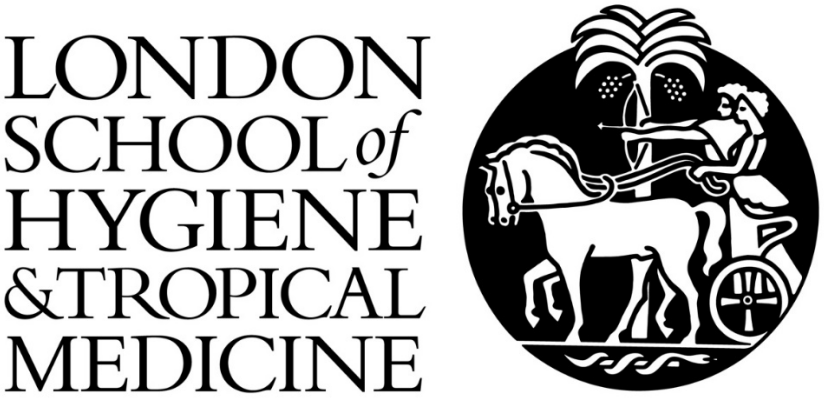


Investigating *Mycobacterium tuberculosis* transmission in rural Malawi

Palwasha Yousafzai Khan

Thesis submitted in accordance with the requirements for the degree of

Doctor of Philosophy

University of London

August 2017

Department of Infectious Disease Epidemiology

Faculty of Epidemiology and Population Health

London School of Hygiene & Tropical Medicine

Funded by the Wellcome Trust: Clinical PhD Training Fellowship (100137/Z/12/Z)

Declaration of work

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

Role of candidate

The idea of using children with incident *Mycobacterium tuberculosis* infection as 'sentinels' of recent transmission in the Karonga demographic surveillance site belonged to Mia Crampin. I wrote the study protocols, designed the questionnaires and led the fieldwork for the tuberculin skin test (TST) cohort and the tuberculosis case-contact household study, with input from Judith Glynn and Mia Crampin. The idea of collecting data prospectively on attendance at gathering places as part of the TST cohort study was mine. I also conceived the idea of using the Rust and Thomas method to estimate *Mycobacterium tuberculosis* infection prevalence using the data from the TST cohort study and the tuberculosis case-contact household study, and conceptualised the idea of examining the effect of antiretroviral treatment on the infectiousness of smear-positive pulmonary tuberculosis in the tuberculosis case-contact household study. I was responsible for all the data management with assistance from Jacky Saul and Keith Branson and I was responsible for all the analyses, with statistical input from Katherine Fielding. With regards to the papers for publication, I was joint first co-author of the review paper and sole first author of all the other four papers, and wrote the initial and final drafts of each of these.

Acknowledgments

I would like to express my sincere gratitude to my supervisor, Professor Judith Glynn, for her support and guidance from my time as an MSc student through to the end of my PhD. I am also grateful to Mia Crampin for her input whilst out in Chilumba. I would like to thank the Wellcome Trust for sponsoring my research training, overseas placement and fieldwork. In addition, I would like to thank Professors David Mabey and Alison Grant for their continued support throughout the Clinical PhD programme and for encouraging me to pursue an academic career.

I am hugely indebted to Dr Katharina Kranzer and Professor Rashida Ferrand for their tremendous support and friendship, as well as the scientific input they have provided along the way. I am also very grateful to Katherine Fielding for her outstanding statistical support, and to Professor Paul Fine for his wise counsel and for being the perfect devil's advocate. I would like to thank my colleagues and friends, Ms Jacky Saul and Mr Keith Branson, without whom I may never have got to grips with the KPS data management. I also wish to thank Mr Yoryos Ponnighaus and Ms Tamara Hurst for their logistics and operational support without which I would have been at sea, and Bill Corner for providing essential formatting input.

This research would not have been possible without the participation of the community in Chilumba and the superb professionalism and hard work of KPS fieldworkers and data officers, to all of whom I am deeply grateful, especially Dominic Mulawa, Regina Chiumya, Ephrida Mwiba and Taniel Njawala. I would also like to thank Dr Olivier Koole, Dr Alison Price and Ms Estelle McLean, whose company and camaraderie in Chilumba was truly valued.

My heartfelt thanks to my parents and brothers (although I know it does not need to be said…) and to Anton, who I am sure is relieved that this *magnum opus* is finally done. Little does he realise that it has only just become!

Abstract

Current control strategies are failing to contain the tuberculosis (TB) epidemic and are limited by our lack of understanding of *Mycobacterium tuberculosis (M.tb)* transmission dynamics, especially in high HIV prevalence settings. *M.tb* infection in children aged under 5 years of age indicates recent transmission, acting as a sentinel for infectious (typically adult) TB and highlights recent failures in community control measures. The overall aim of this research, which was based in a high HIV prevalence rural community in northern Malawi, was to delineate *M.tb* transmission events occurring within the context of a well-implemented TB control programme, and thereby elucidate factors driving transmission that are not being addressed by current control strategies. This thesis presents findings from a series of linked studies, including a longitudinal tuberculin skin-test (TST) study of pre-school children in an area under demographic surveillance, and a household contact study of smear-positive TB cases.

Estimates of the average annual risk of *M.tb* infection (ARTI) in this population of young BCG-vaccinated children varied widely depending on the method used to estimate infection prevalence. A previously overlooked method of estimating *M.tb* infection prevalence, initially published by Rust and Thomas in 1975, appeared to be only method to appropriately adjust for the marked effect of BCG-attributable induration in the youngest children (aged <2 years). Marked differences in the estimates of the risk of *M.tb* infection when using cross-sectional data compared to using longitudinal data are also highlighted.

Age, known contact with a smear-positive TB case and community *M.tb* exposure (defined as the average notification rate of smear-positive TB per 100,000 population) were risk factors for prevalent and incident *M.tb* infection in children. Being HIV-exposed *in utero* was the strongest risk factor for prevalent infection, whilst having an HIV-positive father was strongly associated with incident infection. Additional risk factors for incident infection included church attendance and travel on mini-buses. No evidence was found in the household contact study that smear-positive tuberculosis patients on antiretroviral treatment (ART) were more likely to transmit *M.tb* infection to household child contacts (as inferred from TST positivity) compared to smear-positive tuberculosis patients not on ART. However, child contacts of HIV-negative individuals had nearly three times the odds of having a positive TST compared to child contacts of HIV-positive TB patients not on ART; this was partly explained by differences in the degree of smear positivity.

HIV-related risk factors for prevalent and incident *M.tb* infection in pre-school children highlight that interventions, such as screening for HIV and TB with implementation of isoniazid preventive therapy where indicated within the household of HIV-positive individuals may further reduce the burden of TB at a community level in the longer term. Findings from this research also highlight the need for better infection control practices (improved ventilation) in congregate settings such as churches and mini-buses. There is an urgent need for improved ongoing surveillance to guide the implementation of context-specific TB control strategies.

Acronyms

AIDS	acquired immunodeficiency syndrome
ACF	active case-finding
aOR	adjusted odds ratio
ART	antiretroviral treatment
ARTI	average annual risk of *M.tb* infection
ATS	American Thoracic Society
BCG	bacille Calmette-Guerin
CI	confidence interval
DNA	deoxyribonuclease acid
DOTS	directly observed treatment, short-course
DSS	demographics surveillance site
EM	expectation maximisation
FAST	Finding TB cases Actively, Separating safely, Treating effectively
GPS	global positioning system
HH	household
HIV	human immunodeficiency virus
IGRA	interferon gamma release assay
IQR	interquartile range
IRR	incidence rate ratio
IPT	isoniazid preventive treatment
KHDSS	Karonga Health demographic surveillance site
KPS	Karonga Prevention Study
MDR-TB	multidrug resistant tuberculosis
MEIRU	Malawi Epidemiology and Interventions Research Unit
MIRU-VNTR	mycobacterial interspersed repetitive units-variable number of tandem repeats
NTM	non-tuberculous mycobacteria
NTP	National Tuberuclosis Programme
OR	odds ratio
PAF	population attributable fraction
PCR	polymerase chain reaction

PPD	purified protein derivative
PY	person-years
RLFP	restriction fragment length polymorphism
RR	relative risk/ rate ratio/ risk ratio
SD	standard deviation
SES	socioeconomic status
SNP	single nucleotide polymorphisms
TB	tuberculosis
TST	tuberculin skin test
UCL	University College London
WGS	whole genome sequencing
WHO	World Health Organisation
ZAMSTAR	Zambia South Africa Tuberculosis and AIDS reduction

Table of Contents

1. Introduction

1.1 Tuberculosis control, elimination and eradication

Tuberculosis (TB), caused predominantly by the airborne transmission of the bacillus *Mycobacterium tuberculosis* (*M.tb*),[1] has afflicted humans from prehistoric times,[2] reaching epidemic proportions in Europe and North America during the 18th and 19th centuries.[3] Dramatic declines in TB mortality and disease were evident by the early 20th century, predating the availability of effective treatment.[4,5] Improvements in living and social conditions which occurred in the early 1900s are thought to have led to a decrease in transmission by reducing the average number infected by each infectious TB case,[4] although a comprehensive explanation for this decline remains elusive.[3,6]

By the mid-1930s, in the pre-chemotherapy era, the sustained decline prompted Wade Hampton Frost, the 'father of modern epidemiology',[7] to make the statement:

> "…the eventual *eradication of tuberculosis* only requires that the present balance against it be maintained".

