidence in the neighbouring kebeles, likely suggesting community transmission of TB, as has also been demonstrated in Brazil through mathematical modelling.\textsuperscript{11} The spatial diffusion of infection across space is well described for other communicable diseases, particularly HIV.\textsuperscript{11,23–25} Moreover, the significant spatial autocorrelation observed in our study could also be indicative of contagious spread to neighbouring areas.\textsuperscript{6,8,26} In addition, the emergence of population density as a predictor of smear-positive TB incidence suggests that recent local transmission could also be a driver of TB epidemiology.\textsuperscript{7}

TB incidence was an important predictor of TB incidence in the subsequent year. This finding highlights the persistence of TB in specific locations, with likely factors including persistent community transmission and durable underlying population vulnerabilities. Moreover, it could also imply that
Figure 2  The spatial and temporal trends in rates of tuberculosis by kebele in Sheka Zone, Ethiopia, 2010–2014. TB = tuberculosis. This image can be viewed online in colour at http://www.ingentaconnect.com/content/iuatld/ijtld/2017/00000021/00000001/art00015

Table 2  Predictors of TB incidence by kebele in Sheka Zone, Ethiopia, 2010–2014

<table>
<thead>
<tr>
<th>Predictors</th>
<th>Univariate Coefficient (95%CI)</th>
<th>Multivariate Coefficient (95%CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incidence in adjacent kebeles*</td>
<td>0.61 (0.25 to 0.96)</td>
<td>0.29 (0.12 to 0.47)</td>
</tr>
<tr>
<td>Incidence in the previous year*</td>
<td>0.70 (0.56 to 0.84)</td>
<td>0.55 (0.45 to 0.66)</td>
</tr>
<tr>
<td>Health centre availability†</td>
<td>170 (60 to 280)</td>
<td>84.3 (34.2 to 134.3)</td>
</tr>
<tr>
<td>Population density‡</td>
<td>0.04 (0.01 to 0.07)</td>
<td>§</td>
</tr>
<tr>
<td>Proportion of HIV among TB</td>
<td>140 (–170 to 450)</td>
<td>§</td>
</tr>
</tbody>
</table>

* Cases/100,000 population/year.
† Health centre available as compared to unavailable.
‡ Persons/km².
§ Eliminated by backwards stepwise regression analysis.
TB = tuberculosis; CI = confidence interval; HIV = human immunodeficiency virus.
the observation of higher numbers of TB cases in a specific year occurs partly as a result of insufficient public health actions in preceding years. This hypothesis is consistent with observations from other areas of high TB persistence in Ethiopia, and further work into correlation of public health activity and disease control will be critical to ensure that targeted strategies are developed towards reducing the TB burden.6

The observed TB incidence is also closely linked to the proximity of functional health facilities capable of diagnosing TB. Although significant heterogeneity was observed in the outcomes reported by individual health centres, kebeles with functional health centres consistently reported higher TB incidence than those without. This finding highlights inequalities in TB case detection as a function of distribution of health facilities, and implies a need to account for the potential role of access, especially when notification data are the main source of information for the National TB Control Program planning. Priority setting and resource allocation based solely on notification data, without accounting for issues of access, could mask true incidence in areas with poor access, and could even fuel community transmission by diverting attention and resources to the wrong places. In this study, trends in TB incidence were found to change over the study period. These changing trends could be related to four (31%) health centres that were equipped to provide TB sputum smear testing during the study period,27 and to the community TB programme that was introduced in 2013. However, rather than public health responses leading to a decline in disease burden, the increasing number of presentations over a period during which health care services were expanded is also consistent with access to care uncovering a greater number of cases.

Geographic access to health care poses an immense challenge for early presentation and adherence to care, as most rural kebeles have no road network and many are located more than 2 h walking distance from health facilities.28,29 As travel from almost all rural kebeles is on foot, TB patients from rural kebeles of Northern Ethiopia have reported travelling for 2–10 h for initial diagnosis and treatment.29

As mycobacterial culture is not part of standard of care in Ethiopia, the possibility of missing sputum smear-negative patients who would be diagnosed as culture-positive cases if laboratory facilities were available is unavoidable. The absence of culture and drug susceptibility testing in this setting means that no consideration of multidrug-resistant TB has been possible in this present study. However, the ratio of smear-negative to smear-positive TB in our study is comparable with the ratio of smear-negative, culture-positive cases to smear-positive TB from a population-based survey in Ethiopia.30 Our data set is therefore likely to be consistent with patterns observed broadly across the country. In the absence of sputum culture, it is also impossible to comment on the extent to which drug resistance is driving transmission in this area.

CONCLUSIONS

TB incidence in rural Ethiopia is highly heterogeneous and demonstrates significant spatial autocorrelation. Transmission within and between kebeles, as well as proximity to health care, are likely drivers of these patterns. Identification of local TB hotspots may assist in developing more effective locally targeted strategies, and will be critical for ensuring that limited resources for TB control efforts are optimally engaged in this high-burden setting. The likely under-reporting of TB cases in areas remote from health facilities implies a potentially important role for pro-active interventions, such as active case finding, and highlights the importance of more objective approaches to estimating total burden, such as prevalence surveys.

Acknowledgements

The authors thank Sheka Zone Health Department, Masha, Ethiopian and all health centres for permission to access the data; and the Ethiopian Central Statistical Agency, Addis Ababa, Ethiopia, for the provision of kebele polygons.

Conflicts of interest: none declared.

References


OBJECTIF : Décrire la distribution de la tuberculose (TB) et ses moteurs dans la Sheka Zone, une région géographiquement isolée d’Ethiopie.

MÉTHODE : Nous avons recueilli des données relatives aux patients TB traités entre 2010 et 2014 dans la Sheka Zone. Les facteurs de prédiction de l’incidence de la TB ont été déterminés grâce à un modèle de régression linéaire multivarié généralisé.

RÉSULTATS : Nous avons trouvé une autocorrélation spatiale significative de l’incidence de la TB par kebele (la plus petite subdivision géographique administrative en Ethiopie) (Moran’s I = 0,3 ; P < 0,001). L’incidence moyenne de la TB par kebele allait de 0 à 453 par 100 000 personnes par année et était significativement associée à l’incidence moyenne de la TB dans les kebeles adjacents, à l’incidence de la TB dans le même kebele pendant l’année précédente et à la disponibilité des structures de santé. Toute augmentation de l’incidence de la TB de 10/100 000 personnes/an dans les kebeles adjacents ou dans une année précédente était associée à une augmentation de l’incidence de la TB de 3,0 et 5,5/100 000/an, respectivement. La disponibilité d’un centre de santé a été associée avec une augmentation de l’incidence de la TB de 84,3/100 000.

CONCLUSION : L’incidence de la TB dans l’Ethiopie rurale est très hétérogène et montre une autocorrélation spatiale significative. Ce profil est probablement lié à la fois à la transmission locale et à l’accès aux soins de santé. L’identification de points chauds locaux pourrait contribuer à élaborer et à optimiser des stratégies efficaces de prévention et de lutte.

RESUMEN

OBJECTIVO: Describir la distribución de la tuberculosis (TB) y los factores que favorecen la aparición de la enfermedad en la zona Sheka, una región geográficamente apartada de Etiopía.

MÉTODOS: Se recogieron datos sobre los pacientes con TB del 2010 al 2014 en la zona Sheka. Se definieron las variables independientes de la incidencia de TB mediante un modelo de regresión lineal generalizada multivariante.

RESULTADOS: Se observó una autocorrelación espacial significativa de la incidencia de TB por kebele, que es la subdivisión geográfica administrativa más pequeña en Etiopía (índice de Moran = 0,3; P < 0,001). La incidencia promedio de TB por kebele osciló entre 0 y 453 por 100 000 habitantes por año y se observó una asociación significativa con la incidencia promedio en los kebeles adyacentes, la incidencia de TB en el mismo kebele en el año anterior y la disponibilidad de establecimientos de salud. Cada incremento en la incidencia de TB de 10/100 000/año en los kebeles adyacentes se asoció con un aumento de la incidencia de 3,0 casos/100 000 en un kebele de referencia y un aumento equivalente en 1 año pronosticó un incremento de 5,5 casos/100 000 el año siguiente en un mismo kebele. La disponibilidad de un centro de atención de salud se asoció con un aumento de la incidencia de TB de 84,3/100 000.

CONCLUSIÓN: La incidencia de TB en las regiones rurales de Etiopía es muy heterogénea y exhibe una sólida autocorrelación espacial. La transmisión local y también el acceso a un centro de salud favorecen esta distribución. El reconocimiento de importantes focos de transmisión puede contribuir a elaborar estrategias eficaces de prevención y control y a optimizarlas.
Chapter 3

3 Methods used in the spatial analysis of tuberculosis epidemiology: A systematic review

Chapter overview

Having described the inherent bias arising from missing cases in surveillance data in Chapter 2, this Chapter seeks to understand suitable spatial analysis methods that account for missing cases in TB notification data. Thus, this Chapter highlights how notification data are frequently used in the spatial analysis of TB and the implications for accurate identification of spatial hotspots. In addition, the Chapter thoroughly describes the use and misuse of spatial methods in the spatial analysis of TB and discusses the use of spatial methods to inform public health interventions. The importance of combining spatial analysis methods with genotypic methods to characterise the drivers of TB dynamics in the spatial hotspots for avoiding resource misallocation is also described in detail.
Methods used in the spatial analysis of tuberculosis epidemiology: a systematic review

Debebe Shaweno1,2*, Malancha Karmakar2,3, Kefyalew Addis Alene4,5, Romain Ragonnet1,6, Archie CA Clements7, James M. Trauer2,8, Justin T. Denholm2,3 and Emma S. McBryde1,9

Abstract

Background: Tuberculosis (TB) transmission often occurs within a household or community, leading to heterogeneous spatial patterns. However, apparent spatial clustering of TB could reflect ongoing transmission or co-location of risk factors and can vary considerably depending on the type of data available, the analysis methods employed and the dynamics of the underlying population. Thus, we aimed to review methodological approaches used in the spatial analysis of TB burden.

Methods: We conducted a systematic literature search of spatial studies of TB published in English using Medline, Embase, PsycInfo, Scopus and Web of Science databases with no date restriction from inception to 15 February 2017. The protocol for this systematic review was prospectively registered with PROSPERO (CRD42016036655).

Results: We identified 168 eligible studies with spatial methods used to describe the spatial distribution (n = 154), spatial clusters (n = 73), predictors of spatial patterns (n = 64), the role of congregate settings (n = 3) and the household (n = 2) on TB transmission. Molecular techniques combined with geospatial methods were used by 25 studies to compare the role of transmission to reactivation as a driver of TB spatial distribution, finding that geospatial hotspots are not necessarily areas of recent transmission. Almost all studies used notification data for spatial analysis (161 of 168), although none accounted for undetected cases. The most common data visualisation technique was notification rate mapping, and the use of smoothing techniques was uncommon. Spatial clusters were identified using a range of methods, with the most commonly employed being Kulldorff’s spatial scan statistic followed by local Moran’s I and Getis and Ord’s local Gi(d) tests. In the 11 papers that compared two such methods using a single dataset, the clustering patterns identified were often inconsistent. Classical regression models that did not account for spatial dependence were commonly used to predict spatial TB risk. In all included studies, TB showed a heterogeneous spatial pattern at each geographic resolution level examined.

Conclusions: A range of spatial analysis methodologies has been employed in divergent contexts, with all studies demonstrating significant heterogeneity in spatial TB distribution. Future studies are needed to define the optimal method for each context and should account for unreported cases when using notification data where possible. Future studies combining genotypic and geospatial techniques with epidemiologically linked cases have the potential to provide further insights and improve TB control.

Keywords: Spatial analysis, Tuberculosis, Genotypic cluster

* Correspondence: debebesh@gmail.com
1Department of Medicine, University of Melbourne, Melbourne, Victoria, Australia
2Victorian Tuberculosis Program at the Peter Doherty Institute for Infection and Immunity, Melbourne, Victoria, Australia
Full list of author information is available at the end of the article

© The Author(s). 2018 Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated.
Background

*Mycobacterium tuberculosis* (*Mtb*) transmission often occurs within a household or small community because prolonged duration of contact is typically required for infection to occur, creating the potential for localised clusters to develop [1]. However, geospatial TB clusters are not always due to ongoing person-to-person transmission but may also result from reactivation of latent infection in a group of people with shared risk factors [1, 2]. Spatial analysis and identification of areas with high TB rates (clusters), followed by characterisation of the drivers of the dynamics in these clusters, have been promoted for targeted TB control and intensified use of existing TB control tools [3, 4].

TB differs from other infectious diseases in several ways that are likely to influence apparent spatial clustering. For example, its long latency and prolonged infectious period allow for significant population mobility between serial cases [5]. Thus, *Mtb* infection acquired in a given location may progress to TB disease in an entirely different region, such that clustering of cases may not necessarily indicate intense transmission but could rather reflect aggregation of population groups at higher risk of disease, such as migrants [6]. Similarly, *Mtb* infection acquired from workplaces and other congregate settings can be wrongly attributed to residential exposure, as only an individual’s residence information is typically recorded on TB surveillance documents in many settings [7, 8].

Identifying heterogeneity in the spatial distribution of TB cases and characterising its drivers can help to inform targeted public health responses, making it an attractive approach [9]. However, there are practical challenges in appropriate interpretation of spatial clusters of TB. Of particular importance is that the observed spatial pattern of TB may be affected by factors other than genuine TB transmission or reactivation, including the type and resolution of data and the spatial analysis methods used [10]. For instance, use of incidence data versus notification data could give considerably different spatial pattern [11], as the latter misses a large number of TB cases and could be skewed towards areas with better access to health care in high-burden settings [12, 13]. Thus, spatial analysis using notification data alone in such settings could result in misleading conclusions.

Similarly, the type of model used and the spatial unit of data analysis are important determinants of the patterns identified and their associations [14–16]. That is, different spatial resolutions could lead to markedly different results for the same dataset regardless of the true extent of spatial correlation [15, 17, 18] and the effect observed at a regional level may not hold at the individual level (an effect known as the ecological fallacy) [19]. Therefore, we aimed to review methodological approaches used in the spatial analysis of TB burden. We also considered how common issues in data interpretation were managed, including sparse data, false-positive identification of clustering and undetected cases.

Methods

Data source and search strategy

Our search strategy aimed to identify peer-reviewed studies of the distribution and determinants of TB that employed spatial analysis methods. In this review, studies were considered spatial if they incorporated any spatial approaches (e.g. geocoding, spatial analysis units, cluster detection methods, spatial risk modelling) into the design and analysis of the distribution, determinants and outcomes of TB [20]. We searched Medline, Embase, Web of Science, Scopus and PsycInfo databases from their inception to 15 February 2017 using a combination of keywords and medical subject headings (MeSH) pertaining to our two central concepts: tuberculosis and space. We refined search terms related to the latter concept after reviewing key studies, including a previous systematic review not limited to TB [21]. The full search strategy was adapted to the syntax of the individual database from the following conceptual structure: (tuberculosis OR multidrug-resistant tuberculosis) AND (spatial analysis OR geographic mapping OR spatial regression OR spatiotemporal analysis OR spatial autocorrelation analysis OR geography OR geographic distribution OR geographic information system OR geographically weighted regression OR space-time clustering OR ‘spati’*’ OR ‘hotspots’ OR cluster analysis) and is provided in the Appendix. Studies targeted to special populations (e.g. homeless, migrants, HIV-infected persons) and that considered the entire population of a region were permitted. Additional papers were also identified through hand searching the bibliographies of retrieved articles and from suggestions from experts in the field.

Eligibility, and inclusion and exclusion criteria

We included peer-reviewed papers that incorporated the spatial analysis approaches described above in the study of TB. After exclusion of duplicates, titles and abstracts were screened by two researchers (DS and MK) to identify potentially eligible studies. Of these papers, articles were excluded hierarchically on the basis of article type, whether the method used could be considered spatial or not and the outcomes assessed. No exclusions were made on the basis of the outcome reported, with studies that considered incidence, prevalence or any TB-related health outcome included. Studies were excluded if the language of the publication was not English, the report was a letter, conference abstract or a review or only reported the temporal (trend) of TB. Spatial studies of
non-tuberculous mycobacteria, non-human diseases and population immunological profiles were also excluded. Full-text articles were excluded if they did not provide sufficient information on the spatial analysis techniques employed. There were no exclusions based on study setting or anatomical site of disease.

Data extraction and synthesis
Three independent reviewers (DS, MK, KAA) performed data extraction using pretested data extraction forms and stored these in a Microsoft Excel 2016 spreadsheet (Microsoft Corporation, Redmond, Washington, USA). Disagreements were resolved by consensus. The following information was extracted from each paper: country, publication year, study aim, data type (notifications or survey), type of TB disease (smear-positive pulmonary, smear-negative pulmonary and extrapulmonary), geographic level, spatial methods (map types, cluster detection methods, statistical regression methods, spatial lag, spatial error, spatial smoothing techniques), time scale and outcomes reported (whether quantification of TB cases or TB-related health outcomes, such as mortality, default from care, disability-adjusted life years (DALYs) and key conclusions). In studies which combined geospatial methods with genotypic clustering methods, we also extracted the genotypic cluster identification methods. Spatial analysis techniques were categorised as either visualisation (mapping), exploration (using statistical tests to identify spatial clusters) or statistical modelling [19, 22]. Counts and proportions were primarily used to summarise study findings. The protocol for this systematic review was prospectively registered with PROSPERO (CRD42016036655). Although we adhered to our original published protocol, here we additionally describe the importance of genotypic methods and the application of spatial methods in informing public health interventions in response to requests during peer review.

Results
Study characteristics
A total of 2350 records were identified from the electronic searches, of which 252 full-text articles were assessed. Of these, 168 articles met all inclusion criteria and were included in the final narrative synthesis (Fig. 1). Using a cutoff of 100 TB cases per 100,000 population in reported incidence in 2016, 111 (66%) of the studies were from low-incidence settings.

All references returned by the search strategy were from the period 1982 to 2017, with 71% published from 2010 onwards (Additional file 1: Figure S1). Earlier studies (predominantly in the 1980s and 1990s) tended to be descriptive visualisations, while studies in the last two decades frequently incorporated cluster detection and risk prediction. More recently, a range of statistical techniques including Bayesian statistical approaches and geographically weighted regression have become increasingly popular.

Key objectives of included studies
Spatial analysis was applied to address a range of objectives (Table 1), with the commonest ones including description of the distribution (n = 135), statistical analysis of spatial clustering (n = 73) and analysis of risk factors and risk prediction (n = 64). Spatial methods were also used to determine the relative importance of transmission by comparison to reactivation as a driver of TB incidence (n = 25), the effect of TB interventions (n = 2), barriers to TB service uptake (n = 2), spatial distribution of TB-related health outcomes (mortality, default, hospitalisation) (n = 5), spatial pattern of TB incidence among people living with HIV (PLHIV) (n = 4), HIV-related TB mortality (n = 4), multidrug-resistant TB (MDR-TB) drivers (n = 1), TB outbreak detection (n = 3) and drivers of spatial clustering (including the role of congregate settings, such as social drinking venues and schools) (n = 30).

Types of TB disease analysed
Spatial analysis was most commonly conducted on data for all types of TB (i.e. without distinction between pulmonary or extrapulmonary; n = 121), followed by pulmonary TB only (n = 28) and smear-positive pulmonary TB only (n = 13). Spatial analysis of multidrug-resistant TB (MDR-TB) and extensively drug-resistant TB (XDR-TB) was reported in 15 studies and one study respectively.

Data used and scale of analysis
Nearly all studies used retrospective TB program data (notifications), with the exception of five studies that used prevalence surveys and two prospectively collected data. None of the studies using notification data accounted for undetected/unreported cases. In all included studies, spatial analysis of TB was based on the individual’s residence, except for three studies that explored the effect of exposure from social gathering sites.