This statement was tempered by the additional assertion that:

> "[There are]… only two forces, which singly or together, would check or reverse the striking downward trend of TB. These are: (1) a decrease in human resistance to the disease, or (2) some fundamental change in the adaptation of the tubercle bacillus to its host, tending to favor survival of the parasite."[8]

Regrettably the 'two forces' foretold by Frost that *could* potentially halt the eventual eradication of TB *may well* have materialised in the form of the epidemic spread of the retrovirus, Human Immunodeficiency Virus (HIV) in the 1980s,[9] and the evolution and global dissemination of drug-resistant strains of *M.tb.*[10] However, these are not the only two factors limiting our ability to eradicate *M.tb* today. Failure of case finding and poor treatment adherence hinder the effectiveness of early case detection and effective TB treatment, which are the cornerstone of global TB control strategies.

Despite significant progress through the Directly Observed Treatment, Short-course (DOTS) strategy in the 1990s and the Stop TB strategy since 2006,[11] we have been unable to mitigate the disastrous impact of the HIV pandemic on global TB control predicted decades earlier,[9,12,13] which continues to overwhelm health systems today, especially in sub-Saharan Africa.[14,15] Case detection and treatment completion have been the cornerstone of global TB control, yet evidence points to continued person-to-person *M.tb* transmission, even in settings with well-implemented national TB control programmes.[16-18] The year 2015 marked a transition in global TB control with a step-up from the Stop TB strategy to the ambitious End TB strategy, affirming the ultimate goal of eliminating TB[i] by 2050.[19] It is worth noting that in order to achieve this formidable task, reductions in global TB incidence in the order of ten times the current rate of decline are required.[20]

[i] Elimination is defined as achieving an incidence of < 1 case of all forms of TB per 1,000,000 population per year

This has led to a renewed focus on additional control interventions,[21] which include active case finding (ACF) as a means to interrupt transmission,[22] improved diagnosis and treatment of established *M.tb* infection thereby 'controlling the seedbeds' of future cases,[23] and a greater emphasis on a biosocial approach and sustainable development to put an end to the TB epidemic.[24] Some of these strategies have a robust evidence-base whereas others have weaker grounds for advocating wider implementation.[25] To inform which components should be part of a maximally effective complement of interventions we need a better understanding of *M.tb* transmission dynamics, especially in high HIV burden settings.

1.2 Knowledge gaps in M.tb transmission research

Significant knowledge gaps in *M.tb* transmission research remain despite decades of scientific investigation. Unfortunately, improved understanding of *M.tb* transmission dynamics as a research priority has been underfunded and was excluded from the research and development component of the 2011-2015 Global Plan to Stop TB.[26]

The lack of recognition of *M.tb* transmission research as a research priority area by funders, was one of the reasons that I and a fellow PhD student at University College of London (UCL), Tom Yates, organised an *M.tb* Transmission meeting, which was held in London in November 2014. The event was co-funded by the TB Modelling and Analysis Consortium, UCL Population Health Domain, and the TB Centre at the London School of Hygiene & Tropical Medicine. This meeting led

to a publication of a Review article on *M.tb* transmission in high burden settings, which is presented in Chapter 2. The agenda for the meeting and the list of attendees are included in Appendix I.

Knowledge gaps in *M.tb* transmission research will be covered in detail in Chapter 2 as part of the review. The research priorities which are specifically addressed by the work presented in this thesis include:

i. measuring *M.tb* transmission at a population-level
ii. identifying drivers of recent *M.tb* transmission
iii. providing insight into where *M.tb* transmission occurs in the community
iv. improving our understanding of the effect of HIV and antiretroviral therapy on *M.tb* transmission dynamics

1.3 Research setting

The Karonga Prevention Study, a research group collaborating with the Malawi National Tuberculosis Programme (NTP), has been supporting core TB programme activities and conducting population-level TB epidemiological studies in Karonga district, northern Malawi, since 1986.[27]

Significant findings from previous TB epidemiological studies conducted by the research group include:

- failure of BCG vaccination to protect against adult pulmonary TB, despite protecting against leprosy[28]

- estimation that about two thirds of active TB arises from recent infection[29]
- HIV-associated TB mostly follows recent *M.tb* infection, irrespective of pre-existing latent infection[30]
- recognisable recent contact is responsible for about 10% of disease[31-33]
- proportion of TB cases due to recent transmission has decreased over time, as has the proportion of TB cases transmitting and giving rise to new cases[34]
- drug resistance has remained constant (<10%) over more than 20 years,[35] an indicator of the existence of a well-implemented TB control programme in the district.

Adult HIV prevalence was 4% in 1988-90, reaching a peak of around 13% in early 2000,[27] and has currently plateaued around 9%.[35] Figure 1 is an extension of the original TB incidence analysis of all new episodes of smear-positive pulmonary TB in adults in Karonga District from 1988 to 2011 published in 2013.[35] For the purposes of this thesis, the candidate has updated the analysis with the most recently available data. Adult population denominators for Karonga District were acquired from KPS whole population survey data for the 1980s and from national censuses from 1998 to 2008.[35] An assumption of exponential growth of 3% per annum from 2008 to 2017 was used based on the observed population growth between censuses in 1998 and 2008. The figure illustrates the peak in the incidence of new smear-positive pulmonary tuberculosis in the late 1990s at around 140 per 100,000 population per year and then the continued steady

decline to levels below those seen in the earlier years of the HIV epidemic to around 60 per 100,000 population per year in 2016.

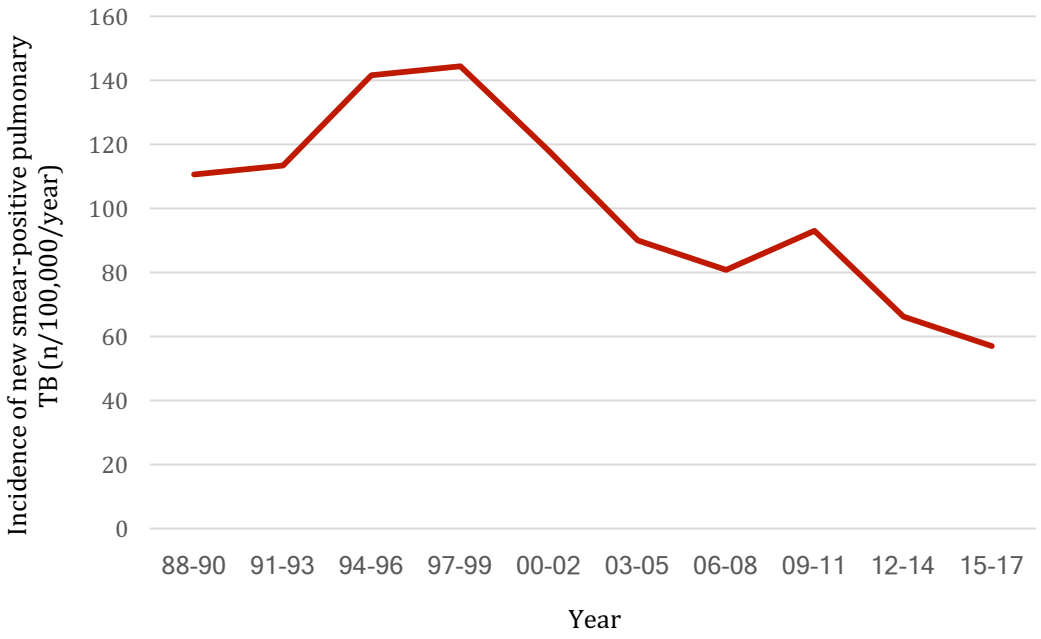
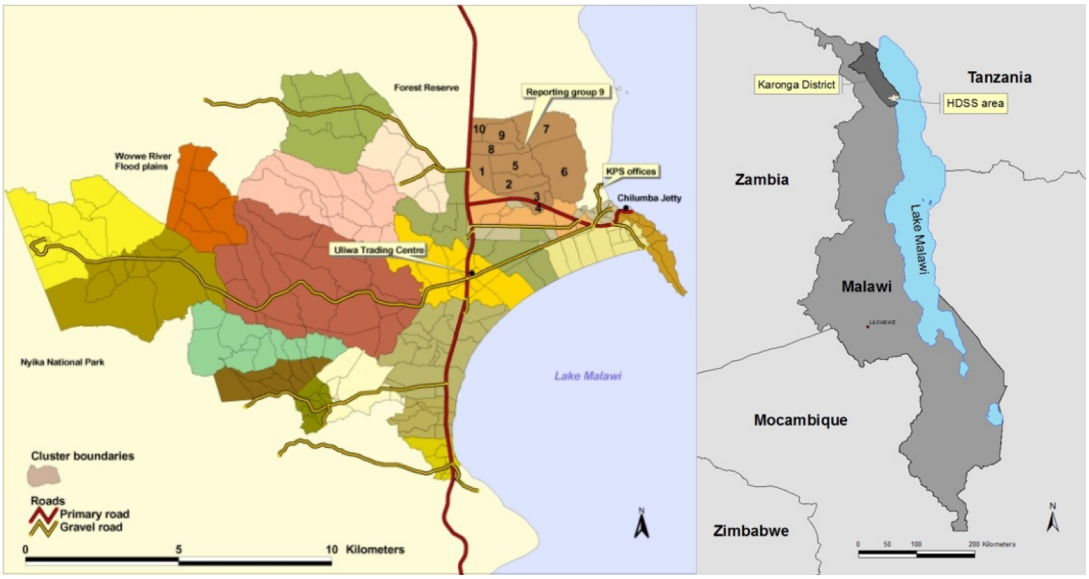


Figure 1. Incidence of new smear-positive pulmonary TB in adults (per 100,000 population/year) in Karonga district from 1988-90 to 2015-17

The greater part of the fieldwork for this research was undertaken in the Karonga demographic surveillance site (DSS: population 39,000). See Figure 2. The demographic surveillance site was set up in the south of the district in 2002, primarily to provide a platform for epidemiological studies of HIV and HIV-associated infectious disease and to monitor the impact of interventions such as antiretroviral therapy (ART).[36]



HDSS health demographic surveillance site

Figure 2. Map of the DSS within Malawi in relation to Tanzania, Zambia and Mozambique (Source: MEIRU/KPS)

BCG vaccination was introduced into the district initially among all ages in a vaccine trial in the mid 1980s. It has been part of the routine Expanded Programme on Immunisation since 1990, and is given at first health systems contact.[27]

Conducting *M.tb* transmission research within the context of a well-implemented TB control programme in a high HIV prevalence setting, such as Malawi, is ideal. It has enabled the elucidation of factors driving community *M.tb* transmission that are not being addressed by current prevention strategies. An in-depth understanding of local TB epidemiology made possible by preceding research has also greatly facilitated detailed interpretation of research findings.

1.4 Study rationale: children as 'sentinels'

The potentially long latent period of *M.tb* infection and the inability to distinguish between recent and historic infection complicates studies examining *M.tb* transmission dynamics.[37] It was therefore decided to undertake this research primarily using pre-school children (aged under 5 years) based on the logic that *M.tb* infection in the very young necessarily results from recent transmission. *M.tb* infection in pre-school children acts as a sentinel of the presence of infectious adult tuberculosis,[1,38] thus having the potential to highlight failure of contemporary TB control measures. [1,39] Identifying recent *M.tb* infection in this subset of the population also increases the chance of being able to pinpoint the source of new infections in space, as young children spend more time near their home and close family than do adults.

The concept of using young children as 'sentinels' to identify potential sources of community *M.tb* transmission is not new. A case-finding study published by Griffiths *et al.* in 1963 used serial annual tuberculin-testing of school children in Cardiff to detect children who were recent 'tuberculin convertors' and went on to trace the source of their infection by radiographically screening all household contacts of 1000 child convertors, identifying 183 contacts with previously undiagnosed tuberculosis.[40] Even as early as 1930, Kielland suggested, in a letter to the *Journal of the Norwegian Tuberculosis Association,* the systematic use of repeated tuberculin tests for early diagnosis and case finding, using tuberculin-positive school children as the starting point for case finding.[41]

1.5 Conceptual framework and definitions of terms used in thesis

Our understanding of the natural history following *M.tb* exposure to established *M.tb* infection through to disease remains open to question.[42-44] However a number of risk factors resulting in established *M.tb* infection and disease following *M.tb* exposure have been identified. Figure 3 (based on a conceptual hierarchical framework as suggested by Victora *et al.*)[45] provides a simplified schematic of the risk factors at the individual, household, societal and environmental level which affect progression at different points along the continuum from *M.tb* exposure through to established infection, disease and ultimately to cure or death from TB disease. Those risk factors relevant to this research, namely those resulting in established *M.tb* infection, have been highlighted in the box with a dashed blue line. Risk factors for prevalent and incident *M.tb* infection which were examined as part of this work are featured in blue font in the figure.

Definition of terms used in this thesis are presented in the Table 1.

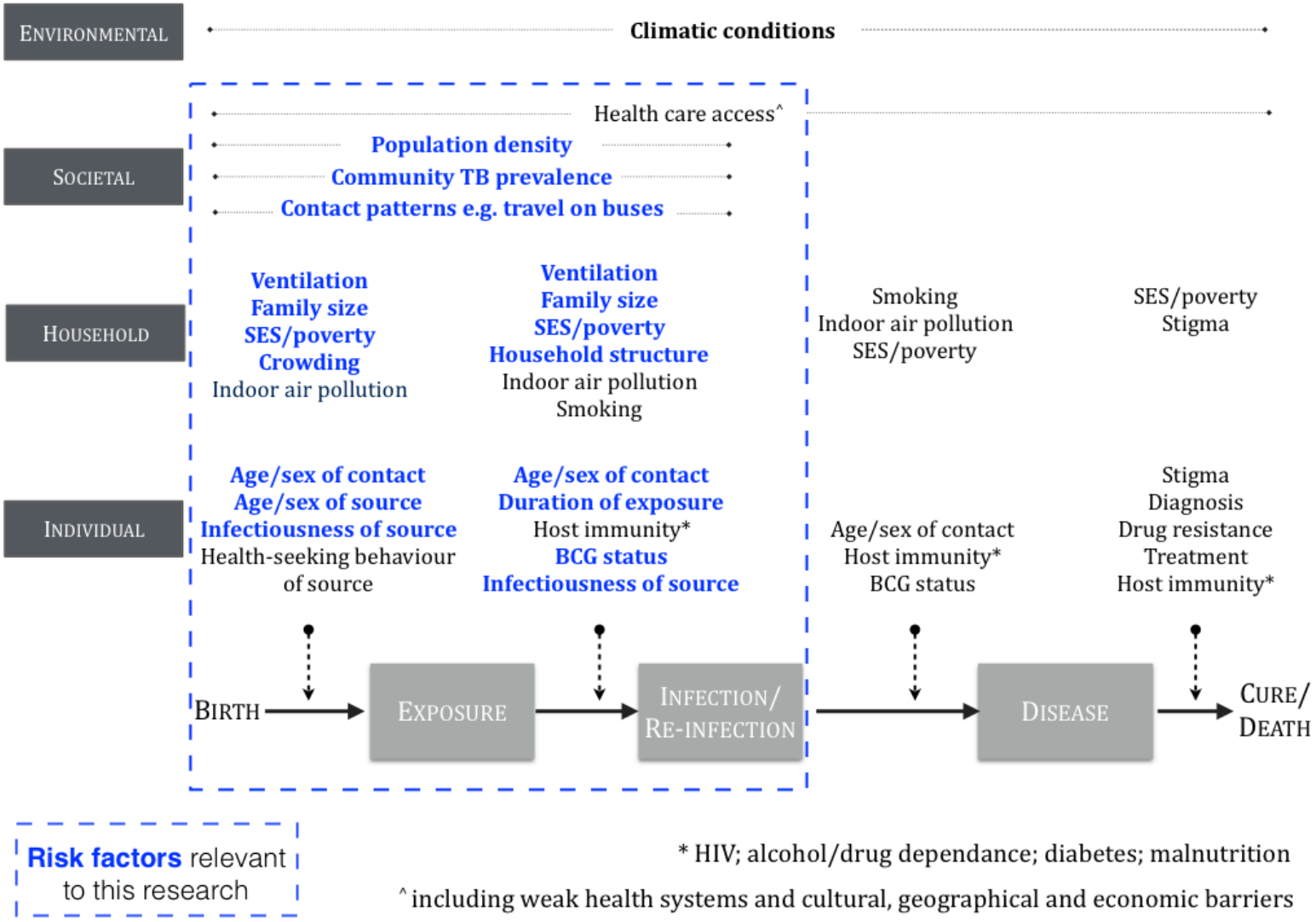


Figure 3. Conceptual framework illustrating risk factors associated with transition from each state from exposure to *M.tb* exposure through to infection and disease [46,47]

Term	Definition	Comment
Infectiousness	Propensity to transmit *M.tb*	This relates to source case characteristics, e.g. active disease *vs* subclinical, presence of lung cavitation *vs* none, bacillary load, HIV status, health-seeking behaviour, and whether on *effective* treatment or not (and not to contact characteristics, e.g. susceptability to infection following exposure) [48]
Recent transmission	*M.tb* transmission which has occurred within the last 5 years	For the purposes of this thesis, all *M.tb* transmission which has occurred within the last 5 years is defined as *recent*, using children aged under 5 years as sentinels of *M.tb* transmission
Exposure	Exposure to *M.tb* bacilli	For the purposes of this thesis, exposure is defined as contact with diagnosed smear-positive pulmonary TB. [49, 50] In the community, this is quantified as an average notification rate of smear-positive pulmonary TB per year per residential area and within the household as proximity to index case, e.g. resident within the same household, slept in the same room, slept in the same bed etc.)
Close contact	Household contact with diagnosed TB case	For the purposes of this thesis, close contact is defined as any form of household contact (including non-residential household contact) with a diagnosed TB case, e.g. a child staying with a grandparent for a 2-week period [50]
Congregate setting/gathering place	Places where >10 persons within a community come together, e.g. at a funeral, in a church, in a mini-bus	Congregate setting/gathering places examined in this research do not include long-term institutionalised congregate settings such as correctional facilities, schools, care homes or refugee camps.