Spatial analysis was generally done using data aggregated over administrative spatial units (n = 131), but the scale of aggregation differed markedly. Common spatial scales included census tract (n = 20), district (n = 15), postal code (n = 15), county (n = 15), neighbourhood (n = 10), health area (n = 7), municipality (n = 11), state (n = 7), province (n = 6), local government area (LGA) (n = 4) and ward (n = 4). Data were analysed at the individual level in 37 studies, while three studies were reported at a continent and country scale.
### Table 1: Application areas of spatial methods in TB studies

<table>
<thead>
<tr>
<th>Spatial method application areas</th>
<th>Methods used</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spatial TB distribution or spatial clustering</td>
<td>Dot maps, rate maps, thematic maps, Moran’s I, GetisOrd statistic, NNI Besag and Newel statistic, k-functions, spatial scan statistic</td>
<td>[1, 2, 7, 8, 12, 16, 23–41, 44–49, 51–54, 57–72, 75, 93–95, 99, 100, 102–176]</td>
</tr>
<tr>
<td>Monitoring spatiotemporal TB trends</td>
<td>Temporal trend maps</td>
<td>[27, 36–39]</td>
</tr>
<tr>
<td>Intervention evaluation</td>
<td>Distance map, kernel density map</td>
<td>[73, 74]</td>
</tr>
<tr>
<td>Barriers to TB care</td>
<td>Rate map, dot map, travel time map, distance map</td>
<td>[12, 187]</td>
</tr>
<tr>
<td>TB program performance</td>
<td>Map (time to detection)</td>
<td>[184]</td>
</tr>
<tr>
<td>HIV-related TB incidence</td>
<td>Rate map, dot map, spatial scan statistic</td>
<td>[40, 166, 186, 190]</td>
</tr>
<tr>
<td>TB treatment outcomes</td>
<td>Spatial empirical Bayes smoothing, kernel density maps, spatial scan statistic, spatial regression</td>
<td>[152, 155, 179, 183, 191]</td>
</tr>
<tr>
<td>Mortality related to TB/HIV coinfection</td>
<td>Rate map, thematic maps, Moran’s I and spatial regression</td>
<td>[42, 43, 174, 192]</td>
</tr>
<tr>
<td>Transmission</td>
<td>Dot maps (congregate settings)</td>
<td>[54, 55, 193]</td>
</tr>
<tr>
<td></td>
<td>Dot maps (cases)</td>
<td>[7, 8]</td>
</tr>
<tr>
<td></td>
<td>Geospatial and genotypic clustering methods</td>
<td>[1, 2, 25, 28, 47, 57, 59–72, 93–95, 169, 194]</td>
</tr>
<tr>
<td>Methodological</td>
<td>Spatial scan statistic</td>
<td>[25]</td>
</tr>
<tr>
<td>TB outbreak detection</td>
<td>Spatial scan statistic</td>
<td>[1, 25, 28]</td>
</tr>
<tr>
<td>Prevalence estimation</td>
<td>Model-based geostatistics</td>
<td>[80]</td>
</tr>
<tr>
<td>Drivers of MDR-TB</td>
<td>k-function</td>
<td>[35]</td>
</tr>
</tbody>
</table>

NNI nearest neighbourhood index, CAR models conditional autoregressive models, GWR geographically weighted regression, PCA principal component analysis, HIV human immunodeficiency virus, MDR-TB multidrug-resistant TB
Methods in the spatial analysis of TB

Table 2 shows the range of spatial methods used. Spatial analysis was used to visualise patterns (n = 154), explore spatial clusters (n = 73) and identify risk factors for clustering (n = 64), with risk prediction undertaken by 11 studies. Of the included studies, six did not explicitly report any of these methods but reported statistical results that implied the use of these methods.

Data visualisation

Data visualisation was the most consistently applied technique, with 154 of the studies using at least one data visualisation method to present TB distribution and/or risk factor patterns across space (Table 1). The TB incidence rate was the commonest indicator mapped (n = 63), followed by event maps (n = 37), which were smoothed using kernel density in seven studies. Data visualisation was based on standardised morbidity ratios (SMR) in 12 studies. Five studies reported maps of trends in TB incidence over time, and thematic maps were used in nine to consider the impact of risk factors on TB incidence by displaying the spatial distribution of other variables. Variables plotted included climate (n = 1), socioeconomic factors (n = 5), diabetes (n = 1) and obesity (n = 1).

Approaches used to account for data sparseness

TB is a relatively rare disease at the population level, and burden is typically expressed in terms of cases per 100,000 population. Various approaches were used to account for this sparseness in the number of cases, such as aggregating cases over administrative geographic levels and over time periods (ranging from 1 to 25 years).

An alternative approach was rate smoothing, although this practice was rare, despite the fact that TB rates were the commonest indicators mapped. In the included

<table>
<thead>
<tr>
<th>Method category</th>
<th>Method</th>
<th>Number</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visualisation</td>
<td>Rate map</td>
<td>63</td>
<td>[12, 16, 23, 26, 27, 29–34, 37, 41, 44–46, 48, 51, 52, 57, 58, 60, 61, 70, 100, 102, 103, 105, 106, 120, 123–146, 164, 165, 170, 173–176, 195, 196]</td>
</tr>
<tr>
<td></td>
<td>Dot map</td>
<td>37</td>
<td>[2, 7, 8, 35, 40, 47, 53, 54, 59, 66, 67, 72, 73, 75, 95, 107–122, 158, 166, 169, 178, 191, 197]</td>
</tr>
<tr>
<td></td>
<td>SMR map</td>
<td>12</td>
<td>[38, 49, 99, 100, 124, 126, 127, 129, 138, 142, 148, 149]</td>
</tr>
<tr>
<td></td>
<td>Kernel density map</td>
<td>7</td>
<td>[35, 37, 62, 93, 120, 147, 171]</td>
</tr>
<tr>
<td></td>
<td>Case counts maps</td>
<td>3</td>
<td>[108, 167, 172]</td>
</tr>
<tr>
<td>Spatial cluster analysis</td>
<td>Global Moran’s I</td>
<td>28</td>
<td>[16, 26, 34, 37, 39, 44, 48, 49, 51, 58, 65, 93, 100, 102, 123, 126, 128, 131, 133, 135, 138, 139, 145, 150, 161, 180, 188, 199]</td>
</tr>
<tr>
<td></td>
<td>Local Moran’s I</td>
<td>14</td>
<td>[16, 41, 44, 49, 51, 93, 100, 123, 126, 131, 135, 138, 145, 192]</td>
</tr>
<tr>
<td></td>
<td>Kulldorff’s spatial scan statistic</td>
<td>43</td>
<td>[1, 2, 23–32, 40, 57, 63, 64, 70, 71, 94, 109–111, 119, 120, 130, 135, 138, 139, 141, 151–160, 163, 164, 166, 191]</td>
</tr>
<tr>
<td></td>
<td>GetisOrd statistic</td>
<td>12</td>
<td>[2, 16, 26, 39, 49, 54, 65, 93, 104, 131, 139, 161]</td>
</tr>
<tr>
<td></td>
<td>k-NN</td>
<td>8</td>
<td>[35, 53, 69, 72, 93, 114, 122, 163]</td>
</tr>
<tr>
<td></td>
<td>k-function</td>
<td>6</td>
<td>[35, 62, 93, 116, 147]</td>
</tr>
<tr>
<td></td>
<td>Besag and Newell statistic</td>
<td>2</td>
<td>[125, 145]</td>
</tr>
<tr>
<td>Statistical modelling</td>
<td>Bayesian CAR models</td>
<td>7</td>
<td>[38, 44, 49, 99, 101, 127, 148]</td>
</tr>
<tr>
<td></td>
<td>Geographically weighted regression</td>
<td>6</td>
<td>[16, 50, 93, 102–104]</td>
</tr>
<tr>
<td></td>
<td>Mixture modelling</td>
<td>2</td>
<td>[142, 149]</td>
</tr>
<tr>
<td></td>
<td>Conventional logistic</td>
<td>15</td>
<td>[8, 40, 70, 71, 94, 95, 111, 112, 120, 141, 161, 177, 178, 187, 189]</td>
</tr>
<tr>
<td></td>
<td>Conventional Poisson</td>
<td>5</td>
<td>[46, 125, 136, 145, 156]</td>
</tr>
<tr>
<td></td>
<td>Conventional linear</td>
<td>5</td>
<td>[12, 47, 129, 137, 176]</td>
</tr>
<tr>
<td></td>
<td>Negative binomial</td>
<td>1</td>
<td>[164]</td>
</tr>
<tr>
<td></td>
<td>Factor analysis</td>
<td>6</td>
<td>[50, 103, 117, 143, 146, 170]</td>
</tr>
<tr>
<td></td>
<td>Regression models with spatial terms</td>
<td>9</td>
<td>[42, 48, 51, 58, 100, 116, 128, 131, 188]</td>
</tr>
<tr>
<td></td>
<td>Spatial prediction</td>
<td>11</td>
<td>[38, 42, 43, 62, 80, 99, 101, 127, 131, 148, 181]</td>
</tr>
</tbody>
</table>

SMR standardised morbidity ratio, k-NN k-nearest neighbourhood test, CAR conditional autoregressive

*Includes maps of disability-adjusted life years (DALYS), survival time, factor scores, probability maps, proportion of cases and regression coefficients
studies, smoothed rates were used in six (4%) studies. Similarly, of 12 studies that analysed SMRs, smoothed SMRs were presented in seven. In the included studies, several different data smoothing techniques were used, including fully Bayesian \((n = 8)\), empirical Bayes \((n = 4)\) and spatial empirical Bayes \((n = 5)\). A significant number of visualisation reports \((n = 30)\) were not complemented by hypothesis testing, either by exploration methods or modelling approaches. In 12 studies (7%), maps were not presented, but a narrative description of TB burden or a tabular presentation of TB distribution by administrative unit was described.

**Spatial cluster (hotspot) identification**

Use of at least one spatial cluster identification method was reported in 73 (43%) studies, with Kulldorff’s spatial scan statistic used most frequently \((n = 43)\), followed by Local Moran test \((n = 14)\) and Getis and Ord's local \(G_I\) statistic \((n = 12)\). Nearest neighbour index (NNI), \(k\)-function and Besag and Newell methods were reported in eight, six and two studies respectively (Table 1). The presence of overall area-wide heterogeneity was assessed most often using global Moran \(I\) \((n = 28)\). In three studies, no globally significant spatial autocorrelation was seen, although there was spatial clustering locally. Although studies used data aggregated over various spatial scales, only one evaluated the impact of spatial scale on the hotspot detection performance of the spatial scan statistic. Use of individual address-level data improved the sensitivity of the spatial scan statistic compared to data aggregated at the administrative level.

Simultaneous use of two spatial cluster detection methods was reported in 11 studies and showed differences in hotspot identification that ranged from complete disagreement to some degree of similarity (Table 3).

**False-positive clustering**

Not all spatial clusters are true clusters. False-positive clusters can arise from various sources, including data and methods used, and unmeasured confounding. Given that notification data were by far the most commonly used data source in the spatial analyses reviewed here, it could not be determined if these clusters represented true clusters of tuberculosis incidence or if they were caused by factors such as pockets of improved case detection. The role of differential TB detection has been documented in some studies from low-income settings, where increased spatial TB burden was linked to improved health care access [12].

In addition, rate was the commonest disease indicator used for disease mapping, as well as cluster detection in this study. As described earlier, rates are liable to stochasticity and can lead to false-positive clustering. However, rate smoothing and stability (sensitivity) analysis of clusters identified using rates was done in only a few studies [23, 24]. This remains an important area of consideration in the future spatial analysis of TB.

<table>
<thead>
<tr>
<th>Author, year</th>
<th>Methods</th>
<th>Outcome</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alene, K. 2017 [49]</td>
<td>Local Moran’s (I) Getis and Ord</td>
<td>Clustered</td>
<td>Clustered</td>
</tr>
<tr>
<td>Alvarez-Hernández, G., et al. 2010 [145]</td>
<td>Local Moran’s (I) Besag and Newell</td>
<td>No significant</td>
<td>Clustered</td>
</tr>
<tr>
<td>Dangisso M, et al. 2015 [26]</td>
<td>Getis and Ord Spatial scan statistic</td>
<td>Clustered</td>
<td>Clustered</td>
</tr>
<tr>
<td>Feske, M., et al. 2011 [93, 178]</td>
<td>Getis and Ord GWR residuals Spatial scan statistic</td>
<td>Clustered</td>
<td>Heterogeneous</td>
</tr>
<tr>
<td>Ge E, et al. 2016 [139]</td>
<td>Getis and Ord Spatial scan statistic</td>
<td>Clustered</td>
<td>Clustered</td>
</tr>
<tr>
<td>Hassanangsee S, et al. 2015 [138]</td>
<td>LISA Spatial scan statistic</td>
<td>Clustered</td>
<td>Clustered</td>
</tr>
<tr>
<td>Li L, et al. 2016 [135]</td>
<td>LISA Spatial scan statistic</td>
<td>No significant cluster</td>
<td>Clustered</td>
</tr>
<tr>
<td>Maceiel ELN, et al. 2010 [131]</td>
<td>LISA, Getis and Ord Model prediction</td>
<td>Clustered</td>
<td>Heterogeneous</td>
</tr>
</tbody>
</table>

GWR geographically weighted regression; LISA local indicators of spatial association
Spatiotemporal analysis

Temporal scale
In the spatial analysis of TB, the time window is an important dimension that influences the spatial pattern of TB [25]. As TB is relatively a rare disease at the population level and has a long incubation period, detection of apparent spatial clusters requires a longer time scale than for acute infectious diseases that may form spatial clusters within days of the start of outbreak. Because of this, the included studies were based on cases that accumulated over considerable time periods, ranging from 1 to 25 years, with use of data aggregated over 5 years being the most frequent practice (20%).

Approaches
Generally, two approaches were used in the space-time cluster analysis of TB. The first uses classical space-time clustering using algorithms which scan space over a changing time window, such as Kulldorff’s spatial scan statistic [23, 25–29]. The second approach is to account for the temporal dimension by repeating the spatial analysis for each time unit [26, 30–35]. In some studies, spatial patterns in temporal trends of TB incidence were determined as increasing or decreasing [27, 36–39].

Spatial statistical modelling
Different statistical modelling approaches were used to describe the relationship between TB and ecological factors in 65 (39%) studies, including nine spatially explicit models using Bayesian approaches. Conditional autoregressive (CAR) models were used in nine models to account for spatial correlation. Classical regression models were used in 33, while non-Bayesian spatial regression models were reported in 12.

Of the regression models that evaluated the effect on model fit of including spatial structure (spatial error or spatial lag), the inclusion of spatial structure improved the performance of the model in seven studies and failed to do so in two (based on deviance information criteria). Spatial lag was explicitly modelled in seven studies and highlighted the significant influence of neighbouring locations on TB distribution.

Traditional models including a Bayesian approach assumed a stationary relationship between TB and its spatial covariates and hence imposed a single (global) regression model on the entire study area. Only six studies used a geographically weighted regression (a local regression model) to accommodate variation in the association between TB and its risk factors from place to place and showed spatially varying (non-stationary) effects (n = 6). Other models used included mixture modelling (n = 2) and factor analysis using principal component analysis (PCA) (n = 4).

Results from spatial analysis

Geographic distribution of TB
The geographic distribution of TB was heterogeneous in all included studies both from low- and high-incidence settings, although no formal hypothesis testing was presented in 55 (33%). An exception was one study from South Africa that reported no significant clustering of cases among HIV patients on ART [40]. Spatial analysis was also used to describe the drivers of drug-resistant tuberculosis, with tighter spatial aggregation of MDR-TB cases compared with non-MDR cases taken as evidence of transmission of MDR-TB [41].

Spatial analyses into both HIV and TB investigated outcomes including HIV-associated TB incidence (n = 4) and spatial patterns of TB/HIV-related mortality (n = 4). All such studies revealed significant spatial heterogeneity. TB/HIV-related mortality in children was linked to areas with low socio-economic status and maternal deaths [42, 43].

Spatial methods used to study the impact of community-based TB treatment showed marked improvement in access compared to health facility-based treatment approaches (n = 1), and similar studies demonstrated travel time and distance to be important barriers to TB control (n = 2).

Correlations with social and environmental factors
The observed spatial patterns of TB were consistently linked to areas with poverty (n = 14), overcrowding and non-standard housing (n = 9), ethnic minority populations (n = 3), population density (n = 2), low education status (n = 2), health care access (n = 3) and immigrant populations (n = 5). However, a minority of studies have also found conflicting or non-significant associations between TB and poverty [44–46], population density [47–49] and unemployment [45, 47].

Four studies (including three from China) examined the correlation of climatic factors with TB incidence, with conflicting results. Two province-level studies in China using data from different time periods found TB burden to be associated with increasing annual average temperature [33, 50], although correlation with humidity was conflicting. Positive associations were observed with average precipitation [33, 50] and with air pressure [33] in these studies, while inverse associations were observed with sun exposure [50] and with wind speed [33]. In contrast, a county-level study which used average monthly climate data within a single province of China found the reverse, with temperature, precipitation, wind speed and sunshine exposure showing associations in the opposite direction [51]. A study that compared TB incidence between regions with different climatic conditions showed higher incidence at dry regions and low incidence in humid regions [52].
**Space-time analysis to detect TB outbreaks**

Studies reporting the application of the spatial methods in the early identification of TB outbreaks were uncommon. Space-time TB studies using retrospective surveillance data in the USA found that the spatial scan statistic and other methods could effectively detect outbreaks months before local public authorities became aware of the problem [25, 28]. However, as space-time clusters of TB can be due to either ongoing transmission or reactivation, characterising the drivers that resulted in the spatial clustering is essential. Findings from studies which compared the timeliness and accuracy of space-time clusters in identifying TB outbreaks varied with spatial resolution and the background population, with two studies from the USA detecting ongoing outbreaks [25, 28], in contrast to false alarms due to reactivation TB among immigrants in a study from Canada [1].

**Spatial analysis of the source of TB infection**

Spatial methods were also used to determine the role of households and congregate settings (e.g. social gathering venues, schools) on TB transmission risk (Table 1). The role of the household was determined by cross-referencing child and adolescent TB infection or disease with adult TB in two studies [7, 8]. In these studies, the importance of household exposure declined with the age of the child, such that TB disease or infection was related to residential exposure to adult TB in younger children but not adolescents.

Congregate settings, which pose increased transmission risk, were identified using multiple techniques that included linking TB cases to social gathering places [53] and mapping the distribution of rebreathed air volume (RAV) [54] (including grading these settings based on TB transmission principles [55]). These approaches identified schools and social gathering sites as high-risk areas.

**Identifying local drivers**

Recent transmission is a critical mechanism driving local TB epidemiology in high-burden settings, while reactivation of remotely acquired infection is thought to predominate in most low-endemic settings [4, 56]. Geospatial clusters may reflect increased disease risk due to geographic proximity, which may correspond to recent transmission, or reactivation of latent TB infection in an aggregate of individuals infected elsewhere or both [57]. In the reviewed studies, spatial methods coupled with other methods were used to identify which of these two mechanisms drives local TB epidemiology in the following three ways.

**Combining spatial clusters with cohort clustering:** TB clustering can occur from ongoing transmission or from reactivation of latent infection among high-risk subgroups due to shared characteristics such as similar country of birth rather than a shared transmission network, a phenomenon known as cohort clustering. Cohort cluster analysis is used to identify selected high-risk population subgroups for targeted interventions based on the relative TB incidence they bear. The Lorenz curve is a simple visualisation tool that compares the clustering (inequality) in the subgroup of interest across regions and over time. One study, which combined such cohort (birth country) cluster analysis using the Lorenz curve of inequality with spatial cluster analysis [31] revealed colocation of these cluster types, suggesting the presence of both transmission and reactivation. Spatial clusters among foreign-born persons covered too large an area compared to clusters among the locally born to be consistent with direct person-to-person transmission. In addition, spatial modelling was also applied to differentiate the role of transmission from reactivation by assessing spatial dependence. The presence of spatial dependence (autocorrelation) was taken to indicate transmission, while its absence was considered to indicate reactivation [58].

**Combining spatial and genotype clustering:** Genotypic clustering of TB may be used as a proxy for recent transmission, such that geospatial clusters in which cases are genotypically clustered may be taken as stronger evidence for locations where recent transmission has occurred. These approaches were combined to quantify the role of recent transmission and determine geographical locations of such transmission in 25 studies. This was done either by determining the spatial distribution of genotypic clusters [25, 28, 59–69] or by assessing the genotypic similarity of cases contained within geospatial clusters [2, 57, 65, 70, 71].

The findings from these studies varied considerably by the country and sub-population studied (locally born versus immigrants) (Table 4). Genotypic clusters were spatially clustered in many studies, providing evidence of recent local transmission. In some studies, cases in geospatial clusters were less likely to be dominated by genotypically similar cases (i.e. were dominated by unique strains) than cases outside the geospatial clusters, implying spatial aggregation of reactivation TB [57]. This finding highlights that geospatial hotspots in low TB incidence settings are not necessarily areas of recent transmission and spatial clustering may be primarily mediated by social determinants, such as migration, HIV and drug abuse [57].
Combinations of multiple methods were typically used for genotyping, with the commonest being IS6110 restriction fragment length polymorphism (IS6110-RFLP) and spoligotyping \((n = 9)\), followed by mycobacterial interspersed repetitive unit variable number tandem repeat (MIRU-VNTR) and spoligotyping \((n = 5)\), although use of a single method was reported in six studies (Table 4). No identified studies reported use of whole genome sequencing.

### Table 4 Overlap between spatial and molecular clustering

<table>
<thead>
<tr>
<th>Authors</th>
<th>Country</th>
<th>Genotyping methods</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bishai WR, et al. 1998 [95]</td>
<td>USA</td>
<td>IS6110-RFLP and PGRS</td>
<td>Genotypic clusters with epidemiologic links were spatially clustered but 76% of DNA clustered cases lack epidemiologic links.</td>
</tr>
<tr>
<td>Mathema B, et al. 2002 [169]</td>
<td>USA</td>
<td>IS6110-RFLP and spoligotyping</td>
<td>Genotypic clusters showed spatial aggregation</td>
</tr>
<tr>
<td>Richardson M, et al. 2002 [72]</td>
<td>South Africa</td>
<td>IS6110-RFLP and spoligotyping</td>
<td>Spatial aggregation of genotypic clusters was limited</td>
</tr>
<tr>
<td>Nguyen D, et al. 2003 [69]</td>
<td>Canada</td>
<td>IS6110-RFLP and spoligotyping</td>
<td>Genotypically similar cases were not more spatially clustered than genotypically unique cases</td>
</tr>
<tr>
<td>Moonan P, et al. 2004 [61]</td>
<td>USA</td>
<td>IS6110-RFLP and spoligotyping</td>
<td>Genotypic clusters were spatially heterogeneous</td>
</tr>
<tr>
<td>Jacobson L, et al. 2005 [59]</td>
<td>Mexico</td>
<td>IS6110-RFLP and spoligotyping</td>
<td>Spatial patterns were similar for both cases categorised as reactivation or recent transmission</td>
</tr>
<tr>
<td>Haase I, et al. 2007 [2]</td>
<td>Canada</td>
<td>IS6110-RFLP and spoligotyping</td>
<td>In spatial TB clusters of immigrants, there was significant genotype similarity</td>
</tr>
<tr>
<td>Feske ML, et al. 2011 [93, 178]</td>
<td>USA</td>
<td>IS6110-RFLP and spoligotyping</td>
<td>Genotypically clustered cases were randomly distributed across space</td>
</tr>
<tr>
<td>Evans JT, et al. 2011 [66]</td>
<td>UK</td>
<td>Spoligotyping and MIRU-VNTR</td>
<td>Genotypic clusters showed spatial aggregation</td>
</tr>
<tr>
<td>Nava-Aguilera E, et al. 2011 [67]</td>
<td>Mexico</td>
<td>Spoligotyping</td>
<td>Genotypic clusters were not spatially aggregated</td>
</tr>
<tr>
<td>Prussing C, et al. 2013 [57]</td>
<td>USA</td>
<td>Spoligotyping and 12-MIRU-VNTR</td>
<td>Cases in geospatial clusters were equally or less likely to share similar genotypes than cases outside geospatial clusters</td>
</tr>
<tr>
<td>Tuite AR, et al. 2013 [94]</td>
<td>Canada</td>
<td>Spoligotyping and 24-MIRU-VNTR</td>
<td>The proportion of cases in genotypic clusters was five times that seen in spatial clusters (23% vs 5%)</td>
</tr>
<tr>
<td>Kammerer JS, et al. 2013 [28]</td>
<td>USA</td>
<td>Spoligotyping and 12-MIRU-VNTR</td>
<td>Genotypically similar cases were spatially clustered</td>
</tr>
<tr>
<td>Verma A, et al. 2014 [1]</td>
<td>Canada</td>
<td>IS6110-RFLP and Spoligotyping</td>
<td>Space-time clusters contained few or no genotypically similar cases</td>
</tr>
<tr>
<td>Izumi K, et al. 2015 [65]</td>
<td>Japan</td>
<td>IS6110-RFLP</td>
<td>Both genotypically similar and unique strains formed spatial hotspots</td>
</tr>
<tr>
<td>Charnie G, et al. 2015 [194]</td>
<td>Uganda</td>
<td>Spoligotyping</td>
<td>Genotypic clusters shared social gathering sites (clinic, place of worship, market or bar)</td>
</tr>
<tr>
<td>Chan-Yeung M, et al. 2005 [47]</td>
<td>Hong Kong</td>
<td>IS6110-RFLP</td>
<td>Spatial locations of genotypic clusters and unique cases did not differ by their sociodemographic characteristics</td>
</tr>
<tr>
<td>Gurjavan U, et al. 2016 [70]</td>
<td>Australia</td>
<td>24-MIRU-VNTR</td>
<td>Spatial hotspots were characterised by a high proportion of unique strains; less than 4% of cases in spatial clusters were genotypically similar</td>
</tr>
<tr>
<td>Ribeiro FK, et al. 2016 [62]</td>
<td>Brazil</td>
<td>IS6110-RFLP and Spoligotyping</td>
<td>Genotypic clusters were spatially clustered</td>
</tr>
<tr>
<td>Seraphin MN, et al. 2016 [64]</td>
<td>USA</td>
<td>Spoligotyping and 24-MIRU-VNTR</td>
<td>22% of cases among USA-born and 3% among foreign-born clustered spatially and genotypically</td>
</tr>
<tr>
<td>Yuen CM, et al. 2016 [68]</td>
<td>USA</td>
<td>Spoligotyping and 24-MIRU-VNTR</td>
<td>Genotype clustered cases were spatially heterogeneous</td>
</tr>
<tr>
<td>Yeboah-Manu D, et al. 2016 [63]</td>
<td>Ghana</td>
<td>IS6110 and rpoB PCR</td>
<td>Genotypic clusters showed spatial aggregation</td>
</tr>
<tr>
<td>Zelner J, et al. 2016 [60]</td>
<td>Peru</td>
<td>24-MIRU-VNTR</td>
<td>Genotypic clusters showed spatial aggregation</td>
</tr>
</tbody>
</table>

**PGRS** polymorphic GC-rich repetitive sequence

**Temporal distribution of genotypically clustered cases**

The temporal pattern of genotypic clustering could provide insights to distinguish between transmission and reactivation. In some studies, the temporal distribution of genotypically clustered cases indicated periods of 1 to more than 8 years between the genotypically clustered cases [1, 72], implying reactivation TB could also show genotypic similarity.
Use of spatial methods to inform public health interventions

In addition to their use in characterising the spatial distribution and determinants of TB, spatial methods have been used to inform TB-related public health interventions. In these studies, spatial analysis methods have proved to be attractive in guiding public health interventions, although their application to TB care beyond research is not well documented. For instance, spatial analysis techniques have been used to identify locations with a high density of TB cases (termed hotspots, although this definition was not based on spatial statistical tests). Community screening was then conducted in these areas, and its yield was compared to that from routine service provision. This GIS-guided screening was found to considerably improve the detection of individuals with latent TB infection and other infectious diseases [73]. Similarly, a study from South Africa highlighted the potential for using GIS to promote community-based DOTS by locating and geographically linking TB patients to their nearest supervision sites, although programmatic implementation of this approach was not reported [74].

The potential for spatial methods to be used for the early detection of TB outbreaks has also been described, although the findings widely varied based on the background population [1, 28]. Spatial cluster analysis using data at higher geographic resolutions improves the method’s performance in cluster detection [25].

Discussion

While a range of methodologies has been employed in divergent contexts, we found that essentially all geospatial studies of TB have demonstrated significant heterogeneity in spatial distribution. Spatial analysis was applied to improve understanding of a range of TB-related issues, including the distribution and determinants of TB, the mechanisms driving the local TB epidemiology, the effect of interventions and the barriers to TB service uptake. Recently, geospatial methods have been combined with genotypic clustering techniques to understand the drivers of local TB epidemiology, although most such studies remain limited to low-endemic settings.

In almost all reviewed studies, retrospective program data (notifications) were used. Notification data, especially from resource-scarce settings, suffer from the often large proportion of undetected cases and are heavily dependent on the availability of diagnostic facilities [12]. None of the spatial studies of TB that used notification data accounted for undetected cases, such that the patterns in the spatial distribution and clustering could be heavily influenced by case detection performance [11]. Hence, distinguishing the true incidence pattern from the detection pattern has rarely been undertaken, despite its importance in interpretation.

The problems of undetected cases could be compounded in the spatial analysis of drug-resistant forms of TB, especially in resource-scarce settings where testing for drug-resistant TB is often additionally conditional on the individual’s risk factors for drug resistance [75]. However, recently, there have been some attempts to account for under-detection in the spatial analysis of TB. A Bayesian geospatial modelling approach presented a framework to estimate TB incidence and case detection rate for any spatial unit and identified previously unreported spatial areas of high burden [11]. Another approach is to estimate incidence using methods such as capture-recapture [76, 77] and mathematical modelling [78]. If case detection rate is truly known for a defined region, incidence can be calculated as notifications divided by case detection rate, although this is rarely if ever the case. Spatial analysis using prevalence data could also be considered in areas where such data are available.

In relation to the data problems outlined above, spatial analysis of TB could benefit from the use of model-based geostatistics, which is commonly used in other infectious diseases [79], although there are few studies that consider Mtb [80]. In particular, measurement of TB prevalence is impractical to perform at multiple locations due to logistic reasons. Therefore, model-based geostatistics can be used to predict disease prevalence in areas that have not been sampled from prevalence values at nearby locations at low or no cost, producing smooth continuous surface estimates.

Mapping of notification rates was the most commonly used data visualisation technique, in which TB cases were categorised at a particular administrative spatial level. This approach has the advantage of easy interpretability, although it can introduce bias because the size of the regions and the locations of their boundaries typically reflect administrative requirements, which may not reflect the spatial distribution of epidemiological factors [19, 22]. In addition, patterns observed across regions may depend on the spatial scale chosen, an effect known as the modifiable areal unit problem (MAUP) [17]. Because the choice of spatial scale mainly depends on the limitations of available data [81], only one study was able to provide a systematic evaluation of the effect of scale on spatial patterns, demonstrating improved performance of Kulldorff’s spatial scan statistic method at a high geographic resolution [25]. Different spatial resolutions could lead to markedly different results for the same dataset regardless of the true extent of correlation, due to averaging (aggregation effect) or other spatial processes operating at different scales.
Assessing the presence of this effect should be a priority for future studies using aggregated data in spatial TB studies.

Bayesian smoothing techniques can mitigate the problems of stochastically unstable rates from areas with small population [81], although such techniques were not widely used in the included studies and so false spatial clustering remains an important consideration. The less frequent use of rate smoothing techniques in the spatial analysis of TB could have various explanations, including lack of software packages that are easily accessible to the wider user (although GeoDa spatial software currently provides an accessible platform to people with limited statistical or mathematical backgrounds [82]). It may also be that most spatial analyses of TB are based on data aggregated over larger geographic areas from several years, such that the problem of statistical stochasticity may not be a major problem, although this was not explicitly discussed in the included studies.

In all studies that applied spatial cluster identification tools, TB cases were clustered irrespective of whether the setting was low or high endemic. However, in studies that incorporated more than one cluster identification method, areas identified as hotspots were not identical, with the extent of agreement between the alternative methods highly variable. This could be partly attributable to different methods testing separate hypotheses, such that these results may correctly support one hypothesis while refuting another. However, there is no consensus on how to interpret these findings appropriately and consistently [82, 83], and method selection did not typically appear to be based on such considerations [84, 85]. Thus, caution is required when considering interventions assessing clusters with one method only, as is frequently undertaken in TB spatial analysis [22].

Use of multiple cluster detection methods and requiring their overlap to represent a truly high-risk area is increasingly recommended [82, 84, 86]. However, this approach could also increase the risk of false-positive spatial clustering when different methods are used serially until significant clusters are observed [85]. Sensitivity analysis of spatial clustering [87, 88] and cluster validation using geostatistical simulations [23, 89, 90] can help identify robust clusters. While methods that adjust for confounding are generally preferred [91], further investigative strategies including data collection and cluster surveillance are required to validate an observed spatial cluster before introducing interventions [84, 85]. Although the focus of this study is TB, several methodological considerations outlined here would remain true for many infectious diseases.

In several studies, presence of spatial clustering or spatial autocorrelation in TB distribution was considered to reflect ongoing TB transmission, while its absence was taken to indicate reactivation [58]. Recently, molecular techniques have been combined with geospatial methods to understand the drivers of local TB epidemiology, although findings from these studies vary by country and the subset of the population studied. While spatial clustering of genotypically related cases was reported in several studies and likely reflected intense local TB transmission [61, 65], spatial clusters were dominated by genotypically unique strains in some studies, implying that reactivation was the dominant process [47, 72]. Hence, the combination of genotypic and geospatial techniques can improve understanding of the relative contribution of reactivation and transmission and other local contributors to burden.

Notwithstanding the general principles outlined above, not all spatial clusters of genotypically related cases will necessarily result from recent transmission, as simultaneous reactivation of remotely acquired infection and limited genetic variation in the pathogen population can also lead to genotypic similarity of spatially clustered cases [2, 92]. In some studies, the time between the first and last diagnosis of the cases in the genetic cluster ranged from 1 to more than 8 years [1, 72], suggesting that genotypic clustering could occur from spatially clustered reactivation. Similarly, limited spatial aggregation of genotypically clustered cases [72, 93, 94] and lack of epidemiological links between genotypically clustered cases in some studies may reflect migration of the human population over the extended time scale over which TB clusters occur [95], although casual transmission creating spatially diffuse clusters is an alternative explanation.

The extent of genotypic similarity between cases also depends on the discriminatory power of the genotyping method and the diversity of the pathogen population. Compared to whole genome sequencing, standard molecular genotyping (spoligotyping, MIRU-VNTR and IS6110) methods generally overestimate TB transmission with a false-positive clustering rate of 25 to 75% based on strain prevalence in the background population [92, 96]. The accuracy of these tests in distinguishing ongoing transmission from genetically closely related strains is very low among immigrants from high TB incidence settings with limited pathogen diversity [92, 97]. Thus, care should be taken when interpreting the genotypic similarity of cases among immigrant groups, as independent importation of closely related strains is possible. The frequent finding of more extensive genotypic than spatial clusters [71, 94] may reflect overestimation by the genotypic methods [98]. On the other hand, TB transmission might not result in apparent spatial clustering due to reasons that include population movement, poor surveillance and unmeasured confounding.
Regression models used for spatial analysis of TB were either conventional regression models or models that incorporated spatial effects. Although the former was more commonly employed, the majority of models incorporating spatial effects confirmed that accounting for spatial correlation improved model fit [11, 33, 44, 58, 99–101]. Conventional regression models assume spatial independence of model residuals and so ignore the potential presence of spatial autocorrelation, such that non-spatial models may lead to false conclusions regarding covariate effects.

The use of the conventional regression models described above may be appropriate for spatial analysis and spatial prediction, in the case that spatial dependence in residuals has been ruled out. Under this approach, the standard procedure is to start with classical ordinary least squares (OLS) regression models and then look for spatial dependence in the residuals, which implies the need for a spatially explicit regression model [82]. Several of the models reviewed here did not appear to adopt this approach, and so, caution is required when interpreting the findings from such analyses.

Most regression models treat the association between TB rates and ecological factors as global and are unable to capture local variation in the estimates of the association. However, geographically weighted regression (GWR) estimates coefficients for all spatial units included [22] and has often found the effect of risk factors on TB incidence to be spatially variable [16, 102–104], implying that global models may be inadequate to consider locally appropriate interventions. Few studies were able to perform explicit Bayesian spatial modelling incorporating information from nearby locations, thereby producing stable and robust estimates for areas with small populations and robust estimates of the effects of covariates [91].

While our review focused on methodological issues, several consistent observations were noted. Most importantly, all studies included in this review demonstrated that TB displayed a heterogeneous spatial pattern across various geographic resolutions. This reflects the underlying tendency for spatial dependence that can be caused by person-to-person transmission, socio-economic aggregation [49] and environmental effects [58, 93]. However, in nearly all included studies, spatial analyses of TB were based on the individual’s residence, although considerable TB infection is acquired from workplaces and other social gathering sites [8, 54]. Such studies could wrongly attribute TB acquired from such sites to residential exposure, leading to resource misallocation.

Several models have shown significant associations between TB rates and demographic, socioeconomic and risk-factor variables, although it is difficult to rule out publication bias favouring studies with positive findings. However, associations observed between TB rates and different factors such as population density, unemployment and poverty at the population level varied across studies. These were recognised as important individual-level risk factors, highlighting the potential for ecological fallacy.

We did not perform individual study level analysis of bias in this review. Analyses in the reviewed studies involved counts and proportions across different spatial distributions, rather than comparisons across different treatment/exposure groups. Standard tools of bias analysis predominantly focus on different treatment groups within cohorts (absent from our included studies) and hence are not applicable to this review. We have however discussed many potential sources of bias in the studies included in our review.

Most of the reviewed studies were from high-income settings, which may either reflect publication bias or a focus of research efforts on such settings. In high-incidence settings, the more limited use of spatial analysis methods could reflect a lack of access to resources (e.g. georeferenced data and spatial software packages) or insufficient expertise in these settings. However, it is these high-transmission settings which stand to gain the most from an improved understanding of TB spatial patterns and also these settings in which geospatial clustering may be most important epidemiologically.

Conclusions
A range of spatial analysis methodologies have been employed in divergent contexts, with virtually all studies demonstrating significant heterogeneity in spatial TB distribution regardless of geographic resolution. Various spatial cluster detection methods are available, although there is no consensus on how to interpret the considerable inconsistencies in the outputs of these methods applied to the same dataset. Further studies are needed to determine the optimal method for each context and research question and should also account for unreported cases when using notifications as input data where possible. Combining genotypic and geospatial techniques with epidemiologically linkage of cases has the potential to improve understanding of TB transmission.

Appendix
Search strings
Search terms used in Embase, Medline, PsycInfo, Scopus and Web of Science
The exp refers to explode which means include all subheadings underneath spatial analysis. When exploded, it contains geographic mapping, spatial regression and spatiotemporal analysis.

Brackets () denote subject headings (MeSH in Medline and Emtree in Embase) terms highlighted by the database.
Medline and PsycInfo

1. (exp spatial analysis) OR (Geographic information systems) OR (Space-time clustering) OR geographic* analys*.mp OR spat*regres*.mp OR spat*temp*.mp OR spat* analys*.mp OR spat* temp* analys*.mp OR spat* temp* pattern*.mp OR geography* distribut*.mp OR spat* temp* distribut*.mp OR heterogen* distribut.mp OR spacetime cluster*mp OR space-time cluster*mp OR hotspot.mp OR hot spots. mp OR GIS OR spat*

2. (tuberculosis) OR (tuberculosis, multidrug resistant) OR TB.mp

3. 1 AND 2

Embase

1. (spatial analysis) OR (geographic mapping) OR (spatial regression) OR (Spatio-temporal analysis OR (spatial autocorrelation analysis) OR (geography) OR (geographic distribution) OR (geographically weighted regression) OR (geographic information systems) OR (cluster analysis) OR geographic* analys*.mp OR spat*regres*.mp OR spat*temp*.mp OR spat* analys*.mp OR spat* temp* analys*.mp OR spat* temp* pattern*.mp OR geography* distribut*.mp OR spat* temp* distribut*.mp OR heterogen* distribut.mp OR spacetime cluster*mp OR space-time cluster*mp OR hotspot.mp OR hot spots. mp OR GIS OR spat*

2. (tuberculosis) OR (multidrug resistant tuberculosis) OR TB.mp

3. 1 AND 2

Scopus

(“Spatial analysis” OR “Spatial-temporal analysis” OR “Spatio-temporal analysis” OR “Geographic Information System” OR “Geographic Mapping” OR “geographic distribution” OR “spatial regression” OR “spatial autocorrelation analysis” OR “Spatiotemporal analysis” OR “geographic distribution” OR “hotspot” OR “hot spot” AND tuberculosis/TB

Web of science

[(Spatial analysis) OR (Spatio-temporal analysis) OR (Geographic Information System) OR (Geographic Mapping) OR (geographic distribution) OR (spatial regression) OR (spatial autocorrelation analysis) OR (Spatiotemporal analysis) OR (hotspot) OR (hot spot)] AND (Tuberculosis)

Additional file

Additional file 1: Figure S1. Trends in the spatial analysis of TB (note—the study included publications up to February 15, 2017). (DOCX 17 kb)

Abbreviations

CAR models: Conditional autoregressive models; GIS: Geographic information system; GWR: Geographically weighted regression; HIV: Human immunodeficiency virus; LISA: Local indicators of spatial association; NNI: Nearest neighbourhood index; PCA: Principal component analysis; TB: Tuberculosis

Acknowledgements

The authors are grateful to the University of Melbourne librarians for their extensive assistance in sourcing articles.

Funding

We did not receive funding for this study. Debebe Shaweno is the recipient of the Melbourne International Research Scholarship and Melbourne International Fee Remission Scholarship. James Trauer is a recipient of an Early Career Fellowship from the NHMRC (APP1142638).

Availability of data and materials

A list of included studies has been made available. The study protocol can be accessed on PROSPERO (CRD42016036655).

Authors’ contributions

DS and EM conceived the study, which was refined by JD and JT. DS developed data extraction checklist, and DS, MK and KA extracted the data. DS drafted the manuscript, and all authors provided input into revisions and approved the final draft for submission.

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Competing interests

The authors declare that they have no competing interests.

Publisher’s Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

1Department of Medicine, University of Melbourne, Melbourne, Victoria, Australia. 2Victorian Tuberculosis Program at the Peter Doherty Institute for Infection and Immunity, Melbourne, Victoria, Australia. 3Department of Microbiology and Immunology, University of Melbourne, Melbourne, Victoria, Australia. 4Research School of Population Health, College of Health and Medicine, The Australian National University, Canberra, Australia. 5Institute of Public Health, College of Medicine and Health Sciences, University of Gondar, Gondar, Ethiopia. 6Burnet Institute, Melbourne, Australia. 7Curtin University, Bentley, Western Australia, Australia. 8School of Public Health and Preventive Medicine, Monash University, Melbourne, Australia. 9Australian Institute of Tropical Health and Medicine, James Cook University, Townsville, Queensland, Australia.
References


A simulation study of three methods.