Table 1. Definitions of terms used in thesis

1.6 Research aim and objectives

The overall aim of the research was to investigate *M.tb* transmission in a rural, high HIV prevalence setting using classical field epidemiological techniques.

The specific objectives were to:

1. Estimate the prevalence and annual risk of *M.tb* infection in pre-school children as a measure of recent *M.tb* transmission
2. Identify risk factors for prevalent and incident *M.tb* infection in pre-school children
3. Identify potential locations of recent *M.tb* transmission in the community
4. Assess the effect of ART on the 'infectiousness' of HIV–positive adults with smear-positive pulmonary TB

1.7 Thesis outline

The format of this thesis is that of a 'research paper style' dissertation. The thesis is composed of manuscripts that have been published, accepted for publication, or are ready for submission, and are bookended by an introduction and discussion section. All papers are formatted according to journal requirements. The chapters are as follows:

Chapter 1 (this chapter) introduces the area of research, the starting point that led up to the project, the study rationale including the research aim and objectives and the outline of the 'research-style' thesis.

Chapter 2 presents the literature review undertaken for this thesis which is a published review paper (joint first author) "**The transmission of *Mycobacterium tuberculosis* in high burden settings**" published in Lancet Infectious Diseases. *Reference*: Yates TA, Khan PY, Knight GM, Taylor JG, McHugh TD, Lipman M, White RG, Cohen T, Cobelens FG, Wood R, Moore DA, Abubakar I. Lancet Infect Dis. 2016 Feb; 16(2):227-38.

Chapter 3 consists of the manuscript titled "**Challenges in the estimation of the annual risk of *Mycobacterium tuberculosis* infection in children aged under 5 years**" published in the American Journal of Epidemiology*. Reference*: Khan PY, Glynn JR, Mzembe T, Mulawa D, Chiumya R, Crampin AC, Kranzer K, Fielding KL. Am J Epidemiol. 2017 May 19. doi: 10.1093/aje/kwx153. [Epub ahead of print].

Chapter 4 is the published paper "**Risk factors for *Mycobacterium tuberculosis* infection in 2-4 year olds in a rural HIV-prevalent setting**" published in the International Journal of Tuberculosis and Lung Disease. *Reference*: Khan PY, Glynn JR, Fielding KL, Mzembe T, Mulawa D, Chiumya R, Fine PE, Koole O, Kranzer K, Crampin AC. Int J Tuberc Lung Dis. 2016 Mar; 20(3): 342-9.

Chapter 5 presents a draft paper **"Incident *Mycobacterium tuberculosis* infection in young children identifies risk factors for recent community transmission in rural Malawi."** This paper is comprised of a risk factor analysis including potential locations of recent community *M.tb* transmission.

Chapter 6 is comprised of the manuscript (*in press*) titled **"Does antiretroviral treatment increase the infectiousness of smear-positive pulmonary tuberculosis?"** which examines the prevalence of *M.tb* infection in child household contacts of adult smear-positive pulmonary TB cases by index case HIV/ART status. Accepted for publication in the International Journal of Tuberculosis and Lung Disease on the 28th July 2017. *Reference:* Khan PY, Crampin AC, Mzembe T, Koole O, Fielding KL, Kranzer K, Glynn JR. Int J Tuberc Lung Disease. 2017 (*in press*).

The discussion, main conclusions, recommendations for further research and potential interventions strategies are presented in **Chapter 7**.

1.8 References

1. Rieder HL. *Epidemiological basis of tuberculosis control.* Paris: International Union Against Tuberculosis and Lung Disease; 1999.
2. Donoghue HD, Spigelman M, Greenblatt CL, et al. Tuberculosis: from prehistory to Robert Koch, as revealed by ancient DNA. *Lancet Infect Dis.* 2004;4(9):584-592.
3. Daniel TM. The history of tuberculosis. *Respiratory medicine.* 2006;100(11):1862-1870.
4. Vynnycky E, Fine PE. Interpreting the decline in tuberculosis: the role of secular trends in effective contact. *Int J Epidemiol.* 1999;28(2):327-334.
5. Styblo K, Meijer J, Sutherland I. [The transmission of tubercle bacilli: its trend in a human population]. *Bull World Health Organ.* 1969;41(1):137-178.
6. McKeown T, Record RG. Reasons for the decline of mortality in england and wales during the nineteenth century. *Population Studies.* 1962;16(2):94-122.
7. Daniel TM. *Wade Hampton Frost, pioneer epidemiologist 1880-1938 : up to the mountain.* Rochester, N.Y.: University of Rochester Press ; [Woodbridge : Boydell & Brewer, distributor]; 2005.
8. Frost WH. How Much Control of Tuberculosis? *American journal of public health and the nation's health.* 1937;27(8):759-766.
9. WHO/IUATLD. Tuberculosis and AIDS. Statement on AIDS and tuberculosis. Geneva, March 1989. Global Programme on AIDS and Tuberculosis Programme, World Health Organization, in collaboration with the International Union Against Tuberculosis and Lung Disease. *Bull Int Union Tuberc Lung Dis.* 1989;64(1):8-11.
10. Keshavjee S, Farmer PE. Tuberculosis, drug resistance, and the history of modern medicine. *N Engl J Med.* 2012;367(10):931-936.
11. WHO. Global tuberculosis control: WHO report 2011. http://www.who.int/tb/publications/global_report/2011/gtbr11_main.pdf. Geneva: World Health Organisation; 2011.
12. Styblo K. The potential impact of AIDS on the tuberculosis situation in developed and developing countries. *Bull Int Union Tuberc Lung Dis.* 1988;63(2):25-28.
13. Schulzer M, Fitzgerald JM, Enarson DA, Grzybowski S. An estimate of the future size of the tuberculosis problem in sub-Saharan Africa resulting from HIV infection. *Tuber Lung Dis.* 1992;73(1):52-58.
14. Harries AD, Zachariah R, Corbett EL, et al. The HIV-associated tuberculosis epidemic--when will we act? *Lancet.* 2010;375(9729):1906-1919.
15. Lawn SD, Harries AD, Williams BG, et al. Antiretroviral therapy and the control of HIV-associated tuberculosis. Will ART do it? *Int J Tuberc Lung Dis.* 2011;15(5):571-581.
16. Murray EJ, Marais BJ, Mans G, et al. A multidisciplinary method to map potential tuberculosis transmission 'hot spots' in high-burden communities. *Int J Tuberc Lung Dis.* 2009;13(6):767-774.

17. Godfrey-Faussett P, Sonnenberg P, Shearer SC, et al. Tuberculosis control and molecular epidemiology in a South African gold-mining community. *Lancet.* 2000;356(9235):1066-1071.
18. Nardell E, Churchyard G. What is thwarting tuberculosis prevention in high-burden settings? *N Engl J Med.* 2011;365(1):79-81.
19. WHO. Global tuberculosis control: WHO report 2015. http://apps.who.int/iris/bitstream/10665/191102/1/9789241565059_eng.pdf?ua=1. Geneva: World Health Organisation; 2015.
20. WHO. End TB Strategy. Draft global strategy and targets for tuberculosis prevention, care and control after 2015. *Documentation for World Health Assembly 67.* http://apps.who.int/gb/ebwha/pdf_files/WHA67/A67_11-en.pdf. Geneva: World Health Organisation; 2014.
21. Das P, Horton R. Tuberculosis--getting to zero. *Lancet.* 2015;386(10010):2231-2232.
22. Yuen CM, Amanullah F, Dharmadhikari A, et al. Turning off the tap: stopping tuberculosis transmission through active case-finding and prompt effective treatment. *Lancet.* 2015;386(10010):2334-2343.
23. Rangaka MX, Cavalcante SC, Marais BJ, et al. Controlling the seedbeds of tuberculosis: diagnosis and treatment of tuberculosis infection. *Lancet.* 2015;386(10010):2344-2353.
24. Ortblad KF, Salomon JA, Barnighausen T, Atun R. Stopping tuberculosis: a biosocial model for sustainable development. *Lancet.* 2015;386(10010):2354-2362.
25. Kranzer K, Khan P, Godfrey-Fausset P, Ayles H, Lonnroth K. Tuberculosis control. *Lancet.* 2016;387(10024):1159-1160.
26. Khan MS, Fletcher H, London School of H, Tropical Medicine TBCSC, Coker R. Investments in tuberculosis research - what are the gaps? *BMC medicine.* 2016;14(1):123.
27. Crampin AC, Glynn JR, Fine PE. What has Karonga taught us? Tuberculosis studied over three decades. *Int J Tuberc Lung Dis.* 2009;13(2):153-164.
28. Ponnighaus JM, Fine PE, Sterne JA, et al. Efficacy of BCG vaccine against leprosy and tuberculosis in northern Malawi. *Lancet.* 1992;339(8794):636-639.
29. Glynn JR, Crampin AC, Yates MD, et al. The importance of recent infection with Mycobacterium tuberculosis in an area with high HIV prevalence: a long-term molecular epidemiological study in Northern Malawi. *J Infect Dis.* 2005;192(3):480-487.
30. Houben RM, Glynn JR, Mallard K, et al. Human immunodeficiency virus increases the risk of tuberculosis due to recent re-infection in individuals with latent infection. *Int J Tuberc Lung Dis.* 2010;14(7):909-915.
31. Crampin AC, Glynn JR, Traore H, et al. Tuberculosis transmission attributable to close contacts and HIV status, Malawi. *Emerg Infect Dis.* 2006;12(5):729-735.
32. Crampin AC, Floyd S, Ngwira BM, et al. Assessment and evaluation of contact as a risk factor for tuberculosis in rural Africa. *Int J Tuberc Lung Dis.* 2008;12(6):612-618.