Maciel ELN, Pan W, Dietze R, Peres RL, Vinhas SA, Ribeiro FK, Palaci M, Rodrigues RR, Zandonade E, Golub JE. Spatial patterns of pulmonary...


150. Onozuka D, Hagiwara A. Geographic prediction of tuberculosis clusters in Fukushima, Japan, using the space-time scan statistic. BMC Infect Dis. 2007;7.


3.1 Supplementary Materials

3.1.1 Trends in the spatial analysis of TB

Figure S 1. Trends in the spatial analysis of TB

(Note - the study included publications up to February 15, 2017)
Chapter 4

A novel Bayesian geospatial method for estimating tuberculosis incidence reveals many missed TB cases in Ethiopia

Chapter overview

As described in the previous two Chapters (Chapters 2 and 3), spatial analysis of TB based on notification data can be misleading because notification data suffer from differential case detection rates depending on health care access. The previous Chapter (Chapter 3) highlighted the universal use of surveillance data in the spatial analysis of tuberculosis without accounting for missing cases. Having identified this issue, this Chapter now proposes a novel Bayesian geospatial approach to account for undetected cases, which uses surveillance data and other relevant covariates as inputs and provides estimates of incidence and case detection rate across space and time as outputs. Using this approach, this thesis is able to unpack the role of factors influencing notification, such as access to health care, and distinguish these from factors that influence the true underlying TB incidence. Applied to notification data collected from rural and remote Ethiopia, the model has identified previously unrecognised high burden areas. Beyond its application to TB, this work provides an important tool for the spatial analysis of any infectious disease displaying a considerable number of missing cases.
A novel Bayesian geospatial method for estimating tuberculosis incidence reveals many missed TB cases in Ethiopia

Debebe Shaweno 1,3*, James M. Trauer 1,2,3, Justin T. Denholm 3,4 and Emma S. McBryde 1,5

Abstract

Background: Reported tuberculosis (TB) incidence globally continues to be heavily influenced by expert opinion of case detection rates and ecological estimates of disease duration. Both approaches are recognised as having substantial variability and inaccuracy, leading to uncertainty in true TB incidence and other such derived statistics.

Methods: We developed Bayesian binomial mixture geospatial models to estimate TB incidence and case detection rate (CDR) in Ethiopia. In these models the underlying true incidence was formulated as a partially observed Markovian process following a mixed Poisson distribution and the detected (observed) TB cases as a binomial distribution, conditional on CDR and true incidence. The models use notification data from multiple areas over several years and account for the existence of undetected TB cases and variability in true underlying incidence and CDR. Deviance information criteria (DIC) were used to select the best performing model.

Results: A geospatial model was the best fitting approach. This model estimated that TB incidence in Sheka Zone increased from 198 (95% Credible Interval (CrI) 187, 233) per 100,000 population in 2010 to 232 (95% CrI 212, 253) per 100,000 population in 2014. The model revealed a wide discrepancy between the estimated incidence rate and notification rate, with the estimated incidence ranging from 1.4 (in 2014) to 1.7 (in 2010) times the notification rate (CDR of 71% and 60% respectively). Population density and TB incidence in neighbouring locations (spatial lag) predicted the underlying TB incidence, while health facility availability predicted higher CDR.

Conclusion: Our model estimated trends in underlying TB incidence while accounting for undetected cases and revealed significant discrepancies between incidence and notification rates in rural Ethiopia. This approach provides an alternative approach to estimating incidence, entirely independent of the methods involved in current estimates and is feasible to perform from routinely collected surveillance data.

Keywords: Tuberculosis, Incidence, Spatial analysis, Binomial mixture models

Background

Population level tuberculosis (TB) prevalence and incidence studies are resource and time-intensive, and so impractical for regular evaluation of TB trends in most settings. Almost universally, data acquired from routine programmatic TB notifications are adjusted by a number of methods, including expert opinion regarding case detection rates (CDR) in the local context, to estimate and report TB incidence [1]. In the few countries with recent and well-conducted prevalence surveys, incidence is calculated from a combination of these survey findings and estimates of the duration of disease, with the latter derived from the pre-chemotherapy era or from mathematical models [2, 3]. Because of uncertainties in the duration of disease, incidence estimates from these methods differ and optimal methods remain elusive [2]. As a result, significant uncertainty exists regarding the true TB incidence and other programatically important unobserved values, such as CDR [4].

Given these uncertainties, notification data with imputation of estimated missing cases have been relied on in many epidemiological TB studies, including in...
spatiotemporal characterization studies [5–8]. Failure to appropriately account for missed cases could reduce the power of such studies to detect factors affecting population-level TB dynamics, and introduce systematic bias of estimates of the effect of covariates towards the null [9]. Both factors could obscure important population-level patterns. Similarly, interpretation of maps from these studies is problematic, as spatial patterns in notification data could reflect spatial dynamics in the true underlying incidence or case detection performance. Thus these maps could be biased to areas with better case detection performance, leading to resource misallocation. One such cause of bias is systematic under-reporting which could also vary depending on the availability of health care facilities.

To address this bias, we aimed to estimate TB incidence and case detection in Ethiopia using an alternative Bayesian methods based on routinely collected surveillance data, without the need for expert opinion regarding programmatic performance or estimates of duration of disease.

Methods

We collected data on TB patients diagnosed between 2010 and 2014 in all 66 kebeles (the smallest geographical administrative unit in Ethiopia) of Sheka Zone (a remote zone in the country). These TB cases were pooled based on kebele of residence and year. We produced population density using census data and kebele shape files obtained from the Central Statistical Agency (CSA). Data on health facility availability were obtained from the Zonal health department. At the time of data collection, close to 50% of health facilities in the Zone did not have sputum microscopy, and there was no access to chest x-ray and culture facilities. Further details of the data and the study area are presented elsewhere [10].

Model development

Hidden Markov models (Bayesian geospatial binomial mixture models) were developed to estimate the true underlying but unknown TB incidence and case detection rates. Hidden Markov models (HMMs) are flexible time series models for sequences of observations that are known to be driven by an underlying state which is hidden from the observer yet is in some way predictable, such as being serially auto-correlated, geospatially correlated or through predictor variables that are observable. The value of HMMs is their ability to predict the underlying process; which they do provided the second process (the relationship between hidden and observed data) is one that itself follows rules that can be defined and in turn estimated.

Here, the underlying process is the factors that drive TB incidence in the study region. We have information including the population density, serial data and geospatial position of the kebele. The process that relates incidence to notification is the case detection rate or the proportion of cases that are notified. Given that (by definition) notification is the product of incidence and case detection rate, the natural model choice is the binomial relationship between the observation (notification) and hidden state (incidence). We also allow for kebele-specific effects on incidence, in effect leading to a (higher variance) beta-binomial distribution for observed notifications in each kebele.

Such a model when challenged by data can lead to issues of identifiability in that both high incidence/low case detection and low incidence/high case detection can explain the same notification rate. However, with sufficient information (e.g. predictors of incidence, changes over time and further information to assist in the observation model, such as presence of a health centre), the precision of model estimates can increase.

The models are informed by spatially and temporally replicated TB case counts and yield estimates of the true incidence and case detection. The parameters of this state process describe the spatiotemporal variation in incidence, which is considered as a latent variable by the model and is our key output of interest.

The state-space (the true underlying incidence) model

The number of incident TB cases (the latent state) in site \(i\) and year \(j\), conditional on the expected mean \(\lambda_{ij}\) is a realisation of a Poisson distribution, where the expected number \(\lambda_{ij}\) is a product of the per capita TB rate \(n_i\) (measured as a probability between 0 and 1) and the susceptible population size in year \(j\) at site \(i\):

\[
\text{Incidence}_{ij} \sim \text{Poisson} (\lambda_{ij})
\]

\[
\lambda_{ij} = \pi_{ij} \times \text{Population}_{ij},
\]

The site index \(i\) runs from 1 to 66, representing the 66 kebeles and the year index \(j\) runs from 1 to 5. The logit transformed probability of incident TB \((\pi_{ij})\) is, in turn, a logit-linear function of the site- and year-specific covariates where population density \((X_i)\), average incidence rate in kebeles that share a border with the index kebele \((Z_{ij})\), and logit transformed incidence rate at a temporal lag of one year \((\pi_{ij-1})\) with intercept \(\beta_0\) and slopes \(\beta_1\), \(\beta_2\) and \(\beta_3\) (eq. 3) were fitted as fixed effects. Extra-Poisson dispersion in the incidence is accounted by specifying two types of random effects: a spatially correlated random effect \((\epsilon)\) and a non-spatially correlated random effect \((\nu)\).

\[
\logit(\pi_{ij}) = \beta + \beta_1 X_i + \beta_2 Z_{ij} + \beta_3 \logit(\pi_{ij-1}) + \nu_{ij} + \epsilon_i
\]

Spatial dependence was introduced into regression in two ways: by introducing spatially structured random
error and spatial lag. The spatially structured random effect, \( \varepsilon = (\varepsilon_1, \ldots, \varepsilon_{64}) \), accounts for the effect of spatial proximity, with the prior distribution taken as a Gaussian Conditional Autoregressive function (CAR), in which the prior probability distribution of the value of \( \varepsilon_i \) has a mean equal to the weighted average of the neighbouring random effects [11, 12] and variance following an inverse gamma distribution (shape = 0.5, scale = 0.0005). Average TB notification rate in the kebeles adjacent to an index kebele was used to define spatial lag.

The case detection model

Notified TB cases have two sources of variation: variation in the true underlying incidence (Incidence\(_{ij}\)) and variation in the case detection process (P\(_{ij}\)). As a description of the detection process giving rise to detected (observed) cases, we assumed detected cases (Y\(_{ij}\)) by health facilities are realisations of a binomial process conditional on the underlying true incidence and case detection probability. The logit transformed case detection probability (P\(_{ij}\)) is a linear function of the covariate of health facility (H\(_i\)) availability.

\[
Y_{ij} \sim \text{Binomial} \left( P_{ij}, \text{Incidence}_{ij} \right),
\]

\[
\logit(P_{ij}) = \theta + \theta_1 H_i + \omega_{ij}
\]

Hence \( \theta_1 \) is the log(odds ratio) of detection in the presence of a health centre compared with no health centre. To account for additional heterogeneity in CDR not captured by this covariate, we fitted a normally distributed random effect that differed by kebele and year (\( \omega_{ij} \)). The prior distribution of this error term had a mean of zero and a standard deviation drawn from the uniform (0, 5) distribution.

A non-informative uniform prior distribution (−10, 10) was chosen for all regression coefficients and intercept parameters (other than those already specified) to express the absence of prior information about model parameters.

We used WinBUGS 1.4.3 to fit the model using Markov chain Monte Carlo (MCMC). MCMC is a simulation tool to draw large samples from the Bayesian posterior distribution of parameters that is typically analytically intractable [13]. The model was executed in R version 3.3.1 using the R2WinBUGS library. To enhance convergence of the MCMC sampler, we standardized population density by dividing the difference between each observation and the group mean by their respective standard deviations and also truncated the normal distributions for over dispersion effects to within (−16, 16) by multiplying with an indicator uniform prior distribution [14]. We ran the model for 250,000 iterations and discarded the first 50,000. We checked whether the priors were too restrictive, by inspecting a histogram of the posterior [15], with no such evidence found.

As we were executing two interdependent logit models, the model initially did not appear to find realistic parameter space for incidence rate and case detection rates (with estimated incidence rates far exceeding observed rates). Thus the model was forced into realistic parameter space [16] by restricting the maximum incidence rate to 1%, corresponding to 1000 per 100,000 per year (this is a realistic upper limit, being five times that of current notification rates and only reported in one country in the world in 2015).

Four candidate Bayesian geospatial models were considered:

- Model 1 Covariate only; logit(\( \pi_{ij} \)) = \( \beta + \beta_1 X_i + \beta_2 Z_{ij} + \beta_3 \logit(\pi_{ij-1}) \)
- Model 2 Covariates and non-spatially correlated random effect; logit(\( \pi_{ij} \)) = \( \beta + \beta_1 X_i + \beta_2 Z_{ij} + \beta_3 \logit(\pi_{ij-1}) + \nu_{ij} \)
- Model 3 Covariate and spatially correlated random effect; logit(\( \pi_{ij} \)) = \( \beta + \beta_1 X_i + \beta_2 Z_{ij} + \beta_3 \logit(\pi_{ij-1}) + \nu_{ij} + \epsilon_i \)
- Model 4 Full Model with covariates and all random effects included logit(\( \pi_{ij} \)) = \( \beta + \beta_1 X_i + \beta_2 Z_{ij} + \beta_3 \logit(\pi_{ij-1}) + \nu_{ij} + \epsilon_i \)

Covariates and random effects were included/excluded to determine if this had an effect on model fit and to determine the extent to which they accounted for the spatial correlation. The effect of spatially correlated random effects was assessed by examining the credible intervals (CrI) of the coefficients of the selected covariates and incidence and case detection estimates.

The deviance information criterion (DIC) statistic was calculated for the models with or without random effect terms to determine if the addition of the geospatial component improved model fit. Maps of the summary statistics of the posterior distributions of predicted incidence and the notified cases were constructed using R.

Goodness-of-fit

We conducted posterior predictive checks to evaluate whether the models considered could likely have generated datasets that are similar to our observed dataset. This procedure uses parameter values estimated by the model using observed data to generate simulated data sets. Chi-square fit statistic was calculated to quantify the lack of fit both for the observed data and for the simulated data sets, and a Bayesian P-value was calculated to quantify the ratio between the fit statistic for the observed data and that of the simulated (perfect datasets under model assumptions) (Additional file 1). A Bayesian P-value close to 0.5 indicates a model fits the data, while a P-value close to zero or one suggests poor fit [15, 17, 18].
**Results**

**Model selection**

We found that a binomial mixture model containing the spatially structured random effect and covariates was the best fitting model, and that the model with no random effects included was the poorest fitting. When a spatially structured random effect was included into the covariate model, the DIC dropped by more than 230 whereas inclusion of the non-spatially correlated random effect reduced the DIC by only 123 (Additional file 2).

**The role of the spatially correlated random effect**

As well as the marked reduction in the DIC, the coefficient for the temporal lag in the state-space model became non-significant while the credible intervals surrounding other coefficients widened, but remained significant, after inclusion of the geospatial component in the state-space model. Similarly, the inclusion of non-spatially structured random effects resulted in non-significant temporal lag. By contrast, the credible intervals surrounding the predicted case detection and incidence rates became more precise when the geospatial component was added to the model (Additional file 2). In the covariate only model, the coefficient for the spatial lag covariate was not statistically significant, but became significant with the inclusion of both random effects.

**Incidence and case detection rates**

Outputs from the best fitting model are presented in Table 1 and show the estimated TB incidence in Sheka Zone was 198 (95% CrI: 187, 233) per 100,000 per year in 2010 and 232 (95% CrI: 212, 253) per 100,000 per year in 2014.

The model demonstrated a wide discrepancy between the estimated incidence rate and the notification rate, with estimated incidence 1.4 (2014) to 1.7 (2010) times the notification rate, placing CDR at 60% (95% CrI: 49, 72) in 2010 and 71% (95% CrI: 60, 81) in 2014 (Fig. 1).

<table>
<thead>
<tr>
<th>Year</th>
<th>Median, 95% credible interval</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimated incidence</td>
</tr>
<tr>
<td>2010</td>
<td>198 (187, 233)</td>
</tr>
<tr>
<td>2011</td>
<td>218 (199, 238)</td>
</tr>
<tr>
<td>2012</td>
<td>216 (200, 234)</td>
</tr>
<tr>
<td>2013</td>
<td>219 (203, 236)</td>
</tr>
<tr>
<td>2014</td>
<td>232 (212, 253)</td>
</tr>
</tbody>
</table>

**Covariate effects on incidence and detectability**

Coefficients of the covariates included in the four candidate hierarchical binomial mixture models are presented in the Additional file and the coefficients for the best fitting model are presented in Table 2. Our methodological approach enabled us to present predictors of incidence and CDR separately.

As shown in Table 2, TB incidence rate was positively associated with population density and spatial lag in the geographically adjacent sites. An increase of 10/100,000/year in average TB incidence in adjacent kebeles predicted a 5.0/100,000/year increment in TB incidence in an index kebele. Similarly, an increase of population size by 10 per square kilometre predicted an increase in TB incidence by 1/100,000/year. Population density remained a significant predictor of TB incidence in all candidate models (Additional file 2). All the models except the covariate only model demonstrated the statistically significant effect of incidence at neighbouring locations. In the best fitting model, TB incidence was not significantly related to incidence at a temporal lag of one year (Table 2).

On the other hand, CDR was related to the presence of a health facility in the best fitting model, as well as in all other candidate models considered in this study. The estimated case detection rate in kebeles with no health facility was 59% (95% CrI: 55, 63), while the rate was 69% (95% CrI: 62, 77) in kebeles with a health facility.

**Spatial distribution**

Maps of the spatial distribution of estimated incidence and notification rates of TB presented in Fig. 2 revealed the presence of undetected TB cases at kebele level in Sheka Zone. The patterns observed in the maps of incidence and notifications appeared broadly correlated. However, the incidence map shows areas of considerably greater burden and identifies new, previously unrecognised areas of high burden. The incidence map identified many rural kebeles without a health facility and urban kebeles as high burden locations, in contrast to the notification map that identified mainly urban kebeles. These locations corresponded to kebeles with high population density and were surrounded by high incidence kebeles. The maps illustrate that estimated incidence rates are higher than notification rates, highlighting that notification data markedly underestimate incidence.

**Goodness of fit of the model**

We calculated a Bayesian $P$-value by comparing a Pearson chi-square for both the simulated datasets and the actual dataset. The proportion of times the discrepancy measure for simulated data is greater than the actual data (a Bayesian $P$-value) from a posterior predictive check was 0.40. The Bayesian P-value suggests that the model fits the observed data satisfactorily [15, 18]. As an internal validity check, we simulated using posterior
parameter values whether our model could reproduce the notification data that were previously used to estimate the model parameters. Comparison of notification data with simulated data indicated a close fit (Fig. 3). In addition, comparison of priors with their posteriors also indicated that our estimates were entirely data-driven (Fig. 4).

**Discussion**

The Bayesian hidden Markov model approach used in this study provided a means to estimate both TB incidence and case detection rate, and identified previously unrecognised TB hotspots in rural Ethiopia. To our knowledge, this is the first report applying spatially explicit binomial mixture methods to estimate TB incidence and case detection using only case notifications distributed over space and time.

By using Bayesian inference, we were able to incorporate the spatial dependence structure, random effects and a hidden state (true TB incidence) into our model. The improved fit achieved by including a geospatial term highlights the importance of accounting for spatial correlation in TB studies [8]. In this study, incorporation of the spatial dependence structure changed estimated relationships between covariates and incidence: some relationships changed from significant to non-significant, while others remained significant but with reduced precision. Like others [19, 20], we conclude that failing to account for spatial effects may lead to inflated regression coefficients and spuriously narrow credible intervals.

In contrast to previous analyses that described patterns in the notified cases (as a proxy for incidence) [7, 8], our model explicitly separated the true incidence (hidden

---

**Table 2** Predictors of incidence, and case detection rate in the binomial mixture model

<table>
<thead>
<tr>
<th></th>
<th>Coefficient, median (95% CrI)</th>
<th>Odds*, Odds ratio (95% CrI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>State-space model</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_0$ (intercept)</td>
<td>−5.5 (−6.1, −4.9)</td>
<td>0.004 (0.002, 0.007)*</td>
</tr>
<tr>
<td>$\beta_1$ (population density)$^a$</td>
<td>0.1 (0.1, 0.14)</td>
<td>1.1 (1.1,1.2)</td>
</tr>
<tr>
<td>$\beta_2$ (temporal lag)$^b$</td>
<td>0.003 (−0.1, 0.1)</td>
<td>1.0 (0.9, 1.1)</td>
</tr>
<tr>
<td>$\beta_3$ (spatial lag)$^b$</td>
<td>0.5 (0.3, 0.7)</td>
<td>1.6 (1.4, 2.0)</td>
</tr>
<tr>
<td><strong>Case detection model</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CDR in kebeles with no health facility</td>
<td>0.59 (0.55, 0.63)</td>
<td></td>
</tr>
<tr>
<td>CDR in kebeles with a health facility</td>
<td>0.69 (0.62, 0.77)</td>
<td></td>
</tr>
</tbody>
</table>

*a Population per square km, $^b$: number of incident TB cases/100,000 population/year
CDR- case detection rate
state) and observation process (case detection). A high or increasing notification rate could reflect either an efficient health system rapidly diagnosing all incident cases or a poor health system failing to detect patients quickly enough to gain control of the epidemic. As these two phenomena lie at opposite poles of programmatic TB control and would be associated with profoundly different case detection rates, it is critical to distinguish them objectively and without the need for opinion-based estimates of case detection [21].