33. Glynn JR, Guerra-Assuncao JA, Houben RM, et al. Whole Genome Sequencing Shows a Low Proportion of Tuberculosis Disease Is Attributable to Known Close Contacts in Rural Malawi. *PLoS One.* 2015;10(7):e0132840.
34. Guerra-Assuncao J, Crampin A, Houben R, et al. Large-scale whole genome sequencing of provides insights into transmission in a high prevalence area. *eLife.* 2015;4.
35. Mboma SM, Houben RM, Glynn JR, et al. Control of (Multi)Drug Resistance and Tuberculosis Incidence over 23 Years in the Context of a Well-Supported Tuberculosis Programme in Rural Malawi. *PLoS One.* 2013;8(3):e58192.
36. Crampin AC, Dube A, Mboma S, et al. Profile: the Karonga Health and Demographic Surveillance System. *Int J Epidemiol.* 2012;41(3):676-685.
37. Esmail H, Barry CE, 3rd, Young DB, Wilkinson RJ. The ongoing challenge of latent tuberculosis. *Philosophical transactions of the Royal Society of London Series B, Biological sciences.* 2014;369(1645):20130437.
38. Detjen AK, Magdorf K. Characteristics of childhood tuberculosis. *Besonderheiten der Kindertuberkulose.* 2009;63(4):207-218.
39. Middelkoop K, Bekker LG, Morrow C, Zwane E, Wood R. Childhood tuberculosis infection and disease: a spatial and temporal transmission analysis in a South African township. *S Afr Med J.* 2009;99(10):738-743.
40. Griffith AH, Bellamy MJ, Davey MF. Case-Finding by Serial Tuberculin-Testing of Schoolchildren. *Br Med J.* 1963;2(5359):717-720.
41. Gedde-Dahl T. Tuberculous infection in the light of tuberculin matriculation. *Am J Hyg.* 1952;56(2):139-214.
42. Vynnycky E, Fine PE. The natural history of tuberculosis: the implications of age-dependent risks of disease and the role of reinfection. *Epidemiol Infect.* 1997;119(2):183-201.
43. Tiemersma EW, van der Werf MJ, Borgdorff MW, Williams BG, Nagelkerke NJ. Natural history of tuberculosis: duration and fatality of untreated pulmonary tuberculosis in HIV negative patients: a systematic review. *PLoS One.* 2011;6(4):e17601.
44. Dowdy DW, Dye C, Cohen T. Data needs for evidence-based decisions: a tuberculosis modeler's 'wish list'. *Int J Tuberc Lung Dis.* 2013;17(7):866-877.
45. Victora CG, Huttly SR, Fuchs SC, Olinto MT. The role of conceptual frameworks in epidemiological analysis: a hierarchical approach. *Int J Epidemiol.* 1997;26(1):224-227.
46. Godfrey-Faussett P. Measuring TB transmission and its impact at community level: what is missing? Paper presented at: HIV/TB research meeting in conjuction with the 19th Conference on Retroviruses and Opportunistic Infections2012; Seattle, USA.
47. Hargreaves JR, Boccia D, Evans CA, Adato M, Petticrew M, Porter JD. The social determinants of tuberculosis: from evidence to action. *Am J Public Health.* 2011;101(4):654-662.
48. Turner RD, Chiu C, Churchyard GJ, et al. Tuberculosis Infectiousness and Host Susceptibility. *J Infect Dis.* 2017;216(suppl_6):S636-S643.

49. National Tuberculosis Controllers A, Centers for Disease C, Prevention. Guidelines for the investigation of contacts of persons with infectious tuberculosis. Recommendations from the National Tuberculosis Controllers Association and CDC. *MMWR Recommendations and reports : Morbidity and mortality weekly report Recommendations and reports / Centers for Disease Control.* 2005;54(RR-15):1-47.
50. Mandalakas AM, Kirchner HL, Lombard C, et al. Well-quantified tuberculosis exposure is a reliable surrogate measure of tuberculosis infection. *Int J Tuberc Lung Dis.* 2012;16(8):1033-1039.

2 *Mycobacterium tuberculosis* transmission in high burden settings: literature review

2.1 Introduction

This chapter is a published review on the transmission of *M.tb* in high burden settings, which was jointly co-first-authored by me and Tom Yates, a PhD student at UCL. The main purpose of the review was to highlight the research gaps in the field of *M.tb* transmission research. In summary, we identified a number of research priorities including the need for the identification of effective strategies for tuberculosis infection control, a call for improved understanding of the transmissibility of drug-resistant strains and of where transmission is taking place, and better estimates of the effect of HIV and antiretroviral therapy on transmission dynamics. We also highlight the different methods of quantifying *M.tb* transmission and their limitations, including measuring transmission at the population-level.

2.2 Research paper I

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT
www.lshtm.ac.uk

Registry
T: +44(0)20 7299 4646
F: +44(0)20 7299 4656
E: registry@lshtm.ac.uk

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

RESEARCH PAPER COVER SHEET

PLEASE NOTE THAT A COVER SHEET MUST BE COMPLETED FOR EACH RESEARCH PAPER INCLUDED IN A THESIS.

SECTION A – Student Details

Student	Palwasha Yousafzai Khan
Principal Supervisor	Professor Judith Glynn
Thesis Title	Investigating Mycobacterium tuberculosis transmission in rural Malawi

If the Research Paper has previously been published please complete Section B, if not please move to Section C

SECTION B – Paper already published

Where was the work published?	The Lancet Infectious Diseases		
When was the work published?	February 2016		
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion	N/A		
Have you retained the copyright for the work?*	**No**	Was the work subject to academic peer review?	**Yes**

**If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.*

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	
Please list the paper's authors in the intended authorship order:	
Stage of publication	Choose an item.

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	The paper was a jointly co-authored following an M.tb transmission meeting in London. The idea for the meeting was my co-author, Tom Yates. I wrote the first draft of the review, including the figure and tables and co-wrote the final draft.

Student Signature: [redacted] **Date:** 26/06/2017

Supervisor Signature: [redacted] **Date:** 28/06/2017

The transmission of *Mycobacterium tuberculosis* in high burden settings

Tom A Yates, Palwasha Y Khan*, Gwenan M Knight, Jonathon G Taylor, Timothy D McHugh, Marc Lipman, Richard G White, Ted Cohen, Frank G Cobelens, Robin Wood, David A J Moore, Ibrahim Abubakar*

Unacceptable levels of *Mycobacterium tuberculosis* transmission are noted in high burden settings and a renewed focus on reducing person-to-person transmission in these communities is needed. We review recent developments in the understanding of airborne transmission. We outline approaches to measure transmission in populations and trials and describe the Wells–Riley equation, which is used to estimate transmission risk in indoor spaces. Present research priorities include the identification of effective strategies for tuberculosis infection control, improved understanding of where transmission occurs and the transmissibility of drug-resistant strains, and estimates of the effect of HIV and antiretroviral therapy on transmission dynamics. When research is planned and interventions are designed to interrupt transmission, resource constraints that are common in high burden settings—including shortages of health-care workers—must be considered.

Lancet Infect Dis 2016; 16: 227–38

*Contributed equally

Centre for Infectious Disease Epidemiology, Research Department of Infection and Population Health, University College London, London, UK (T A Yates MSc, Prof I Abubakar FRCP); Wellcome Trust Africa Centre for Population Health, Mtubatuba, South Africa (T A Yates); Department of Infectious Disease Epidemiology, Faculty of Epidemiology and Population Health (P Y Khan MBBCh, Prof R Wood DSc [Med]), Tuberculosis Centre (P Y Khan, G M Knight PhD, R G White PhD, Prof R Wood, Prof D A J Moore MD), Tuberculosis Modelling Group (G M Knight, R G White), and Department of Clinical Research (Prof D A J Moore); London School of Hygiene & Tropical Medicine, London, UK; Karonga Prevention Study, Chilumba, Malawi (P Y Khan); National Institute for Health Research Health Protection Research Unit in Healthcare Associated Infection and Antimicrobial Resistance, Imperial College London, London, UK (G M Knight); UCL Institute for Environmental Design and Engineering, Bartlett School of Environment, Energy and Resources (J G Taylor PhD), Centre for Clinical Microbiology (Prof T D McHugh PhD), Division of Medicine (M Lipman MD), and MRC Clinical Trials Unit at University College London (Prof I Abubakar), University College London, London, UK; Department of Epidemiology of Microbial Diseases, Yale School of Public Health, New Haven, CT, USA (T Cohen MD); Department of Global Health, Academic Medical Center, Amsterdam, Netherlands (Prof F G Cobelens PhD); KNCV Tuberculosis Foundation, The Hague, Netherlands (Prof F G Cobelens);

Introduction

Sustained reductions in disease incidence of up to 20% per year are required to meet the targets set out in the WHO End TB Strategy.[1,2] However, incidence is currently only estimated to be reducing at 1·5% per annum.[3] This trend is consistent with model predictions with respect to the probable effect of present control strategies,[4] which focus on case detection and treatment completion.[5] Even in areas with good rates of case finding and treatment completion, evidence suggests that transmission is an issue. Although quality data for active tuberculosis in children younger than 5 years are scarce, the incidence of paediatric cases indicate continuing high levels of transmission.[3,6,7] Tuberculin surveys in high prevalence countries estimate annual risks of *Mycobacterium tuberculosis* infection of 0·3–2·2%,[8–12] but exceeding 5% in some parts of southern Africa.[13,14] Test reversions (negative tests in people who previously had a positive test) mean such cross-sectional surveys might underestimate transmission.[15] Data for *M tuberculosis* transmission derived from molecular typing methods from high burden areas are limited to a small number of research active settings. Nevertheless, these data suggest more disease results from recent transmission than from reactivation of latent tuberculosis,[16,17] particularly in people living with HIV.[18] The rapid rebound in tuberculosis incidence after the discontinuation of isoniazid preventive treatment (IPT) in southern African studies suggest continuing transmission is important in high burden settings,[19,20] although models predict a contribution from reactivation disease implying IPT might not sterilise.[21,22]

To achieve the goals of the End TB Strategy,[2] an increased emphasis on reducing person-to-person *M tuberculosis* transmission in high burden settings is needed. This Review summarises research into *M tuberculosis* transmission in these settings. We focus on the biology of airborne *M tuberculosis* transmission, measuring transmission in populations, and modelling transmission with the Wells–Riley approach. We conclude by identifying research priorities. We do not discuss transmission-blocking vaccines or mixed infections, each the subject of a recent review article.[23,24] Of note, no international consensus exists for tuberculosis incidence or prevalence thresholds that define high burden, although a tuberculosis incidence of 100 cases per 100 000 people per year has been used by WHO.[25] Most of the studies we review were implemented in communities with a tuberculosis incidence of 100 cases or more per 100 000 people per year.

Airborne *M tuberculosis* transmission

Although *M tuberculosis* complex organisms can be spread through unpasteurised milk, direct inoculation, and other means, we focus on the predominant route, airborne transmission. The fundamentals of airborne *M tuberculosis* transmission were described by William Frith Wells, Richard Riley, Robert Loudon, Rena Roberts, and others, more than 60 years ago.[26] Recent progress in basic and clinical sciences has improved our understanding of *M tuberculosis* transmission, which had remained largely unchanged for more than 50 years. However, much remains unknown. Interruption of any process in the natural history of *M tuberculosis* will reduce rates of transmission at a population level (figure).

Individuals with pulmonary tuberculosis aerosolise *M tuberculosis*, placing their contacts at risk of infection (figure). This aerosolisation occurs at a faster rate during coughing.[27] Some evidence suggests that speech and singing are effective aerosol generating activities,[28,29] although these studies focused only on droplets originating in the mouth. Although the largest respiratory droplets fall to the ground, rapid evaporation means many droplets attain a sufficiently low mass before settling so that they remain suspended in air currents until either inhaled or ventilated out of the room.[30] Important new insights into *M tuberculosis* transmission have come from cough box experiments.[31] In these studies, tuberculosis patients were asked to cough as frequently as was comfortable for 5 min

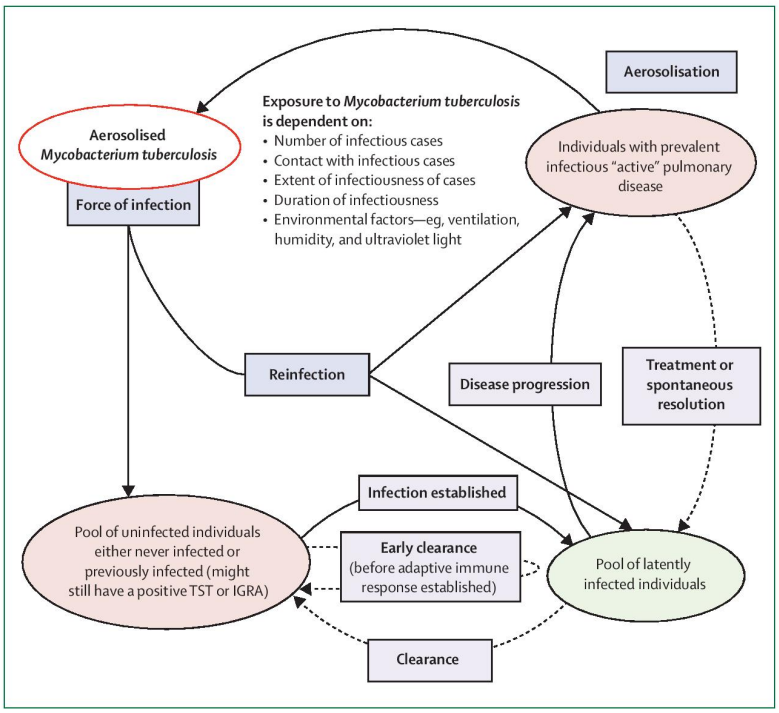


***Figure:* *Mycobacterium tuberculosis* transmission cycle**
IGRA=interferon-gamma release assay. TST=tuberculin skin test.

and The Desmond Tutu HIV Centre, Institute of Infectious Disease and Molecular Medicine, University of Cape Town, Cape Town, South Africa
(Prof R Wood)

Correspondence to:
Dr Tom A Yates, Centre for Infectious Disease Epidemiology, Research Department of Infection and Population Health, University College London, London WC1E 6JB, UK
t.yates@ucl.ac.uk

into a cough aerosol sampling system. Although this procedure might not represent real physiological processes, these experiments suggest that most *M tuberculosis* that is aerosolised during coughing is in droplets small enough, even without evaporation, to remain suspended in the air.[31] These cough box experiments,[31,32] consistent with studies from guineapig facilities[33,34] and molecular epidemiological observations,[35–37] suggest that some people with tuberculosis might be much more infectious than others. Early animal experiments showed that *M tuberculosis* in smaller droplets more readily produced tubercles in the lung than did *M tuberculosis* in larger droplets, presumably because these small droplets escape filtration in the upper airways.[38,39] The use of surgical masks by patients with tuberculosis has been shown to reduce transmission to guineapigs by 56%, suggesting they partly block aerosolisation of the relevant respiratory droplets.[40]

The quantity and characteristics of the inhaled droplet can predict clinical outcomes, with early experiments showing that the infectious dose predicts the risk of infection and progression to disease.[41] In the cough box experiments, the quantity of aerosolised *M tuberculosis* produced by individuals predicted infection in household contacts better than the smear grade or time to culture positivity.[32] Larger droplets, settling in the upper airway, might result in immune memory and a positive test for infection but little risk of progression to disease.[41] The fact that some highly exposed individuals do not develop positive tuberculin skin tests (TSTs) or interferon-gamma release assays (IGRAs) and the discovery of genetic loci that predict TST positivity suggest that the body can clear *M tuberculosis* infection without development of an adaptive immune response—so-called early clearance.[42] This process is probably important epidemiologically but, because no evidence of the early clearance is left, is difficult to study. Informed by animal studies and advanced imaging techniques, appreciation is growing for the opinion that a binary classification of tuberculosis into latent infection and active disease might be too simplistic.[43,44] Some individuals with positive tests for infection might have cleared the mycobacteria and periods of active replication in latent tuberculosis have been reported.[43,44] Pulmonary *M tuberculosis* infection towards the more active end of the infection-to-disease range is probably necessary for infectiousness.

A widely-held view is that the infectiousness of patients diminishes sufficiently after 2 weeks of antituberculous treatment such that transmission to contacts is unlikely.[45] Many guidelines rely on proxy measures of infectiousness—eg, smear status or culture conversion. Of note, median time to culture conversion in patients treated with daily direct observation for drug-susceptible tuberculosis in Peru was much longer than 2 weeks at 37 days.[46] However, in patients on treatment, the association between sputum smear or culture status and infectiousness is not straightforward,[47–51] and is likely to be influenced not only by mycobacteria viability in respiratory secretions but also by the capacity to generate aerosolised *M tuberculosis* through coughing.[27] Because cough frequency diminishes with treatment,[52] assumptions about infectiousness on the basis of culture conversion times might overestimate risk. Furthermore, organisms that propagate in culture might not thrive when exposed to a hostile immune system in the alveoli.