Applied to five years of TB surveillance data in Sheka Zone Ethiopia, our model demonstrated a wide discrepancy between the estimated incidence rate and notification rate in areas with no health centres, which is consistent with WHO estimates [22, 23], highlighting the importance of strengthening surveillance systems to reduce missed cases [23, 24]. Unlike the WHO estimates, however, we estimated that the incidence of tuberculosis is increasing, and that increased notification rates reflect both increased detection and increased incidence.
The estimated TB incidence in this study is highly spatially heterogeneous, replicating previous reports [10, 20, 25], and associated with average incidence in the neighbouring kebeles (spatial lag) and population density, reflecting an extended duration of contact between individuals as a driver of TB epidemiology in high-burden settings [26, 27]. In contrast, TB incidence at a temporal lag of one year in this study failed to replicate the statistically significant association observed in the previous analysis using GLMs [10].

In our study, health facility availability predicted high TB case detection. Higher notification rates in some of the studied kebeles were attributable to both underlying higher incidence and higher case detection in the setting of health facility availability.

In contrast to methods used by WHO to estimate TB incidence, our approach uses only routinely collected surveillance data and does not incorporate costly prevalence surveys and expert opinion regarding CDR, which is criticized for its insensitivity to recent changes and the potential for bias [1]. Expert opinion is a potentially important strategy to use in a Bayesian models where information is lacking, but has a recognised limitations in accurate adjustment of disease estimates [3] and is widely considered the lowest standard of empirical evidence [28]. In addition, it does not exist at fine geographical levels, such as the one we are assessing here, and so is impractical to use in this context. Hence, our approach uses expert opinion only for the purpose of developing a vague prior probability distribution. We aim to improve on expert opinion through the use of the structured hidden Markov Model and so to develop a feasible alternative approach that can be used for regular monitoring of TB incidence and is reactive to recent changes.

Our case detection model assumes that individual TB cases are detected at a fixed rate and independently (conditional on a given incidence). This assumption may not be valid in the case of concerted efforts at contact tracing. However, as contact tracing is not systematically implemented in Ethiopia, the proportion of cases in our study arising from contact tracing would be small and this assumption would be a minor consideration. Moreover, we included a random effect term ($\omega_{ij}$) in our model to allow for changes in detection rate over place and time. Sparse data were accounted for by including random effects to explain extra-model variability. Further work is in progress to build models that account for over dispersion, which may arise from non-independent detections of individuals.

**Conclusions**

By using Binomial mixture models, we were able to investigate different epidemiological questions related to the size, trend and predictors of TB incidence and CDR. Our model demonstrated a wide discrepancy between incidence rate and notification rates, and identified previously unrecognised TB hotspots in rural Ethiopia, but was broadly consistent with official estimates. This approach provides an alternative approach to estimating incidence, entirely independent of the methods involved in current case detection rate estimates and is feasible to perform from routinely collected surveillance data.
Additional files

Additional file 1: Model code. (DOCX 14 kb)
Additional file 2: Outputs from Candidate Models. (DOCX 21 kb)

Abbreviations
CDR: Case detection rate; CrI: Credible Interval; DIC: Deviance information criteria; TB: Tuberculosis

Acknowledgements
The authors thank Sheka Zone Health Department, Ethiopia, and all health centres for permission to access the data; and the Ethiopian Central Statistical Agency, Ethiopia, for the provision of kebele polygons. The authors are grateful to Mr Tamrat Shaweno, Mr Elesa Moges and Mr Sewnet Dasho for data collection and coordination.

Availability of data and materials
The spatial data that support the findings of this study were made available by the Central Statistical Agency (CSA) of Ethiopia, and so are not publicly available. The code required to run the above analysis using R2WinBUGS library is provided as an additional file.

Authors’ contributions
DS and EM wrote the code in WinBUGS, DS drafted the initial study concept and EM, JD & JT refined this further. DS wrote the initial draft of this manuscript, and all authors provided input into revisions and approved the final draft and submission for publication.

Ethical approval and consent to participate
The study was approved by the Melbourne University Health Sciences Human Ethics Subcommittee, Melbourne, VIC, Australia, and the Zonal Health Department, Sheka Zone, Masha, Ethiopia.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References
4.1 Supplementary Materials

4.1.1 Model code

model {
    # Priors
    tau.p <- 1/(sd.p * sd.p)
    sd.p ~ dunif(0, 5)
    beta0 ~ dunif(-10, 10)
    beta1 ~ dunif(-10, 10)
    beta2 ~ dunif(-10, 10)
    beta3 ~ dunif(-10, 10)
    theta1 ~ dunif(-15, 15)
    ss[1:N] ~ car.normal(adj[], weights[1:sumNumNeigh], num[], tau.s);
    tau.s ~ dgamma(0.5, 0.0005);
    for (k in 1: sumNumNeigh) {
        weights[k] <- 1
    }
    for (i in 1 : N) {
        for (j in 1 : 5) {
            v[i, j] ~ dnorm(0, tau.p) I(-10, 10);
            beta[i, j] <- beta0
        }
    }
    # Likelihood
    for (j in 1:1) {
        for (i in 1 : N) {
            # True State
            Incidence[i, j] ~ dpois(lambda[i, j]);
            lambda[i, j] <- m[i, j] * Pop[i, j];  # m is probability of infection
            logit(m[i, j]) <- beta + beta1*X[i] + beta3*Z[i, j] + ε[i] + ν[i, j]
            # Observation Process
            y[i, j] ~ dbin(p[i, j], Incidence[i, j]);
            lp[i, j] <- theta + theta1*Xhc[i] + ω[i, j]
            theta0[i, j] ~ dunif(-10, 10)
        }
    }
}

p[i, j] <- exp(lp[i, j])/(1 + exp(lp[i, j]))
}
}
}
}
for (j in 2:5) {
for (i in 1:N) {
# True State
Incidence[i, j] ~ dpois(lambda[i, j]);
lambda[i, j] <- m[i, j]*Pop[i, j]
logit(m[i, j]) <- beta + beta1*X[i] + beta2*Z[i, j] + beta3*logit(m[i, j-1]) + e[i] +v[i,j]
# Observation Process
y[i, j] ~ dbin(p[i, j], Incidence[i, j]);
lp[i, j] <- theta + theta1*Xhc[i] + o[i, j]
p[i, j] <- exp(lp[i, j])/(1 + exp(lp[i, j]))
theta0[i, j] ~ dunif(-10, 10)
}
}
# Model Goodness of fit
# Posterior predictive distributions of \chi^2 discrepancy: assess model fit using \chi^2 discrepancy
for (j in 1:5) {
for (i in 1:N) {
# Compute \chi^2 statistic for observed data
Eval[i, j] <- p[i, j] * Incidence[i, j];  # Expected notification using model parameters
E[i, j] <- pow((y[i, j] - Eval[i, j]), 2) / (Eval[i, j] + 0.5);
y.new[i, j] ~ dbin(p[i, j], Incidence[i, j]);  # generate notification using the model
E.new[i, j] <- pow((y.new[i, j] - eval[i, j]), 2) / (eval[i, j] + 0.5)
}
}
fit <- sum(E[, ])
fit.new <- sum(E.new[, ])
bpv <- step(fit.new - fit);  # Bayesian p-value
}
" , fill = TRUE)
### 4.1.2 Outputs from Candidate Models

Table S1. Incidence and case detection estimates from candidate models

<table>
<thead>
<tr>
<th>Year</th>
<th>Incidence, mean, 95% CrI</th>
<th>Case detection rate, 95% CrI</th>
<th>DIC</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Model 1: Covariate only model (no random effects)</strong></td>
<td></td>
<td></td>
<td>762</td>
</tr>
<tr>
<td>2010</td>
<td>246 (202, 301)</td>
<td>0.508 (0.421, 0.593)</td>
<td></td>
</tr>
<tr>
<td>2011</td>
<td>342 (299, 392)</td>
<td>0.454 (0.388, 0.527)</td>
<td></td>
</tr>
<tr>
<td>2012</td>
<td>332 (297, 371)</td>
<td>0.412 (0.388, 0.488)</td>
<td></td>
</tr>
<tr>
<td>2013</td>
<td>332 (297, 372)</td>
<td>0.519 (0.451, 0.590)</td>
<td></td>
</tr>
<tr>
<td>2014</td>
<td>334 (299, 376)</td>
<td>0.536 (0.465, 0.608)</td>
<td></td>
</tr>
<tr>
<td><strong>Model 2: Covariate and non-spatially structured random effects model</strong></td>
<td></td>
<td></td>
<td>649</td>
</tr>
<tr>
<td>2010</td>
<td>215 (159, 289)</td>
<td>0.659 (0.545, 0.755)</td>
<td></td>
</tr>
<tr>
<td>2011</td>
<td>223 (178, 278)</td>
<td>0.692 (0.559, 0.794)</td>
<td></td>
</tr>
<tr>
<td>2012</td>
<td>218 (176, 269)</td>
<td>0.629 (0.505, 0.753)</td>
<td></td>
</tr>
<tr>
<td>2013</td>
<td>226 (187, 275)</td>
<td>0.685 (0.591, 0.769)</td>
<td></td>
</tr>
<tr>
<td>2014</td>
<td>261 (219, 312)</td>
<td>0.714 (0.624, 0.786)</td>
<td></td>
</tr>
<tr>
<td><strong>Model 3: Covariate and spatially structured random effects model</strong></td>
<td></td>
<td></td>
<td>526</td>
</tr>
<tr>
<td>2010</td>
<td>198 (187, 233)</td>
<td>0.606 (0.486, 0.717)</td>
<td></td>
</tr>
<tr>
<td>2011</td>
<td>218 (199, 238)</td>
<td>0.646 (0.528, 0.767)</td>
<td></td>
</tr>
<tr>
<td>2012</td>
<td>216 (200, 234)</td>
<td>0.587 (0.470, 0.707)</td>
<td></td>
</tr>
<tr>
<td>2013</td>
<td>219 (203, 236)</td>
<td>0.669 (0.559, 0.770)</td>
<td></td>
</tr>
<tr>
<td>2014</td>
<td>232 (212, 253)</td>
<td>0.708 (0.596, 0.806)</td>
<td></td>
</tr>
<tr>
<td><strong>Model 4: Full Model (both random effects in)</strong></td>
<td></td>
<td></td>
<td>657</td>
</tr>
<tr>
<td>2010</td>
<td>215 (160, 289)</td>
<td>0.659 (0.534, 0.758)</td>
<td></td>
</tr>
<tr>
<td>2011</td>
<td>223 (178, 278)</td>
<td>0.692 (0.563, 0.796)</td>
<td></td>
</tr>
<tr>
<td>2012</td>
<td>218 (176, 268)</td>
<td>0.638 (0.509, 0.756)</td>
<td></td>
</tr>
<tr>
<td>2013</td>
<td>226 (187, 275)</td>
<td>0.689 (0.592, 0.772)</td>
<td></td>
</tr>
<tr>
<td>2014</td>
<td>260 (219, 311)</td>
<td>0.712 (0.619, 0.786)</td>
<td></td>
</tr>
</tbody>
</table>
Table S 2. Regression coefficients and predicted case detection rates from the candidate models

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Coefficient, median (95% credible interval)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model 1</td>
</tr>
<tr>
<td>$\beta_0$</td>
<td>-5.3 (-5.5, -4.9)</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>0.09 (0.08, 0.11)</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>-0.08 (-0.14, -0.02)</td>
</tr>
<tr>
<td>$\beta_3$</td>
<td>0.08 (-0.10, 0.36)</td>
</tr>
<tr>
<td>$\text{CDR}_{HC}$</td>
<td>0.40 (0.36, 0.44)</td>
</tr>
<tr>
<td>$\text{CDR}_{NHC}$</td>
<td>0.57 (0.51, 0.64)</td>
</tr>
</tbody>
</table>

CDR$_{HC}$ - case detection rate in areas with health center, CDR$_{NHC}$ - case detection rate in areas with no health center, $\beta_1$ – population density, $\beta_2$ - temporal lag, $\beta_3$ - spatial lag

**Model 1**- Covariate only (no random effects),  
**Model 2**- Covariate and spatially un-structured random effects,  
**Model 3** - Covariate and spatially structured random effects,  
**Model 4** – Full model (Covariates and both random effects)

N.B- $\beta$ and $\beta_0$ are used interchangeably for the intercept in the published article
Chapter 5

5 The role of geospatial TB hotspots in the spatial spread of TB in Ethiopia: a mathematical model

Chapter overview

In order to determine the role of spatial hotspots in the spatial transmission of TB in Ethiopia, this Chapter utilises the information generated in the preceding chapters to develop spatially structured mathematical models. In particular, the incidence data generated using the novel method presented in Chapter 4 is used to identify geospatial TB hotspots which in turn are used to create spatially structured TB transmission model. The Chapter evaluates several approaches to simulating spatial transmission by presenting several candidate models for capturing local transmission patterns, and goes on to quantify the degree of transmission from spatial hotspots to neighbouring regions. The Chapter also provides estimates of the magnitude of infection in neighbouring locations that are caused by infectious individuals in a spatial hotspot. Findings from this Chapter provide useful insights into the potential implications of targeting spatial TB hotspots in Ethiopia.
The role of geospatial hotspots in the spatial spread of tuberculosis in rural Ethiopia: a mathematical model

Debebe Shaweno \textsuperscript{1,4}, James M. Trauer \textsuperscript{1,3,4}, Justin T. Denholm \textsuperscript{2,4} and Emma S. McBryde \textsuperscript{1,5}

\textsuperscript{1}Department of Medicine, and \textsuperscript{2}Department of Microbiology and Immunology, University of Melbourne, Melbourne, Victoria, Australia
\textsuperscript{3}School of Public Health and Preventive Medicine, Monash University, Melbourne, Victoria, Australia
\textsuperscript{4}Victorian Tuberculosis Program at the Peter Doherty Institute for Infection and Immunity, Melbourne, Victoria, Australia
\textsuperscript{5}Australian Institute of Tropical Health and Medicine, James Cook University, Townsville, Queensland, Australia

Geospatial tuberculosis (TB) hotspots are hubs of TB transmission both within and across community groups. We aimed to quantify the extent to which these hotspots account for the spatial spread of TB in a high-burden setting. We developed spatially coupled models to quantify the spread of TB from geographical hotspots to distant regions in rural Ethiopia. The population was divided into three ‘patches’ based on their proximity to transmission hotspots, namely hotspots, adjacent regions and remote regions. The models were fitted to 5-year notification data aggregated by the metapopulation structure. Model fitting was achieved with a Metropolis–Hastings algorithm using a Poisson likelihood to compare model-estimated notification rate with observed notification rates. A cross-coupled metapopulation model with assortative mixing by region closely fit to notification data as assessed by the deviance information criterion. We estimated 45 hotspot-to-adjacent regions transmission events and 2 hotspot-to-remote regions transmission events occurred for every 1000 hotspot-to-hotspot transmission events. Although the degree of spatial coupling was weak, the proportion of infections in the adjacent region that resulted from mixing with hotspots was high due to the high prevalence of TB cases in a hotspot region, with approximately 75\% of infections attributable to hotspot contact. Our results suggest that the role of hotspots in the geospatial spread of TB in rural Ethiopia is limited, implying that TB transmission is primarily locally driven.
1. Introduction

Tuberculosis (TB) is now the world’s leading infectious killer with an estimated 10.4 million cases and 1.7 million deaths in 2016. In the same year, TB caused 182,000 cases and 30,000 deaths in Ethiopia [1]. TB demonstrates marked spatial heterogeneity in distribution at any geographical scale and transmission often occurs in households and the general community, leading to the formation of localized transmission hotspots which act as hubs of TB transmission both within and across community groups [2,3]. Consequently, area-based TB control has been recommended rather than conventional contact investigations [4–7] and the identification and targeting of these spatial hotspots has been emphasized to achieve TB elimination goals [8].

In modelling spatial effects on the spread of disease, distinction is made between diffusion (spatially continuous) models and dispersal (metapopulation) models [9]. The first assumes random diffusion of infective individuals into adjacent areas. Dispersal models are used when the considered space is discrete with the population divided into patches [10,11] and assume cross-infection between infective and susceptible populations in different spatial subdivisions [9,12]. Because of spatial diffusion, spatial heterogeneities in disease burden can have important implications for the persistence of infections in a community. Even if disease dies out in some regions, in the presence of spatial structures the coupling between different regions can lead to repeated reintroductions, offsetting local control efforts [9,13].

Several spatial analyses of TB have identified localized transmission hotspots associated with areas of overcrowding and poverty [14–17]. However, the extent to which these geospatial hotspots drive the spatial spread of tuberculosis has not been documented, especially in settings with dispersed population settlements. Although considerable advancements in the methods used to investigate the spatial diffusion of infectious diseases have been made in the last decades [12,18,19], only a few modelling studies were able to apply such methods in the investigation of spatial transmission of TB by incorporating spatial structure. In addition, these studies were limited to overcrowded urban areas [20], and specific cross-border settings [21]. In this study, we aimed to understand the geographical spread of TB from hotspots to regions located at different distances in a remote region of Ethiopia.

2. Methods

2.1. Identification of spatial hotspots

In our previous spatial analysis of TB in Sheka Zone, a remote region of Ethiopia, we observed considerable spatial heterogeneity [16], with disease clustered in five kebeles (the smallest geographical administrative unit in Ethiopia) that contained 20% of the Zonal population but accounted for 53% of notifications. The clusters were identified using local Moran’s I at 95% confidence interval (electronic supplementary material, figure S1).

2.2. Formulation of the cross-coupled metapopulation model

To capture the spatial diffusion of TB from hotspots, we divided the Zonal population into three discrete spatial regions (patches) based on their proximity to hotspots, namely hotspots, adjacent and remote patches. While hotspots constituted areas identified as significant clusters on spatial analysis, adjacent regions comprised regions that share a border with hotspots (electronic supplementary material, figure S1, panel B). Kebeles not meeting the criteria for hotspots or adjacent areas were termed remote regions. The average case notification rates per 100,000 population in these patches respectively were 377, 77 and 97 per year.

2.3. Model assumptions on spatial coupling

The dynamics of TB were assumed to be identical in the three patches, except for the transmission parameter, which we calibrated on the basis of notification rates in each region. To capture transmission from a hotspot region to the other two regions and vice versa, we developed a cross-coupled metapopulation model by defining contact matrices referred to as WAIFW (‘who acquires
infection from whom\(^+)\) matrices that represent the strength of interaction within and between regions [22] as

\[
\begin{bmatrix}
\beta_{11} & 0 & 0 \\
0 & \beta_{22} & 0 \\
0 & 0 & \beta_{33}
\end{bmatrix}
\begin{bmatrix}
\beta_{11} & \rho_{12} & \rho_{13} \\
\rho_{21} & \beta_{22} & \rho_{23} \\
\rho_{31} & \rho_{32} & \beta_{33}
\end{bmatrix}
\begin{bmatrix}
\beta_{11} & \beta_{22} & \beta_{33} \\
\beta_{21} & \beta_{22} & \beta_{23} \\
\beta_{31} & \beta_{32} & \beta_{33}
\end{bmatrix}
\begin{bmatrix}
\beta_{11} & \beta_{22} & \beta_{33} \\
\beta_{11} & \beta_{22} & \beta_{33} \\
\beta_{11} & \beta_{22} & \beta_{33}
\end{bmatrix}
\]

where the diagonal elements \(\beta_{ii}\) (\(\beta_{11}, \beta_{22}\) and \(\beta_{33}\)) of the WAIFW matrix represent the average per capita effective contact rates per year that an individual in region \(i\) makes with individuals in region \(i\), while the off-diagonal elements (\(\beta_{ij}\)) represent the average per capita effective contact rates per year that an individual in region \(j\) makes with individuals in region \(i\). The degree of coupling between patches is assumed to be smaller than the degree of interaction within patches, such that the values of the off-diagonal elements are found by multiplying the appropriate within-group transmission parameter by the appropriate coupling proportion according to the following four different scenarios considered regarding coupling between regions.

In the first scenario, we considered no coupling between regions (figure 1, model A), with the force of infection determined solely by the prevalence in the index region. Next, we introduced coupling between regions with a single mixing parameter, such that neighbouring regions (i.e. hotspot with adjacent and adjacent with remote) have equal mixing and non-neighbouring regions mix to a lesser degree (i.e. hotspot with remote) (figure 1, models B and C). Third, we introduced coupling between regions with three different coupling parameters, so that each region can exert an independent force of infection on others at different distances (figure 1, model D). These mixing approaches are undertaken either under the assumption of different effective contact rates in each region (figure 1, models C and D) or with only hotspots having a greater transmission parameter (figure 1, model B). In models B, C and D, we assumed that any infective individual in a hotspot region is able to infect susceptible individuals in either of the other two regions and vice versa.

For model D, \(\rho_{1A}\) and \(\rho_{AR}\) are coupling proportions between adjacent regions while \(\rho_{HR}\) is between non-adjacent regions, with \(\rho_{HR} < \rho_{1A}\). These scaling parameters may take values between 0 and 1, spanning the range from completely uncoupled to maximally coupled systems [12].