Infectiousness can be studied in guineapig facilities in which the number of animals infected after exposure to air exhausted from isolation rooms containing patients with tuberculosis is measured. Such experiments show that effective treatment is associated with decidedly fewer *M tuberculosis* infections than are reported before treatment is initiated or when isolates are not fully susceptible to the treatment regimen.[34,51] However, the guineapigs in these experiments were exposed to the patients for many weeks. Thus, these experiments have not yet reliably established a time window after which a patient can be deemed to be no longer infectious. Most, if not all, household contacts are more likely to be infected by an index patient in the pretreatment period than once treatment is initiated due to both the likely longer duration of exposure and the greater infectiousness pretreatment.[53] This situation might also be true for patients with drug-resistant tuberculosis, in whom culture conversion times will typically be longer than those with drug-susceptible tuberculosis, particularly if initiation of effective treatment is delayed.

Measuring transmission in populations

Even in the highest burden communities, the prevalence and annual incidence of active tuberculosis disease rarely exceed 2%. Infrequent outcomes, in combination with incomplete surveillance data and poor tests for infection, make the measurement of *M tuberculosis* transmission in populations a challenge.[54] Inference about transmission therefore relies on proxy measures, assumptions, and combination approaches. The best approach to measure *M tuberculosis* transmission in trials of control interventions is controversial.[55–57] Short-term reductions in disease prevalence, for example, are difficult to interpret as prevalence is affected not only by transmission but also by progression from infection to disease and disease duration.[55] We have summarised frequently used tests for *M tuberculosis* infection (panel) and approaches to measure the *M tuberculosis* transmission in populations (table).

Traditional approaches to measure transmission include tuberculin surveys in schoolchildren. Such surveys are a typical means of estimating *M tuberculosis* transmission at a population level. Although age assortative mixing might mean paediatric infections do not fully reflect *M tuberculosis* transmission between adults,[55,58,67] repeated TST surveys might still allow estimates of the trend in force of infection over time to be attained. Trend estimates based on tuberculin surveys are fairly robust and not greatly influenced by the proportion of children with BCG vaccine scars or the cut point used to define a positive test for infection.[67,69]

Molecular approaches include methods for strain typing *M tuberculosis* such as spoligotyping, which has low resolution; restriction fragment length polymorphism and mycobacterial interspersed repetitive units-variable number of tandem repeats,[70–72] both of which have been widely used; and whole genome sequencing (WGS). These molecular epidemiological techniques provide evidence for or against potential linkages between two or more cases of active tuberculosis and have led to several crucial insights into *M tuberculosis* transmission.[73] Since these techniques need a bacterial isolate, molecular epidemiology, with rare exceptions,[74] captures only infections that have progressed to active tuberculosis disease.[75] Molecular epidemiology cannot distinguish changes in transmission intensity from changes in the rate of progression to active tuberculosis shortly after infection resulting, for example, from varying levels of immunosuppression. The high resolution of WGS and steep reductions in cost mean this technique is likely to eventually replace existing strain typing techniques. However, molecular epidemiology with WGS will need an understanding of the rate at which mutations occur. Recent studies[37,76–78] suggest that, in active disease, single nucleotide polymorphisms (SNPs) emerge, on average, at half an SNP per *M tuberculosis* genome per year or slower. Most of the patients in these studies were on tuberculosis treatment and substantial variation was reported in the rate at which mutations occurred. A primate study suggests a similar mutation rate and that the mutation rate might not differ substantially between active and latent infection.[79] However, scarce data suggest that, in man, mutations accumulate more slowly during latent infection than in active infection.[80] Occasional accelerated intrapatient microevolution events[81] and the slow rate at which SNPs accumulate might make inference of chains of transmission from *M tuberculosis* genotypes alone challenging. Probabilistic models that also incorporate epidemiological and clinical data might be needed.[82] With molecular epidemiology by WGS, as with older strain

Panel: Diagnostic tests of *Mycobacterium tuberculosis* infection

- No gold standard diagnostic test exists for *Mycobacterium tuberculosis* infection.
- The two widely used diagnostic tests are tuberculin skin tests (TSTs) and interferon-gamma release assays (IGRAs). Positive tests are interpreted as showing a previous adaptive immune response to mycobacterial infection.
- Detection is not possible for infections cleared by the innate immune system before an adaptive response,[42] nor is it possible to distinguish cleared infections that leave a lingering immunological footprint from persistent infection.
- Neither test is able to distinguish between latent infection and disease.

Tuberculin skin tests

- TSTs use an intradermal injection of a standardised purified protein derivative then measurement of any induration after 48–96 h.
- Sensitivity and specificity are dependent on the number of millimetres of induration chosen as the cut point and the prevalence of non-specific reactions resulting from exposure to environmental mycobacteria or previous BCG vaccination.[58] The BCG effect wanes in children vaccinated in infancy.
- Low sensitivity can be seen with advanced age and with immunosuppression as a result of malnutrition or HIV.
- New statistical techniques can suggest appropriate cut points for given distributions of reaction sizes.[59]
- Test reversions do occur and are more common in young individuals, probably reflecting an initial false positive test.[60]

Interferon-gamma release assay

- IGRAs require a blood sample to be taken from patients. T cells are then exposed to antigens that are found in *M tuberculosis* but not in BCG or in most environmental mycobacteria.
- Interferon-gamma released by cells that recognise these antigens is then assayed in the supernatant after incubation or by counting the number of interferon-gamma producing cells in an enzyme-linked immunospot assay.
- IGRAs are a more specific test for *M tuberculosis* infection but less precedent exists for their use in transmission studies. The need for phlebotomy and the high cost of the test are also disadvantages.
- Test reversions are common and the clustering of results around the threshold for positivity means the choice of cut point can substantially affect sensitivity, specificity, and prevalence estimates.[15]

New methods

- A new and, hopefully, more specific skin test based on similar antigens to those used in IGRAs is now being tested in phase 3 trials.[61]
- RNA expression signatures have been developed that might distinguish disease from latent infection, tuberculosis from other diseases, and that might revert after successful treatment of active tuberculosis.[62–65] These signatures need further validation.

	What is measured?	Advantages	Disadvantages
Prevalence of tuberculosis infection	Typically measured with tuberculin skin tests in school-age children	Cheap and well established; infections must have occurred within an individual's lifetime, hence, in young children, this test is a measure of recent infection; prevalence can be converted into an annualised incidence (ie, ARTI);[58] repeated surveys or continuous measurement of infection prevalence in the same age group can quantify changes in *Mycobacterium tuberculosis* transmission over time	Does not capture early clearance; poor sensitivity and specificity, uncertainty with respect to cut points plus conversions and reversions of test positivity can affect estimates in some populations; from a study at only one timepoint, age and cohort effects cannot be separated; Styblo's rule, which states that ARTI and the incidence of tuberculosis disease have a fixed association, is no longer thought to be valid;[66,67] WHO no longer recommends single tuberculin skin test surveys
Incidence of tuberculosis infection	Testing cohorts for tuberculosis infection longitudinally	Older children can be included and inferences still made about recent transmission; an incidence cohort including older children and adults provides a general insight into transmission in the community even if mixing patterns are strongly age assortative	Does not capture early clearance; needs larger sample or longer duration of follow-up than measurements of infection prevalence; losses to follow-up might reduce power and bias estimates; exclusion of those who are positive at baseline might exclude those at highest risk—a particular issue in older individuals in high burden settings
Tuberculosis notifications	Notifications of tuberculosis disease to the national treatment programme	Data are routinely captured; enhancement of capacity to diagnose and notify cases of tuberculosis might be possible for the purposes of research, although substantial biases and quality problems inherent to routinely collected data are likely to persist	Serious problems with data quality in most high burden settings; captures only tuberculosis transmission that progresses to disease; captures only individuals who access a diagnosis and whose diagnosis is notified; might capture individuals who do not have tuberculosis—poor specificity is a particular issue where tuberculosis diagnosis is mainly on the basis of chest radiographs, such as in children
Prevalence of tuberculosis disease	Typically measured in large surveys by sputum culture with or without prescreening for symptoms and/or with a chest radiograph	Well established; undiagnosed individuals can be referred for treatment	Substantial and expensive undertaking; captures only tuberculosis transmission that progresses to (pulmonary) disease; whether changes in prevalence are a result of differences in transmission, in progression from infection to disease, or in disease duration, might not be clear; prevalence surveys are active case finding interventions and will transiently change local tuberculosis epidemiology; sputum culture has low sensitivity in children
Incidence of tuberculosis disease	Measured in established cohorts or by two prevalence surveys	Allows changes in incidence to be disaggregated from changes in disease duration	Except in established cohorts in high burden settings, the measurement of incidence needs more than one large prevalence survey, which is rarely feasible; captures only tuberculosis transmission that progresses to (pulmonary) disease; whether differences in prevalence are a result of variation in transmission or in progression from infection to disease is not clear
Molecular epidemiology (proportion clustered)	The proportion of isolates that have the same strain type usually with RFLP, MIRU-VNTR, or WGS*	Allows inferences to be made about the proportion of tuberculosis resulting from reactivation versus recent infection; strain typing can disprove or provide evidence to support putative transmission events	Needs advanced laboratory capacity; captures only transmissions that progress to disease and isolates that are sampled; biased estimates can be obtained if the sampling fraction is low, if the study is not of sufficient duration, or if substantial in or out migration of participants takes place

ARTI=annual risk of tuberculosis infection. RFLP=restriction fragment length polymorphism. MIRU-VNTR=mycobacterial interspersed repetitive units-variable number of tandem repeats. WGS=whole genome sequencing. *These techniques type strains of *M tuberculosis*.