The models were then fitted to recorded notification data over 5 years. Best fit solutions were used to find the \(\beta\) and \(\rho\) values. The force of infection on the \(i\)th group, \(\lambda_i\), assumes a frequency-dependent transmission and is therefore a weighted sum of infectious TB prevalence in the different spatial groups:

\[
\lambda_i = \sum_{j=1}^{n} \frac{\phi_i \beta_{ij} I_j}{N_j}.
\]

\(I_j\) and \(N_j\) refer to the number of individuals with active TB and total population in spatial group \(j\) respectively, while \(\beta_{ij}\) is the average number of effective contacts per year that an individual in region \(j\) makes with individuals in region \(i\).

In this model, the population in each of the three sub-regions is divided into five disjoint states based on TB status, according to a model structure we adapted from a recent publication by Trauer et al. [23]. Susceptible (S) represents individuals who have never been exposed to \(Mycobacterium tuberculosis\), while the early latent (E) compartment comprises individuals who have recently been infected (such that progressions to \(L\) from this compartment correspond to primary TB) [24], and the persistent latent (L) compartment represents individuals who were remotely infected but have not yet progressed to active TB. The active TB compartment (I) denotes individuals with active TB, while a recovered compartment (R) represents individuals who were cured by previous treatment or natural recovery (figure 2).

Susceptible individuals are replenished by births and depleted by infection through contact with infective cases at a rate proportional to the fraction of persons in the active state. Once infected, all individuals transition to the early latent compartment from which they either rapidly progress to...
active TB (I, at rate \( \varepsilon \)) or transition to the late latent state (L, at rate \( \kappa \)). From the late latent state, they may reactivate to join the active TB state (I) at a rate \( \nu \). Once an individual has progressed to active TB (I), they either experience natural recovery (\( \gamma \)), die from TB-related mortality (\( \mu_d \)) or are detected and commence treatment (\( \delta \)). Recovered individuals may relapse at rate \( \omega \), while recovered and latently infected individuals are subject to reinfection at a reduced rate. The rate of non-disease-induced mortality is constant (\( \mu \)), while the additional death rate due to disease affects only class I and is also constant (\( \mu_d \)).

We consider a closed population, with births replacing both TB-related and non-TB-related deaths, such that demographic effects only act to slowly replenish the susceptible pool over time. In the model, only a fraction of diseased individuals comprising all smear-positive and 22% of smear-negative TB patients is considered infectious, in line with a previous report that 17–22% of smear-negative TB cases are infectious [25]. The dynamic transmission model only captures drug-susceptible TB.

The coupling between regions and the rate of flow through compartments is described by the following system of ordinary differential equations:

\[
\begin{align*}
\frac{dS_i}{dt} &= \psi_i - \lambda_i S_i - \mu S_i, \\
\frac{dE_i}{dt} &= \lambda_i S_i + \lambda_i (1 - p)(L_i + R_i) - (\varepsilon + \kappa + \mu)E_i, \\
\frac{dL_i}{dt} &= \kappa E_i - (\lambda_i (1 - p) + \nu + \mu) L_i, \\
\frac{dI_i}{dt} &= \varepsilon E_i + \nu L_i + \omega R_i - (\delta_i + \gamma + \mu_d) I_i, \\
\frac{dR_i}{dt} &= (\delta_i + \gamma) I_i - (\lambda_i (1 - p) + \omega + \mu) R_i,
\end{align*}
\]

where

\[
\psi_i = \mu (S_i + E_i + L_i + R_i) + \mu_d I_i
\]

\[
\lambda_i = \sum_{j=1}^{n} \frac{d \beta_{ij} I_j}{N_j},
\]

The system of ordinary differential equation was solved using a Runge–Kutta algorithm. The numerical solutions were obtained using ode45 in Matlab 2015b.

### 2.4. Model fitting

The metapopulation model was fitted to 5-year TB case-notification data collected from clinical records on all TB patients diagnosed between 2010 and 2014 in health facilities of Sheka Zone, Ethiopia [15]. For model fitting, we aggregated data by patches and year based on patients’ places of residence.
When fitting the models to data, we used a Metropolis–Hastings algorithm, with a likelihood function that considered the observed TB notification rate in region \( i \) and year \( y \) as the realization of a Poisson process. The expectation of the Poisson distribution \( \lambda_{ij} \) is the mean notification rate from the dynamic TB transmission model at different effective contact rate values within and between regions. The mean notification rate is the flow rate per unit time from the \( I \) to the \( R \) compartment through case detection. The site index \( i \) runs from 1 to 3, representing the 3 patches and the year index \( j \) runs from 1 to 5 representing 5-year notification data. The Metropolis–Hastings algorithm was initialized by setting initial parameter values and assigning a uniform prior distribution between 0 and 1 for all proportions and all rates for which plausible values were less than 1 (for the parameter set). Then the likelihood of the parameter set given the data was calculated. We then iterated to find the best fitting parameter set, where at each iteration a new candidate parameter set \( \theta^* \) is randomly generated using a multivariate normal proposal distribution centred at the previous accepted parameter set, \( \theta \). Then the likelihood of the new candidate parameter set \( \theta^* \) given the data \( L(\theta^*) \) is determined using the Poisson distribution. The new parameter set \( \theta^* \) is accepted with probability \( p = \min(1, L(\theta^*) / L(\theta)) \). The same procedure was followed for each of the four models and the best fitting model was selected using the deviance information criterion (DIC). DIC assesses models in terms of both their goodness of fit and their parsimony, penalizing models according to the effective number of parameters [26]. We calculated DIC as the sum of the expected posterior value of deviance (2 times the average of the log-likelihood ratios) and the effective number of parameters in the model, described as the difference between the posterior mean of the deviance and the deviance at the posterior means of the parameters (likelihood ratio evaluated at the average of the parameters) of interest [26]. We ran the model over 120 000 iterations and discarded the first 110 000 as a burn-in. The model was coded in Matlab-R2015b (The MathWorks, 2015) and some of the plots were produced using ggplot2 library in R v. 3.3.1. We collected data used in this study after obtaining ethical approval from the University of Melbourne Health Sciences Human Ethics Subcommittee and the Zonal Health Department of Sheka Zone, Ethiopia.

2.5. Simulation of the impact of interventions in hotspots on the spatial spread of tuberculosis

To estimate the extent to which hotspots contributed to TB transmission in the two non-hotspot regions, we first ran the best candidate model to equilibrium using the best fitting parameter values. We subsequently modified the case detection rate only in the hotspots from 65% (baseline value) using increments of 5% to understand the effect this might have on the burden of TB in the two non-hotspot regions.

2.6. Parametrization

We considered identical model parameter values across each of the three patches except for case detection rate (table 1). The baseline case detection rate (CDR) of 65% was considered in the hotspot region, while a lower CDR (60%) was considered in the two non-hotspot regions.

2.7. Goodness-of-fit

We simulated notification data using fitted posterior parameter values to determine whether our model could reproduce the notification data that were previously used to estimate the model parameters. Estimated mean notification rates from the model using fitted parameter values were used as the mean of the Poisson distribution to yield simulated notification rates. We overlayed the observed notification rates on the histograms of simulated notification rates.

3. Results

3.1. Model comparison

Of the four models, the model assuming equal effective contact rates in adjacent and remote regions (B) was the poorest fitting model based on the DIC, with the simulated data from this model fitting the observed data poorly. The other three models which assumed different contact rates in each patch
(A, C, D) demonstrated better fit based on DIC values, including the two models (C and D) which incorporated coupling and the one which did not (A). The model incorporating spatial coupling with a single coupling parameter (C) had a slightly lower DIC than the non-coupled model (A) and the spatially coupled model with separate between-region coupling parameters (D) (table 2). Simulations of notification data based on parameters fitted using models B, C and D produced similar notification rates.

3.2. Parameter estimation

The best model was model C, which assumed different effective contact rates in each region and the proportion of coupling between two non-adjacent regions to be the square of the coupling proportion between adjacent regions. The results of this model are the focus of the rest of this Results section (for outputs from other candidate models, see electronic supplementary material, figures S4–S7). Model C estimated the mean effective contact rates in the hotspot, adjacent and remote regions to be 55.5 (95% credible interval (95% CrI): 52.9, 58.9), 2.2 (95% CrI: 0.2, 7.4) and 14.7 (95% CrI: 13.0, 16.4) per year (table 1). Similarly, this model estimated the strength of coupling between hotspot and adjacent regions to be 4.5% (95% CrI: 2%, 6%), and between hotspots and remote regions to be 0.2%. This means that for every thousand hotspot-to-hotspot transmission events, 45 transmission events occur from hotspot-to-adjacent regions and 2 transmission events occur from hotspot-to-remote regions.

Each of the four models estimated similar effective contact rates in hotspot and remote regions, which is also true for estimated notification rates except for model B. The posterior distributions of estimated parameters from the best fitting cross-coupled metapopulation model (C) were well-fitted by a normal distribution, except for the effective contact rate parameter in the adjacent region that was well-fitted by a gamma distribution (shape parameter $\theta = 1.87$, scale parameter $\lambda = 1.19$) (electronic supplementary material, figure S2).

Our model could satisfactorily reproduce the notification rates that were previously used to estimate the set of model parameters (figure 3). We also tested if both the observed and simulated data came from the same underlying distribution using Kolmogorov–Smirnov test. The test indicated that both the observed data and the simulated data came from the same underlying distribution for all the three patches (hotspots: $D = 0.36$, $p$-value = 0.54; adjacent region: $D = 0.54$, $p$-value = 0.11; remote region: $D = 0.31$, $p$-value = 0.71).

### Table 1. Numerical values of model parameters.

<table>
<thead>
<tr>
<th>Fixed input parameters</th>
<th>Value</th>
<th>References</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>natural mortality rate, $\mu$</td>
<td>0.0154</td>
<td>[27]</td>
<td>year$^{-1}$</td>
</tr>
<tr>
<td>fast progression rate, $\varepsilon$</td>
<td>0.4</td>
<td>[28]</td>
<td>year$^{-1}$</td>
</tr>
<tr>
<td>stabilization rate, $\kappa$</td>
<td>3.6</td>
<td>[28]</td>
<td>year$^{-1}$</td>
</tr>
<tr>
<td>reactivation rate, $\nu$</td>
<td>0.002</td>
<td>[28]</td>
<td>year$^{-1}$</td>
</tr>
<tr>
<td>untreated mortality rate, $\mu_d$</td>
<td>0.125</td>
<td>[29]</td>
<td>year$^{-1}$</td>
</tr>
<tr>
<td>natural recovery rate, $\gamma$</td>
<td>0.205</td>
<td>[29]</td>
<td>year$^{-1}$</td>
</tr>
<tr>
<td>proportion of incident TB smear-positive</td>
<td>0.33</td>
<td>[30]</td>
<td>proportion</td>
</tr>
<tr>
<td>proportion of incident TB smear-negative</td>
<td>0.35</td>
<td>[30]</td>
<td>proportion</td>
</tr>
<tr>
<td>case detection rate, $\delta$</td>
<td>65%$^a$, 60%$^b$</td>
<td>[31]</td>
<td>proportion</td>
</tr>
<tr>
<td>relapse, $\omega$</td>
<td>0.002</td>
<td>[32]</td>
<td>year$^{-1}$</td>
</tr>
<tr>
<td>fraction of smear-negative TB infectious</td>
<td>0.22</td>
<td>[25]</td>
<td>proportion</td>
</tr>
<tr>
<td>fraction of infectious cases, $\phi$</td>
<td>0.40</td>
<td></td>
<td>proportion</td>
</tr>
<tr>
<td>protection against infection from latency, $p$</td>
<td>0.79</td>
<td>[33]</td>
<td>multiplier</td>
</tr>
</tbody>
</table>

$^a$In hotspots.

$^b$In both non-hotspot regions.

$^c$The fraction of active cases that are infectious is calculated as the proportion of smear-positive TB (0.33) plus 0.22 times the proportion of smear-negative TB (0.35). The remaining fraction of active cases (0.32) are extrapulmonary and non-infectious.
3.3. Analysis of covariance of fitted model parameter posterior probability distributions

We further assessed the behaviour of model C by examining the relationship between parameter pairs. The effective contact rates in hotspot ($\beta_{11}$) and remote regions ($\beta_{33}$) were not correlated with the coupling term ($\rho$). However, the number of effective contacts in an adjacent region ($\beta_{22}$) and the coupling term ($\rho$) demonstrated a strong negative correlation (correlation coefficient: $-0.7$) (figure 4). The presence of considerable correlation between $\beta_2$ and $\rho$ reduced the precision of the estimates for these two parameters.

---

Table 2. Credible intervals of estimated parameters and outputs from candidate models. Model A—no coupling; model B—coupled, and similar mixing in the two non-hotspot regions; model C—coupled, with area-specific contact rates; model D—coupled, with area-specific contact rates and three separate coupling terms between regions. $\rho_{HA}$: hotspot—adjacent region coupling; $\rho_{AR}$: adjacent—remote coupling; $\rho_{HR}$: hotspot—remote coupling.

<table>
<thead>
<tr>
<th>parameters</th>
<th>median (95% CrI) of posterior distributions of model parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>model A</td>
</tr>
<tr>
<td>$\beta_{11}$</td>
<td>55.2 (52.8, 58.0)</td>
</tr>
<tr>
<td>$\beta_{22}$</td>
<td>14.5 (13.2, 16.1)</td>
</tr>
<tr>
<td>$\beta_{33}$</td>
<td>15.3 (13.8, 16.9)</td>
</tr>
<tr>
<td>$\rho_{HA}$</td>
<td>a</td>
</tr>
<tr>
<td>$\rho_{AR}$</td>
<td>a</td>
</tr>
<tr>
<td>$\rho_{HR}$</td>
<td>a</td>
</tr>
<tr>
<td>notification rate,</td>
<td>384 (347, 424)</td>
</tr>
<tr>
<td>hotspot</td>
<td>a</td>
</tr>
<tr>
<td>notification rate,</td>
<td>87 (69, 106)</td>
</tr>
<tr>
<td>adjacent</td>
<td>a</td>
</tr>
<tr>
<td>notification rate,</td>
<td>97 (78, 117)</td>
</tr>
<tr>
<td>remote</td>
<td>a</td>
</tr>
<tr>
<td>DIC</td>
<td>197</td>
</tr>
</tbody>
</table>

*aNot applicable.

Figure 3. Comparison of simulated notification rate (histograms) and observed notification rates over 5 years (vertical dashed blue lines).
3.4. Simulation of the impact of improving case detection in hotspots on the neighbouring regions

In a cross-coupled metapopulation formulation, infection in a given patch is a function of disease prevalence in all sub-regions (patches) and the degree of coupling between and within regions. Thus, although the extent of spatial coupling between hotspots and the other two regions is weak, the proportion of TB infection in an adjacent region due to mixing with hotspots was about 76% (95% CrI: 52%, 92%); and was 2.5% (95% CrI: 0.7%, 3.2%) in the remote regions in the first 5 years at a baseline CDR in hotspots (figure 5). That is, out of 771 (95% CrI: 258, 1736) secondary infections in the adjacent region, 593 (95% CrI: 237, 904) were generated due to mixing with hotspots. As a result, improving CDR in the hotspot regions has significant impact on the neighbouring region. For instance, the proportion of infections in the adjacent regions due to mixing with hotspot regions drops by 10% down to 65% (95% CrI: 37%, 86%) when case detection rate in the hotspot regions goes from 65% (the baseline) to 70%. Similarly, when case detection rate goes from 65% to 95%, the proportion of infections in the adjacent regions due to mixing with hotspot regions drops down to 16% (95% CrI: 7%, 39%). However, the effect of the same intervention is only marginal in the remote regions (figure 5). This is because of extremely large number of prevalent infectious cases in hotspots who are available to make cross-contact.

4. Discussion

Using cross-coupled metapopulation models, we quantified the role of TB hotspots in the spatial spread of TB in rural Ethiopia, demonstrating that spatial coupling between TB hotspots and the surrounding regions is limited. However, despite limited mixing between hotspots and adjacent regions, the very high rate of transmission in hotspots means that they contribute significantly to disease in immediately adjacent surrounding areas.

Of the models considered, the model that assumed spatial coupling with different effective contact rates by region attained the best combination of accuracy and parsimony, although a similar model with different transmission rates by region but without spatial coupling also demonstrated a
reasonable fit to data. The best fitting model predicted 45 hotspot-to-adjacent region transmission events and two hotspot-to-remote region transmission events for every 1000 hotspot-to-hotspot transmission events, although it was not possible for us to confirm this with additional data on mobility or organism genotypes. However, we consistently predicted a coupling proportion of less than 10% between adjacent regions using all of our cross-coupled models.

In our models, the extent of population mixing between regions was modest, implying that the probability of contact with externals is much lower than the probability of local contact and that TB transmission in rural Ethiopia is predominantly locally driven. This is consistent with our expectations, as the study was conducted in a rural area with a dispersed population and limited long-distance population movement (most population movements are by foot), as well as its close fit to data. Thus, TB control efforts targeting hotspots in rural Ethiopia may not achieve the anticipated impact on community-wide TB control, although this is the subject of further investigation.

Although coupled metapopulation models are standard approaches in the presence of heterogeneity, there is no universally accepted approach to quantifying epidemiological coupling between different regions [9,18,19]. Previous works (mainly of measles) have used different approaches: considering a range of coupling parameters [12], estimating coupling by trial and error from simulation [34] or using intuition alone. For measles, the coupling term has been estimated between $10^{-2}$ and $10^{-1}$ in higher resource settings [12,18,19]. In common with many of the studies described above, our study used a simulation approach by developing an algorithm to find the best fitting coupling term and transmission parameters. However, the strong correlation between the transmission term in adjacent regions and the coupling term might have reduced the amount of information available on cross-coupling. Thus, in the presence of trade-off between the effective contact rate in adjacent regions and the coupling term, the accurate interpretation of the extent of coupling or the effective contact rate in the adjacent region remains a challenge. Future works would require genotypic data to validate the extent of transmission from hotspots to neighbouring communities, although this information is often unavailable in high incidence settings.

Our coupling term is considerably lower than hotspot-to-community transmission term used in a modelling study of urban Rio de Janeiro, in which each infective in a hotspot region was assumed to cause 0.5 transmission events outside of the hotspot for each event caused within the hotspot [20]. However, it is important to note that this parameter in the Rio de Janeiro study was derived from a study that was based on a very small sample size ($n = 10$) and geographical clustering was not statistically defined [35]. Rather, geographical clusters were defined by at least two cases sharing a
common molecular structure and from the same or close neighbourhoods. Moreover, it should be noted that these urban settings could be much more strongly coupled than the broad geographical region considered in this study, such that these estimates may both be accurate.

Recent works have recommended quantification of coupling terms by measuring population mobility patterns [19]. An important modelling study describing the impact of cross-country border population mobility on TB burden in a low incidence setting concluded that TB in the Australian Torres Strait region is driven by the TB dynamics in Papua New Guinea [21]. This was not unexpected given that 98% of the mobility is to Australia for cultural practices as well as for healthcare. Similarly, another cross-border study also suggested that expanding the directly observed treatment short course programme in the neighbouring high incidence settings (Mexico, Haiti and Dominican Republic) could reduce TB-related morbidity and mortality among migrants to the United States [36]. In contrast, in our study setting, there is no similar reason to assume large population movements between rural regions which do have similar features in terms of healthcare access.

Although the extent of coupling observed in this study is low (less than 5%), the proportion of infections in the adjacent region due to mixing with hotspots is considerable (more than 75%). This emerges from the proportional relationship between the coupling term and intra-region mixing (prevalence), with the large number of infectious cases in a hotspot region relative to the adjacent regions leading to a considerable contribution to infection in the adjacent region. This is consistent with a previous analysis indicating that infectious disease persistence in a community can be due to either high intragroup mixing, strong coupling or both [13]. Thus, TB control efforts in extremely high transmission regions may provide additional TB risk reduction to surrounding regions.

Our study has important limitations. Our classification of the study region into three spatially discrete groups may be overly simplistic, although increasing the number of patches simulated would present additional challenges in fitting many parameters. More complex models, such as spatially continuous models or agent-based models, may be useful to investigate these links at finer spatial scales. In addition, the model presented here assumes only a single strain of drug-susceptible TB in this setting where drug susceptibility testing is unavailable. Considering multidrug-resistant TB, HIV and urban dynamics into the model is limited by available data, but will be the subject of future research.

5. Conclusion

Our study suggests that TB in rural Ethiopia is primarily driven by local transmission, rather than spillover from hotspot regions. However, the epidemiology of tuberculosis in regions adjacent to transmission hotspots is considerably contributed by these hotspots. Control efforts in high transmission regions may provide some additional TB risk reduction in surrounding regions, although locally focused measures remain essential.

Acknowledgements. The authors thank Sheka Zone Health Department, Ethiopia, and all health centres for permission to access the data.


5.1 Supplementary Materials

5.1.1 Spatial TB clusters in Sheka Zone, Ethiopia

Figure S 2. Spatial distribution of TB in Sheka Zone, 2010-2014

Panel A shows the spatial clustering according to local Moran’s I, where High-High (also called hotspot) indicates that high incidence regions are adjacent to high incidence regions. Similarly, Low-Low indicates that areas with significantly lower diseases burden are neighbouring regions with significantly low values. High-Low or low high clusters indicate spatial outliers where high incidence regions are adjacent to low incidence regions. Panel B shows the spatial metapopulation structure used in this analysis. The selected hotspot region constitutes a larger city and its surroundings.

5.1.2 Deviance information criterion

In this study, we computed the deviance information criterion as follows based on an approach presented in Spiegelhalter et al. (2002):
Deviance was defined as $-2$ times the loglikelihood and can be written as:

$$D(y, \theta) = -2\log p(y|\theta) \ldots[147]$$

Having defined deviance as above, the deviance information criterion (DIC) can be computed as follows:

$$\text{DIC} = \bar{D} + PD \ldots[2]$$

$$PD = E_{\theta|y}[D] - D(E_{\theta|y}[\theta]) \ldots[3]$$

$$= \bar{D} - D(\bar{\theta}) \ldots[4]$$

where, $\bar{D}$ is the expected posterior value of the deviance (the average of log-likelihood ratios) and PD is the effective number of parameters in the model. PD is calculated as the difference between the posterior mean of the deviance and the deviance at the posterior means of the parameters (likelihood ratio evaluated at the average of the parameters) of interest. Adding the PD to the posterior mean deviance ($\bar{D}$) gives a deviance information criterion for comparing models [148].

Models with the lowest DIC estimates indicate models with a better performance, although there is no agreement on the magnitude of the difference between different DIC values that constitutes a meaningful difference. A general rule of thumb suggests that differences of more than 10 should definitely rule out the model with the higher DIC, while differences between 5 and 10 are substantial. However, if the difference is less than 5 and if different inferences would be drawn from the models being compared, then it is misleading to report the model with the lowest DIC only [149].

### 5.1.3 Calculation of the likelihood

In this analysis we used Poisson likelihood in the model fitting as follows:

$$\text{Likelihood} = \prod_{i=1}^{3} \prod_{j=1}^{5} \frac{\lambda^k e^{-\lambda}}{k!}$$

The subscript $i$ takes values from 1 to 3 to index each spatial patch, and the subscript $j$ takes values from 1 to 5 to index years. $\kappa$ refers to the observed notification and $\lambda$ represents the expected Poisson mean, which is the model estimated notification.
5.1.4 Outputs from the best fitting model

Figure S3. Posterior distribution of fitted parameters under the assumption of mixing with different contact rates in each region

(μ- mean, σ- standard deviation parameters of normal distributions; α and β are the shape and scale parameters of a gamma distribution)
5.1.5 Outputs from the candidate models

1. Model A: Non- spatially coupled model ($\rho = 0$)

   Posterior distributions of fitted parameters: no spatial coupling

   Figure S 5. Posterior distribution of fitted parameters.

   $\beta_{11}$-hotspot, $\beta_{22}$-adjacent, $\beta_{33}$-remote region

2. Model B: Coupled model with equal contact rates in the two non-hotspot regions
3. Model D: Spatially coupled model with separate coupling parameters across different regions

Figure S 6. Posterior distribution of fitted parameters;

i.e. $\beta_{11}$-hotspot, $\beta_{22}$-adjacent & remote, $\rho$- coupling term.

Figure S 7. Correlation between pairs of parameters of model D
Figure S 8. Posterior distribution of parameters, model D
Chapter 6

6 Impact of geographically targeted interventions for TB control in Ethiopia: A mathematical model

Chapter overview

This Chapter utilises data generated in the preceding chapters, and especially from Chapter 6, to develop spatially structured mathematical models to determine the impact of different interventions strategies targeting spatial TB hotspots in Ethiopia. The Chapter compares three intervention scenarios: spatial hotspot targeting, non-spatial hotspot targeting and spatially un-targeted interventions. Effectiveness of these interventions were determined by calculating the effort it takes to administer each intervention. The effort here refers to the proportion of the total population that needs to be screened under each of these intervention strategies to achieve a given reduction in the overall TB prevalence in the study region.
6.1 Abstract

**Background:** Tuberculosis (TB) exhibits considerable spatial heterogeneity, occurring in clusters that may act as hubs of community transmission. We evaluated the impact of an intervention targeting spatial TB hotspots in a rural region of Ethiopia.

**Methods:** To evaluate the impact of targeted active case finding (ACF), we used a spatially structured mathematical model that has previously been described. From model equilibrium, we simulated the impact of a hotspot-targeted strategy (HTS) on TB incidence ten years from intervention commencement and the associated cost-effectiveness. HTS was also compared with an untargeted strategy (UTS). We used logistic cost-coverage analysis to estimate cost-effectiveness of interventions.

**Results:** At a community screening coverage level of 95% in a hotspot region, which corresponds to screening 20% of the total population, HTS would reduce overall TB incidence by 52% compared with baseline. For UTS to achieve an equivalent effect, it would be necessary to screen more than 80% of the total population. In a scenario in which, case detection rate (CDR) is increased to 75% in each strategy (from the baseline level of 65% in hotspots and 60% in non-hotspots), the cost per averted case of active TB was estimated at USD 800 for HTS and USD 2,700 for UTS. Compared with hotspot targeting, the UTS approach would avert 24 new TB cases at an additional total cost of USD 162,200, which translates to an incremental cost-effectiveness ratio of USD 6700 for each additional case averted. Where regional TB program spending was capped at current levels, the intervention impact increased with increasing budget allocation to hotspots. Increased regional spending was associated with further incidence reductions, with maximum gains seen when an increased regional budget (five times higher than currently available) was shared between hotspots and non-hotspot regions in the ratio of 60% to 40%.

**Conclusions:** Our analysis suggests that a spatially-targeted strategy is efficient and cost-effective, with the potential for significant reduction in overall TB burden.

**Keywords:** Geospatial clustering, mathematical modelling, tuberculosis, active TB case finding, cost-effectiveness analysis
6.2 Background

Geographic heterogeneity is a defining characteristic in tuberculosis (TB) epidemiology [150, 151]. This means that a small fraction of the population bears the highest burden of disease, while the majority of the population carries a considerably lower burden – raising the possibility that geographically targeted interventions may be particularly effective. However, failure to target resources to these locations has been described as one of the reasons for slow progress in TB incidence reduction [152].

Current TB control programs miss a considerable number of cases, estimated at 36% in 2017, which poses an important challenge for global TB control [4]. Therefore, active screening will likely be critical to reach the missed cases, although indiscriminate mass screening is expensive and is currently discouraged [19]. Instead, active case finding (ACF) in geographical areas with a high disease prevalence is favoured [10, 19, 140], although the details of how to implement such a strategy are not fully defined. Few studies have evaluated the impact of targeting these locations, and there remains no consensus on how to define a spatial TB hotspot [153]. A proof of concept study in urban Brazil that evaluated the impact of a hypothetical intervention aimed at reducing time to detection showed that spatial targeting was superior to community-wide approaches [137]. In contrast, results from a study that evaluated the impact of hypothetical vaccine targeting spatial hotspots in India showed modest benefits and the extent of heterogeneity was an important determinant of impact [138], implying that the effectiveness of spatial targeting was unpredictable and setting-dependent.

In this study, we estimated the impact of spatially targeted active TB contact screening in a remote zone in Ethiopia.
6.3 Methods

The analysis we present here extends our previous work which explored the implications of spatial heterogeneity for the spatial spread of TB [153].

6.3.1 Overview of past work

In our previous study we used spatial analysis techniques [13] to identify TB hotspots in Sheka Zone, Ethiopia, which comprised about 22% of the Zonal population but contained 50% of notified TB cases. The contribution of hotspots to overall dynamics was determined by dividing the overall study region into three sub-regions based on disease burden and location. These were designated hotspots (statistically significant spatial clusters), non-hotspot regions adjacent to hotspots (having a shared border with hotspots) and remote regions (no shared border) [153] (Supplementary Material, Figure S2).

We then used a compartmental transmission dynamic TB model which divided the population into five epidemiological compartments depending on TB-related infection or disease progression status, and three geographic patches [153] (Supplementary Material, Figure S2). The transmission dynamics of TB were assumed to be identical in each spatial subdivision, except that the per capita effective contact rate ($\beta$) was calibrated to the local notification rate and the baseline case detection rate.

*Mycobacterium tuberculosis* (*Mtb*) transmission within and across regions was captured by constructing models with different contact mixing matrices to represent the strength of interaction within and between regions [153]. We parameterised the model by considering identical model parameter values across each of the three regions except for case detection rate (CDR) and transmission parameters. The baseline CDR of 65% was considered in the hotspot region, while a lower CDR of 60% was applied in the non-hotspot regions, based on our earlier findings. A full list of model parameters used in or estimated from our previous work including the cross-region coupling rate of 4.6% are provided in the Supplementary Material, Table S3.

6.3.2 Intervention strategies

We simulated an ACF strategy because the current national TB control program in Ethiopia misses a considerable number of TB cases [4]. We modelled ACF as a community intervention assuming door to door enquiry for chronic cough ($\geq$ 2 weeks) followed by two sputum samples for microscopy [154]. Individuals with other TB suggestive symptoms (chronic weight loss,
fever ≥ 2 weeks and night sweets ≥ 2 weeks) are referred to the nearest health facility for clinical evaluation and specimen culture. The average cost per screened individual is estimated at USD 2.5 after accounting for inflation [155].

This community screening intervention (ACF) was conducted under both a spatially targeted strategy (HTS) and a spatially untargeted strategy (UTS). Under HTS, we explored the role of spatially targeted ACF by incrementally increasing CDR in the hotspot region from the baseline value of 65%, while maintaining the baseline CDR at 60% in the remaining non-hotspot regions. Under UTS, CDR was increased from the baseline value in all spatial subdivisions. The model was implemented from equilibrium, with outputs of spatially targeted case finding strategy (HTS) compared against the base-case scenario and the spatially untargeted case finding strategy (UTS) ten years after intervention [153].

The intervention increases TB CDR by reaching the proportion of cases that are missed during their course of illness. Thus, the proportion of cases detected and treated under the intervention is the sum of the baseline CDR (under existing programmatic conditions) and the proportion of missing cases detected by the intervention:

\[ \tau_i = p_i (1 - \tau_{0i}) + \tau_{0i} \]

where \( i \) may take two values representing spatial hotspots or the overall region, \( p_i \) represents the coverage of population screening in the spatial subdivision of interest, \( \tau_i \) represents the intervention CDR in the subdivision, and \( \tau_{0i} \) refers to the baseline CDR in the subdivision under the national TB program. The value of \( p \) ranges from 0 (no intervention) to 1 (full coverage). The intervention case detection proportion (\( \tau_i \)) generated in this way was converted to case detection rate (\( \delta_i \)) to be used in the model as follows:

\[ \delta_i = \frac{\mu_i (\mu_{di} + \gamma_i)}{1 - \tau_i} \]

where \( i \) may take two values representing spatial patch, \( \mu_{di} \) represents TB related mortality, \( \tau_i \) represents the intervention CDR (proportion) in the subdivision, and \( \gamma_i \) refers to natural recovery of TB.

Because HTS and UTS operate in regions with different population sizes, direct comparison of the impact of an increase in local screening coverage could be misleading. Thus, to account for differences in efforts of intervention strategies such that impacts are comparable, we used two
approaches to estimate the efficiency of intervention strategies. In the first approach, we translated local screening coverage in spatial hotspots into the proportion of entire population reached (total screening coverage) by multiplying local coverage by the proportion of total population in it. In the second approach, we determined efficiency using cost-coverage and cost-effectiveness analysis as described below.

### 6.3.3 Cost-coverage analysis and cost-effectiveness analysis

The community intervention screening coverage ($p_i$) in the spatial subdivision of interest was used as an input to a logistic cost-coverage function that linked spending on programmatic interventions to intervention coverage [156]. Spending on each intervention strategy was estimated using a unit cost of USD2.5 per TB review in Ethiopia (in 2018) reflecting the direct costs of TB diagnosis incurred by a TB control program adjusted for inflation [155]. This cost estimate is based on a study from a single facility and hence may not be representative for the country. Thus, a sensitivity analysis of allocative efficacy to changes in unit-cost per person screened for TB is provided (*Supplementary Material- Figures S9 and S10*). The cost-coverage relationship is then determined by fixing the saturation (maximum possible coverage of screening) at 80%.

We also calculated average cost-effectiveness ratios (CER) for each intervention strategy by comparing cost and impact (reduction in incidence) at a given coverage with the respective values at baseline. The incremental cost-effectiveness ratio (ICER) was calculated as the ratio of the difference in cost to the difference in the number of active TB cases averted between the two interventions [157]. CER and ICER in this study represent cost per active TB case averted.

Often national TB programs have a fixed budget and it is unclear what proportion of this budget should go to the spatial hotspots in contrast to the non-hotspot regions for a maximum possible impact. To maximise the impact of the available budget, we estimated annual TB funding at approximately USD 65,000 across the study Zone (based on available budget in the Zone, 2018) and identified the optimal resource allocation strategy for this funding envelope. This was achieved by implementing a model such that for every proportion of available budget going to spatial hotspots, the remaining proportion goes to the non-hotspot regions. Similarly, hypothetical annual budget scenarios were considered to define the ideal maximum potential gain of spatially targeted resource allocation.
6.3.4 Sensitivity analysis

To account for the extent to which variations in parameter values were related to variations in outcome variable (incidence rate), we carried out a multivariate sensitivity analysis using Latin hypercube sampling [158]. Sensitivity analysis was done for epidemiological parameters governing transitions between compartments, case detection ratio and population mixing, while keeping the intervention (population screening coverage) constant. These included the within region transmission parameters $\beta_{ii}$ ($\beta_{11}$ within hotspots, $\beta_{22}$ in adjacent regions and, $\beta_{33}$ within remote regions), between region coupling parameter ($\rho$); fast progression rate ($\epsilon$); stabilisation rate ($\kappa$); reactivation rate ($\nu$); case detection ratio (CDR) and relapse rate ($\omega$). We calculated partial rank correlation coefficients (PRCC) to quantify the relationship between model input parameters and the projected impact of targeted active case finding (overall incidence). The parameter values and their ranges used in the sensitivity analysis are provided in the Supplementary Material, Table S3. The model was coded in Matlab-R2015b.
6.4 Results

In our previous work, using our baseline parameter values, we estimated TB incidence at 538 cases per 100,000 per year in a hotspot containing one-fifth of the Zone’s population, and 110 and 150 cases per 100,000 population per year in adjacent and remote regions respectively. The population-wide prevalence and incidence rates were estimated to be 267 cases per 100,000 and 221 cases per 100,000 population per year, respectively [153].

6.4.1 Impact of intervention strategies on TB epidemiology

HTS resulted in a reduction of overall TB incidence by 52% with 20% of the total population reached, which corresponded to achieving a CDR of 95% in the hotspot region. However, to achieve the same benefit from UTS, the proportion of people that needed to be reached was about four-fold the coverage in the HTS (Figure 2).

![Figure 2](image)

Figure 2. Proportion of entire population screened and reduction in TB incidence for the entire study region, hotspot regions and the two non-hotspot regions.

Note in figure 2 that if interventions in the hotspot region are maximally implemented, they would cover only 22% of the entire population, such that the solid lines terminate at 20%. While HTS that reached 20% of the total population (corresponding to a CDR of 95% in the hotspot region) reduced TB incidence by 78% in the hotspot region, it reduced incidence
by 59% in the adjacent region and by only 2.7% in the remote region after 10 years of implementation. However, UTS had the potential for greater impact by extending coverage throughout the region and could reduce TB incidence by more than 74% in the region at the programmatic coverage value of 95%.

### 6.4.2 Sensitivity analysis

When the intervention parameter is kept constant, the estimated incidence was highly sensitive to changes in several model parameters. As expected, the parameters governing TB transmission ($\beta$’s and $\rho$), fast progression rate, and reactivation rates ($\nu$) were positively correlated with estimated incidence, while changes in the baseline case detection rate were negatively correlated in the opposite direction, with partial correlation coefficient of 0.73 (Figure 3).

---

**Figure 3.** Sensitivity of predicted overall TB incidence to variations in selected model parameters

($\kappa$ - stabilisation rate, $\nu$ - reactivation rate, $\omega$ - relapse rate, CDR - proportion of cases detected in hotspots, $\rho$ - coupling parameter, $\epsilon$ - fast progression rate, $\beta$'s – within region per-capita effective contact rates) while the intervention parameter (screening coverage) was kept constant at 0.
6.4.3 Cost-coverage curves

Figure 4 presents the effect of increased program spending (cost) on intervention coverage, reduction in TB incidence, number of people screened and averted active TB cases under the two intervention strategies - UTS and HTS. The logistic cost-coverage plot (panel A) shows that the HTS (solid line) saturates early compared to the UTS (broken line), which requires considerably more spending to approach saturation because of the greater population to target. The intervention spending associated with specific coverage values from the logistic cost-coverage curves were retrieved and used to calculate the cost-effectiveness ratios as described in the next section.

![Cost-coverage curves](image)

Figure 4. Impact of program spending

In the above figure, panel A compares cost-coverage analysis of HTS and UTS. Panels B, C and D show impacts of increased program spending on the number of people reached, averted number of incident TB cases and incidence.
6.4.4 Cost-effectiveness analysis

HTS that aims to increase CDR to 75% in hotspot regions is expected to avert 52 new active TB cases at an increased cost of USD 42,400. The resulting average CER is USD 800 per additional case of active TB averted. Similarly, UTS will avert a total of 76 new cases at an increased cost of USD 204,600, resulting in average CER of USD 2,700 per additional case averted. Compared with a hotspot targeting strategy, the UTS approach would avert 24 additional active TB cases at an additional cost of USD 162,200, which translates to USD 6,700 per averted active TB case.

6.4.5 What proportion of TB budget should go to hotspots for a maximum impact?

At the available total budget of USD 65,000, the simulations suggest that when all the budget goes to the non-spatial hotspots, the overall incidence would decline by 7.7%. In contrast, in the extreme opposite situation where all the available budget goes to the spatial hotspots, impact increases with a reduction in overall TB incidence of 31.5%. However, when the available budget is shared half-half between the two regions, incidence decreases by 23%.

We also explored the relation between cost and its impact on TB incidence using various assumed budget envelopes: USD 100,000, USD 200,000 and USD 300,000. Using these envelopes, the relation between the proportion of budget going to spatial hotspots and impact (reduction in incidence) is not linear, particularly for the USD 200,000 and USD 300,000 envelopes. Under the USD 300,000 envelope, maximum impact was observed when about 58% of the total budget was allocated to spatial hotspots. When the proportion of the budget allocated to hotspot exceeded 60% population incidence rises markedly (Figure 5).
Figure 5. Impact of geographic funding allocation on population TB incidence under four budget envelope scenarios.
6.5 Discussion

Using a spatially structured mathematical model, we found that targeting spatial hotspots is efficient compared with spatially untargeted intervention up to a ceiling of around 20% overall coverage of the population. Compared with the UTS, HTS was predicted to be more efficient and cost-effective under most scenarios. Our results suggest that to obtain comparable reductions in TB incidence from HTS and UTS, the required coverage of the UTS intervention was about four-fold that of HTS before approaching saturation in hotspots. While less efficient, this strategy has the advantage of allowing further scale-up to reach a greater proportion of the population and could achieve a theoretical impact of a 60% reduction in TB incidence if 87% of the population were reached. However, it should be remembered that this level of coverage in ACF is much higher than current levels and has not been described to date.

The impact of spatially-targeted ACF is primarily confined to regions in close proximity to the hotspots, suggesting that spatial targeting would be effective in rural-remote settings such as the ones we studied here. The impact of hotspot targeting that resulted in significant reduction in the overall TB incidence as well as incidence in the proximal regions reflects significant transmission from hotspots to the other regions. However, given that the fitted cross coupling rate is low (4.6%), the impact of spatial targeting in reducing the overall incidence of TB primarily reflects the disproportionate burden of TB in hotspot regions, which is 3-4 times the disease burden in remote and adjacent regions. Consistent with this explanation, passive CDR and entire population screening in a hotspot region lead to saturation of effect and an incidence rate that remains above 110 cases per 100,000 population. This implies that spatial targeting alone would not achieve disease elimination, as a considerable burden transmission persists in non-hotspot regions [153].

ACF could yield considerable gains when targeted to population at increased risk of developing active TB including household contacts of patients with active TB [159], areas with high HIV prevalence [160] and people in congregate settings (eg. prison) [161]. Although geographic heterogeneity in TB risk has been documented well [150], evidence has been lacking on the epidemiologic and economic impact of spatially targeted ACF and hence remained unimplemented. The study we presented here shows, like other high TB risk groups explored in previous studies[152, 159, 160], geographically targeted ACF would provide considerable gains in TB control.
The potential effectiveness of spatially targeted ACF is a consequence of two main phenomena. First, current passive approaches to case detection rely on the index case seeking health care and so miss many people with TB. Health care seeking behaviour usually occurs in advanced stages of clinically apparent disease and so may have limited impact on transmission [162]. Conversely, indiscriminate population-wide screening could produce high levels of coverage at a markedly increased cost and so may not be feasible [163]. Thus, this provides a rationale for considering spatially-targeted ACF. Moreover, spatially-targeted ACF could be feasible because routinely collected data and the techniques we propose could be used to identify the hotspots [164].

Our study suggested that when the available budget is limited, as in our study setting, the intervention impact increases with increasing budget allocation to hotspots. At about five times higher budget envelope than currently available in the study region, maximum gains in incidence reductions were observed when budget is shared between the regions in the ratio of 60% to 40%. In contrast, at budget envelopes that are up to three times the current budget, maximum reduction in overall TB burden consistently occurs when all the available budget is allocated to hotspots. However, this raises equity concerns as spatial targeting has considerable impact locally, in adjacent regions as well as in the overall population, but little or no impact on the remote region of the zone. Such equity concerns could be offset by improving passive case detection across the entire zone [151].

Although we predict that spatial targeting could be effective in a high burden setting, we may not have fully captured the mechanism leading to spatial clustering in TB incidence, which may result from intense localised transmission or aggregation of cases among groups of individuals sharing risk factors for progression [165, 166]. However, others have also argued that concentration of disease in high burden settings is mainly driven by localised transmission [5] and hence our study assumes transmission to be the predominant mechanism driving TB epidemiology. An improved understanding of drivers of spatial heterogeneity in TB incidence would be useful to inform targeted control interventions and hence data that describe the extent to which transmission accounts for the local TB epidemiology through the use of genotypic methods would further assist in the design of intervention studies.
6.6 Conclusions

In summary, our analysis suggests that spatial hotspot target strategy is efficient and cost-effective with the potential for significant reduction in overall TB burden.

6.7 Authors contributions

DS wrote the initial code, TND, JMT and ESM added additional lines of code. DS drafted the initial study concept and ESM, JTD and JMT refined this further. DS wrote the initial draft of the manuscript, and all authors provided input into revisions and approved the final draft and submission for publication.

6.8 Funding

No specific funding was received for this study. Debebe Shaweno is the recipient of the Melbourne International Research Scholarship and Melbourne International Fee Remission Scholarship. James Trauer is a recipient of an Early Career Fellowship from the NHMRC (APP1142638).

6.9 Competing interests

The authors declare that they have no competing interests.
6.10 References


6.11 Supplementary materials

Table S 3. Parameter and their ranges for sensitivity analysis

<table>
<thead>
<tr>
<th>Fixed input parameters</th>
<th>Value</th>
<th>References</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fast progression rate, $\varepsilon$</td>
<td>0.4 [0.31, 0.54]</td>
<td>[118]</td>
<td>Year$^{-1}$</td>
</tr>
<tr>
<td>Stabilisation rate, $\kappa$</td>
<td>3.6 [3.1, 5.1]</td>
<td>[118]</td>
<td>Year$^{-1}$</td>
</tr>
<tr>
<td>Reactivation rate, $\nu$</td>
<td>0.002 [0.0009, 0.004]</td>
<td>[118]</td>
<td>Year$^{-1}$</td>
</tr>
<tr>
<td>Natural recovery rate, $\gamma$</td>
<td>0.205 [0.15, 0.25]</td>
<td>[167]</td>
<td>Year$^{-1}$</td>
</tr>
<tr>
<td>Case detection rate, hotspots, $\delta_1$</td>
<td>0.65 [0.5, 0.95]</td>
<td>[168]</td>
<td>Proportion</td>
</tr>
<tr>
<td>Case detection rate, non-hotspots, $\delta_2$</td>
<td>0.60 [0.5, 0.75]</td>
<td>[168]</td>
<td>Proportion</td>
</tr>
<tr>
<td>Relapse, $\omega$</td>
<td>0.002 [0.003, 0.0025]</td>
<td>[169]</td>
<td>Year$^{-1}$</td>
</tr>
<tr>
<td>Effective contact rate, hotspots ($\beta_{11}$)</td>
<td>55 [10, 60]</td>
<td>Fitted to data</td>
<td>Year$^{-1}$</td>
</tr>
<tr>
<td>Effective contact rate, adjacent ($\beta_{22}$)</td>
<td>2.6 [1, 16]</td>
<td>Fitted to data</td>
<td>Year$^{-1}$</td>
</tr>
<tr>
<td>Effective contact rate, remote ($\beta_{33}$)</td>
<td>14 [1, 16]</td>
<td>Fitted to data</td>
<td>Year$^{-1}$</td>
</tr>
<tr>
<td>Coupling proportion, $\rho$</td>
<td>0.046 [0.0, 0.2]</td>
<td>Fitted to data</td>
<td>Proportion</td>
</tr>
</tbody>
</table>

All transitions between compartments were rates and case detection proportion was converted to case detection rate using the following formula:

$$
\delta_i = \frac{\mu_i (\mu_{di} + \gamma_i)}{1 - \mu_i}
$$

where $i$ may take two values representing spatial patch, $\mu_{di}$ represents TB related mortality, $\tau_i$ represents the intervention CDR (proportion) in the subdivision, and $\gamma_i$ refers to natural recovery of TB.
6.11.1 Cost-coverage curves

A generalised logistic function is used to calculate the coverage ‘cov’ associated with a specific spending ‘cost’ on an intervention [156]:

$$\text{Cov} = A + \frac{\text{Saturation} - A}{\left(1 + e^{-B \cdot \text{cost}}\right)^\alpha}$$

where saturation represents the maximum coverage possible; $\alpha$ is a parameter affecting the shape of the logistic curve (equal to 1 by default); and $A$ and $B$ are described below.

$$A = \frac{\text{Saturation}}{1 - 2^\alpha}$$

$$B = \frac{2^{\alpha+1}}{u \cdot p \cdot \alpha \cdot (\text{Saturation} - A)}$$

where $u$ denotes the unit cost and $p$ the size of the population to which the intervention could potentially apply.

Figure S9. Sensitivity of cost-coverage curves to a unit cost parameter

Figure S10 below shows sensitivity of impact of geographic funding allocation on population TB incidence to variations in unit cost per person screened. At USD 65,000 budget envelope, lower unit costs have higher population level impacts if the higher proportion of funding goes to spatial hotspots.
Figure S10. Sensitivity analysis of intervention impact unit cost per person screened.

Figure S11. Sensitivity of predicted TB incidence to changes in selected model parameters.
Chapter 7

7 Conclusions and future directions

This thesis proposed a novel method that can accurately identify spatial TB hotspots, and examined the role of spatial hotspots in the spread and control of TB in highly endemic settings. These objectives were achieved by integrating spatial analysis approaches with transmission dynamic models of TB. This general approach harmonises with the current global emphasis on optimising available tools for diagnosing and treating people infected with *Mtb* [30, 32, 139], including a need for more accurate identification and targeting of spatial clusters of TB transmission [140].

Although the idea of targeting spatial TB clusters is an attractive approach and could improve efficiency, there are important practical challenges in the identification and interpretation of epidemiologically relevant TB spatial clusters (Chapter 3). Thus, to ensure optimal resource allocation, there are important epidemiological questions that need to be answered:

- In the absence of complete incidence data, how should spatial TB clusters be identified?
- What is the role of spatial hotspots in the spatial spread of TB in highly endemic settings?
- Is spatial targeting an efficient strategy in such settings?

To address these questions, I integrated spatial analysis and TB transmission dynamic models. The spatial analyses adopted in this thesis include spatial cluster detection and spatial regression models, which span both frequentist and Bayesian frameworks. A Bayesian approach was adopted because there is no opportunity to unpack the drivers of spatial patterns in notification data under a frequentist framework. This is important as the spatial patterns in notification data could reflect either variations in TB case detection efforts or true underlying incidence. The Bayesian approach adopted in this thesis uses a hidden Markov model framework that offers a platform to understand the true underlying dynamics from the observations, which could vary according to local case detection efforts. Moreover, the Bayesian approach is also flexible, enabling the model to include spatially structured random effects, to smooth the effect of stochasticity and of unmeasured covariates.
In the subsequent sections, I present key lessons from this thesis, including how the studies described here have advanced methods in the spatial analysis of TB.

### 7.1 Lessons learned

#### 7.1.1 Limitations in approaches to spatial analysis of TB

This thesis described various methodological approaches used in spatial analysis, including spatial cluster detection methods and spatial modelling. Several spatial cluster detection methods are available, although information is lacking as to when to use each method. Applied to the same TB dataset, these methods often locate clusters at different spatial locations (Chapter 3). However, it can often be argued that each method is applied appropriately but is reflecting different interpretations, which poses practical challenges in using these methods to guide prevention and control efforts. In this regard, further studies are required to develop principles to guide method selection.

In contrast, the principles of spatial modelling have been well described in the literature, although adherence to these algorithms has been disappointing. In particular, the use of conventional regression methods that assume spatial independence without first assessing for spatial autocorrelation was common (Chapter 3) and can lead to inflated regression coefficients [78]. That means, associations may be identified even when they do not exist, as shown in Chapter 4. Conventional regression models may be appropriate for spatial analysis if spatial dependence in the residuals or dependent variable is ruled out. However, this approach was rarely reported, such that it is important to critically analyse findings from spatial analysis in light of the model selection approach presented in Section 1.1.3.

Almost all previous spatial analysis studies of TB depended largely on retrospective descriptions of spatial and temporal patterns, and there was no attempt to project the spatial patterns into the future (Chapter 3). In addition, in these studies, only attributes of the neighbourhoods where the individual reside at the time of TB diagnosis were incorporated into spatial analyses. Today, a wealth of evidence from molecular studies shows that more than 80 percent of TB transmission events in high endemic settings occur outside households – often in various social congregate settings [170, 171]. Therefore, it is likely that infections acquired from these settings could be misclassified as household transmission leading to misdirected interventions. Furthermore, an individual’s residence at the time of diagnosis may not reflect...
the place where exposure has occurred. This is particularly important because of the extended latency period of \textit{Mtb} infection and more so in areas where individuals are highly mobile (e.g. urban) that can result in exposure misclassification, biased estimates of association and imprecise risk estimation. In this regard, adoption of modern global positioning system (GPS)-defined longitudinal spatiotemporal exposure measurements or longitudinal residential information might be useful, although their practicality in high TB incidence settings remains challenging [172].

\subsection*{7.1.2 Data driven spatial TB heterogeneity}

Chapter 2 documented previously undescribed spatially heterogeneous TB distribution in a sparsely populated remote region of a highly endemic setting. This finding replicates spatial TB heterogeneity universally observed in different settings which is described in Chapter 3 raising the possible effectiveness of targeted interventions [137]. On the other hand, the mechanisms that drive these heterogeneities, which could arise from the characteristics of the infectious host, the organism, the susceptible host, the environment and program capacity that could lead to deferential case detection rates remain incompletely understood [151].

With reference to differential case detection rate, use of notification data in the spatial analysis of TB has been an established practice, although previous studies have not been able to account for undetected cases. Mainly due to lack of access to TB diagnostic and treatment centres, in 2017, notification data missed an estimated 36 percent of incident TB cases [4]. In line with this, the positive and strong correlation between health care access and TB burden in a remote region of Ethiopia (Chapter 2), which is likely to generalise to other highly endemic settings, illustrates the potential bias arising from use of notification data. That is, the spatial burden of TB appears greater in areas with access to a health care facility, implying that the spatial pattern does not necessarily reflect true underlying heterogeneity in case rates, but may rather reflect differences in case detection efforts [173]. Thus, priority setting, and resource allocation based solely on notification data that does not account for access and undetected cases could mask true incidence hotspots, and could even fuel community transmission by diverting attention and resources to the wrong places.

\subsection*{7.1.3 A novel approach to deal with data driven heterogeneity}

In addressing the potential bias in spatial analysis of TB due to differential case detection inherent in surveillance data, the novel Bayesian geospatial approach presented in this doctoral
thesis can serve as an important methodological development (Chapter 4). An interesting feature of this method is that it can estimate both incidence and case detection rates simultaneously by incorporating both an incidence model and an observation model. The current approaches to incidence estimation described in Section 1.1.3 are not sufficiently flexible to provide incidence estimates for different geographical resolutions and are less useful for monitoring local TB epidemiology across time and space. However, the ability of my model to predict incidence and case detection rates for any geographic and temporal resolutions is a major advantage over currently used incidence estimation tools and hence could be used to monitor local disease trends across space and time. Applied to Ethiopian TB surveillance data, this method demonstrated that notification data overlooks disease burden in areas with limited health care access, which is consistent with a positive correlation between notification data and health facility availability reported in Chapter 2. This approach is attractive for three reasons – 1) it accounts for missing TB cases, 2) it estimates incidence and case detection rates across various spatial and temporal units and so could be an alternative approach to incidence estimation, and 3) it allows smoothing of neighbourhood effects to provide more accurate disease incidence maps. As a result, this method has recently been acknowledged as a significant methodological advancement [174] and I hope that it will become a paradigm that can be used more routinely in the future analysis of surveillance data in high burden settings. The incidence estimate from this method is consistent with incidence estimate from the global TB disease burden estimates conducted by the Institute of Health Metrics and Evaluation (IHME) which also uses a Bayesian regression model [53].

In addition, the novel method proposed in this work is applicable to other infectious diseases as well. The approach can be used to track and monitor disease trends using data from patients who present to health facilities, although this would require implementation of standardised data collection formats.

7.1.4 Importance of mathematical modelling to capture spatial TB spread

Although several mathematical TB models exist, the link between spatial and mathematical models has been weak. Spatial epidemiologists and mathematical modellers have been working largely independently and thus, mathematical TB models often fail to account for universally documented spatial heterogeneity in TB distribution reported in Chapter 3. These mathematical models tended to describe who becomes infected and why they are infected but not where the infection occurs and where it is spreading. Although descriptions of who, when and why are
important to determine how resources can be targeted to treat cases and institute preventive measures, predicting where infection spreads is important in determining how best to place programmatic interventions to prevent further spatial spread. In this regard, the work presented in this doctoral thesis, integrates spatial analysis and mathematical models, and provides a potential source of TB infection and where the disease is spatially spreading. Such studies are particularly important to go beyond descriptions of past spatial patterns which are often the cases in spatial analysis approaches of TB (Chapter 3) to predictions of future course of spatial TB spread. Moreover, these studies are needed as population mobility is the main mechanism by which disease spreads across space and is ever increasing.

Studies in Chapters 2 and 4 show that TB incidence in a given location correlates strongly with its incidence in adjacent locations, likely reflecting spatial spread of infection and highlighting the potential effectiveness of spatially targeted interventions in such settings. The spatially structured mathematical TB model introduced in Chapter 5 was able to quantify the spatial spread potentially driving the spatial correlation in TB incidence reported in the two studies described above. In modelling the spatial spread of infectious diseases, determining the spatial contact pattern is challenging and has been a major barrier, as detailed data on specific contact structures are typically unavailable.

Apart from projecting the future behaviour of epidemics, mathematical models also offer a means to reconstruct contact patterns from existing data. In this regard, the model in Chapter 5 uses existing TB surveillance data to reconstruct spatial contact patterns, which is vital to determine pathways of TB diffusion that would help to design effective control strategies. This was achieved by fitting spatially structured mathematical models to TB surveillance data, such that best-fit solutions were used to estimate spatial contact patterns. Validation of such models would require collection of additional data on population mobility or contact patterns in the study region, although this is often infeasible. One relatively straightforward approach is to assess whether different diseases will lead to similar inter-regional contact (mobility) patterns, where similarity justifies the use of existing data sources to estimate contact patterns [114]. However, a caveat to this prescription is that the relevant contact patterns will be specific to particular infections due to various reasons and consequently locations of spatial clusters of different diseases are often variable. Thus, in understanding the spatial spread of infectious diseases including TB, development of efficient approaches that would generate local contact patterns are important contributions.
In developing spatially structured metapopulation models to understand the spatial spread of infectious disease there is no consensus on how many spatial structures to consider. The studies presented in Chapters 5 and 6 used three spatial subdivisions to effectively capture the waning spatial autocorrelation as distance from spatial hotspots increases, in contrast to previous studies that compared two spatial structures [137, 138]. Limited by data availability, the model used three spatial groups, which may be overly simplistic, but with additional spatial structures the model could become complex and lose the advantages associated with simpler meta-population models. In addition, in the absence of adequate information about contact patterns between spatial groups, models with a large number of spatial patches may be unable to yield realistic estimates. Generally, more complex models, such as spatially continuous models or agent-based models, are useful to investigate the spatial spread at finer spatial scales.

As spatially structured mathematical TB modelling is a newly developing research field that gained attention only recently, there is also no consensus on how to define spatial TB hotspots. In some previous studies, spatial clusters were taken to indicate high incidence regions, although an unusual disease incidence in a given location may not necessarily imply spatial hotspots [137, 138]. In contrast, spatial hotspots used in this thesis were determined by using spatial cluster detection technique Chapter 5 that uses a statistical test that accounts for potential spatial dependence [80]. The application of spatial cluster detection methods to identify meaningful spatial clusters and further to quantify the interaction amongst the three meta-populations is a powerful tool. The estimates found in the inference framework described above, can be used –as was done in this thesis to reliably inform simulation models of targeted interventions aimed at hotspots and hence to estimate the impact of targeting such hotspots.

The thesis demonstrates that spatial hotspot targeting in Ethiopia can be an efficient strategy. Although such spatial targeting could maximise the impact of interventions within a given budget range, resource allocation to such locations has raised economic and equity concerns. Thus, while allocating resources to high burden locations, policy makers need to ensure that other individuals outside the hotspots do not face inappropriate disparities in health care access, and financial burden to access care [151].

7.2 Future directions and limitations

This doctoral thesis has several potential limitations. Even when the missing data problems are resolved using alternative data sources such as prevalence surveys and a novel method
presented in this thesis, the identified spatial clusters of TB may not necessarily mean that transmission is driving the observed pattern. In this thesis I did not verify the extent to which transmission drives the dynamics of TB in geospatial hotspots, although generally transmission is thought to account for most new cases in high burden settings [5]. It would be ideal if genotypic data were used to quantify the extent to which transmission drives TB dynamics in the hotspots as well as in the community for effective public health interventions. In addition, no mobility data were available in the study region to quantify population spread and should be considered in the future studies. Moreover, notification data used in this thesis contains lots of clinically detected cases, some of which might not be TB cases. All analyses presented in this thesis do not account for such potential misclassification.

In this thesis the role of spatial hotspots in the spatial spread of TB was determined based on spatial hotspots identified using TB data aggregated at the kebele level. Given the absence of individual-level TB data, it was not possible to evaluate the impact of spatial scale on the stability of these hotspots. In addition, the classification of the study region into three spatially discrete groups may be overly simplistic, although increasing the number of patches simulated would present additional challenges in fitting many parameters. More complex models, such as spatially continuous models or agent-based models, may be useful to investigate these links at finer spatial scales and to overcome the role of spatial scale.

Almost all spatial analysis of TB conducted so far depended on describing past TB epidemics using retrospectively collected data (Chapter 3). There has been no attempt to project potential spatial TB patterns into the future beyond the time of the study. Predictions of the future spatial course of TB epidemics can possibly be achieved by linking spatial data to transmission dynamic models or statistical forecasting models, although validation of such models is very challenging. Another study limitation is that, limited by data availability, transmission dynamic models presented in this thesis do not consider multidrug-resistant TB, HIV and age, and these will be the subject of the future research.

Overall, this thesis has highlighted the importance of accounting for missing TB cases in spatial analysis so as to develop spatially targeted effective public health responses. Previous spatial analyses of TB aimed at identifying spatial TB clusters for spatial targeting predominantly used surveillance data. These studies were clearly informative of the underlying spatial TB distribution in low endemic settings. However, as has been shown in
this thesis, spatial analysis of TB using surveillance data in highly endemic settings has been less informative as teasing out the role underlying incidence and variations in case detection efforts has remained a challenge. The method introduced in this thesis advances current approaches to spatial analysis and provides a means to account for problems of undetected cases. GIS based spatial analysis can be key to policy formulation if data used reflects the actual underlying incidence. In this regard, data sources other than notifications including incidence estimates from our novel method could be quite useful for effective public health responses.

Although identification of spatial hotspots is a useful initial step, characterising what constitutes a spatial TB hotspot in specific context is critical in designing effective public health responses. This is important because spatial TB hotspots are not always caused by transmission and hence interventions that aim to interrupt disease spread in the identified hotspots may not achieve the intended goals. Genotypic methods can be used to characterise the drivers, but these tools are not available in highly endemic settings.

Importantly, the study also concludes that the role of spatial TB hotspots in the remote regions of Ethiopia is limited and transmission is predominantly local. Hence, interventions strategies that are spatially targeted may not achieve anticipated outcomes, although the overall effect of these interventions is considerable due to extremely high prevalence in hotspot regions.
References

12. Murray J: The Industrial Revolution and the decline in death rates from tuberculosis. The international journal of tuberculosis and lung disease: the official


149. DIC: Deviance Information Criteria [https://www.mrc-bsu.cam.ac.uk/software/bugs/the-bugs-project-dic/]


173. Dangisso MH, Datiko DG, Lindtjorn B: **Accessibility to tuberculosis control services and tuberculosis programme performance in southern Ethiopia.** *Glob Health Action* 2015, 8:29443.