Table: **Measures of *Mycobacterium tuberculosis* transmission in populations**

typing techniques, adequate study duration, a high sampling fraction, and careful documentation, follow-up, and reporting are important.[16,83–85] However, novel approaches to the analysis of sequence data might allow population level inferences to be made from a smaller number of people than the older genotyping techniques.

Direct detection of aerosolised *M tuberculosis* in samples of room air is of great interest. This ability might allow quantification of *M tuberculosis* exposure in putative sites of transmission. A few demonstration studies[86–88] of PCR with room air filtrate suggest that this detection might be feasible. Although PCR detection of *M tuberculosis* DNA does not necessarily mean organisms are viable, it would suggest that individuals with pulmonary tuberculosis have produced bioaerosols in the space. This finding should be, at least in theory, a reasonable proxy for transmission risk.

The Wells–Riley equation

Room ventilation and social contact patterns predict whether other individuals are exposed to *M tuberculosis* that has been aerosolised. The Wells–Riley equation is used to model the transmission of respiratory pathogens,[89] such as *M tuberculosis*, that are spread by crowd rather than close contact. Transmission risk in a defined space over time t is modelled as a Poisson process:

$$\text{Probability of transmission} = 1 - e^{-Iqpt/Q}$$

where I is the number of infectious individuals present, q is the rate at which infectious individuals produce quanta, p is the rate at which susceptible individuals breathe, and Q is the rate at which air from the space is exchanged with uncontaminated air (ventilation).

Riley and colleagues defined quanta as "the number of infectious airborne particles required to infect which may be one or more airborne particles".[89] The parameter q is often assumed or fitted to data. Various attempts have been made to empirically estimate q for tuberculosis by venting air exhaled by patients with tuberculosis over experimental animals. Two sets of experiments in the pre-HIV era estimated quanta production at 0·62–0·82 and 1·25 per h with a heterogeneous group of tuberculosis patients.[40,90–92] More recently, Escombe and colleagues[34] obtained an estimate of 8·2 quanta per h in a group of patients living with HIV in Lima, Peru. These averages disguise huge heterogeneity in infectiousness with the most infectious patients in each study producing 60 and 226 infectious quanta per h.[34,91,93] High rates of quanta production have been measured in patients with advanced multidrug-resistant tuberculosis,[40] and very high rates estimated in outbreaks related to aerosol generating procedures.[94] Interestingly, estimates of q obtained by fitting to data from high burden communities are lower than those obtained empirically.[95] This finding might be because untreated patients in the community are at an earlier stage in their illness and hence less infectious than the diagnosed patients used in the animal studies.

The Wells–Riley equation has several important limitations. The equation assumes that air in the space is fully mixed and does not account for heterogeneity in infectiousness or susceptibility to infection. Adaptations to the equation have been published. One widely used variant uses a rebreathed fraction, the fraction of inhaled air that has been exhaled previously by someone in the building.[96] This rebreathed fraction can be obtained from paired indoor and outdoor carbon dioxide (CO_2) measurements. This process avoids the need to measure Q, which can be technically challenging. In many settings, spatial and temporal variations in CO_2 concentrations are substantial. Not obtaining contemporaneous CO_2 measurements from directly outside the buildings studied might be reasonable in the absence of alternative CO_2 sources and where wind speeds restrict local spatial and temporal variations in CO_2.[95,97] The absence of such contemporaneous measurements would not be reasonable in other circumstances. Several important insights have been derived from the Wells–Riley approach. For example, one study[98] suggested that active case finding could not control high levels of *M tuberculosis* transmission in a South African prison if levels of overcrowding and poor ventilation were not also addressed. Another South African study,[95] which used the equation to predict settings in which *M tuberculosis* transmission might occur, is described later in this Review.

Research priorities

Much is still to be learned about *M tuberculosis* transmission. Approaches to interrupting *M tuberculosis* transmission include active case finding, the provision of IPT, and tuberculosis infection control. Large trials have been completed into active case findings and mass IPT to interrupt *M tuberculosis* transmission. The ZAMSTAR[56] result may be the first empirical data suggesting that active screening for tuberculosis disease reduces *M tuberculosis* transmission at a population level.[99] The Thibela tuberculosis trial,[20] implemented in a setting with a very high incidence of infection, reported mass administration of IPT protected individuals whilst on treatment but had no effect on tuberculosis incidence in the wider community.[22]

Infection control

Little research into tuberculosis infection control has been undertaken. Tuberculosis infection control is conventionally described in three domains—administrative controls (which aim to minimise contamination of shared air by infectious subjects—eg, cough triage, early diagnosis, and treatment), environmental controls (which aim to minimise exposure to *M tuberculosis* through disinfection or removal of contaminated air), and personal protection measures (which aim to minimise inhalation of contaminated air—eg, N95 respirators).[100] A review of observational and animal studies[101] concluded that the evidence is strong in support of the role of ventilation as an environmental control in reducing the risk of airborne transmission of *M tuberculosis*. Of the many ways to increase ventilation, increased mechanical ventilation,[102] natural ventilation through increased window opening,[103] and wind-driven roof turbines have been considered specifically as means to reduce *M tuberculosis* infection risk.[104] Natural ventilation has been recommended by WHO as an effective way to reduce infection.[100] Air disinfection methods, particularly upper room ultraviolet germicidal irradiation, have also been studied leading to steep reductions in *M tuberculosis* transmission from tuberculosis patients to experimental animals.[105,106]

Implementation of environmental controls is not always straightforward. An increase in indoor levels of outdoor pollution, security concerns, exposure to outdoor hazards such as disease vectors, a loss of thermal comfort, energy loss through the exfiltration of conditioned (heated or cooled) indoor air, and high running and maintenance costs of mechanical systems, are side-effects of increased ventilation[107] and might make such measures unacceptable to occupants. Therefore, the ideal retrofit and design measures used in a building should account for occupant patterns, numbers, and preferences, the climate and surrounding environment, building geometry, and the materials and budget available. Building simulation instruments might be used to predict the optimum design or retrofit of buildings to maximise ventilation according to specified criteria.[108,109]

The FAST approach to tuberculosis infection control in congregate settings has been promoted and advocates

"Finding TB cases Actively, Separating safely, and Treating effectively".[110] A trial of the FAST approach is about to commence in Peru (NCT 02355223) using TST negative to positive conversions in health-care workers as an endpoint. However, the absence of a comparator group might limit the strength of the conclusions that can be reached. Although an association between ventilation rate and tuberculosis transmission is clear, little empirical evidence exists for the effectiveness of infection control interventions in reducing transmission, with most studies using animal surrogates or ventilation measurements as a proxy for transmission risk. A notable exception was the Tuberculosis Ultraviolet Shelter Study,[111] which showed that environmental modifications can be safely implemented at scale.[111] Although too few TST conversions occurred in residents of the shelters to show an effect on *M tuberculosis* transmission, similar studies in high burden settings would be valuable to quantify the effect of tuberculosis infection control interventions on transmission to human occupants and, potentially, on transmission in the surrounding community.[112] Whether household infection control implemented at diagnosis can mitigate against secondary infections in patients managed in the community is not known. To our knowledge, no trials of such interventions have been implemented. This question is important, perhaps in multidrug-resistant tuberculosis and certainly in extensively drug-resistant tuberculosis, in which chemotherapy might not promptly reduce infectiousness and where the consequences of transmission might be severe.

Locating *M tuberculosis* transmission

Historically, households have been deemed to be a major focus of *M tuberculosis* transmission. However, three molecular epidemiological studies[17,113,114] from sub-Saharan Africa and one from rural Vietnam,[115] all suggest that most transmission occurs between, rather than within, households. In these high burden settings, this finding probably reflects a high transmission risk outside the home rather than a reduction in the risk of transmission to household contacts. Other evidence also suggests that, overall, most transmissions occur outside the household,[116,117] but studies suggest this result is age dependent with young children more likely to have been infected by a household member.[118–120] Transmission in indoor congregate settings is probably important in high burden settings. For example, time working in public transport is strongly associated with TST positivity in Lima, Peru.[121] Understanding which settings are important should be a research priority, because this knowledge would allow infection control and active case finding interventions to be better targeted.[112] The Wells–Riley equation and its variants have been used to estimate *M tuberculosis* transmission risk by location. Studies adopting these approaches have estimated ventilation or likely exposure to exhaled bioaerosols based on CO_2 levels. These methods have been applied to study transmission in a Cape Town, South Africa, township. Data were collected for CO_2 concentrations in various settings visited by 63 adolescents. 93% of total exposure to rebreathed air was estimated to occur in a few locations: own home, visited homes, public transport, work, or school.[97] The same research group used a modified Wells–Riley equation, CO_2 measurements, and social contact pattern data to estimate the proportion of overall *M tuberculosis* transmission by location. The researchers reported that, in the same Cape Town township, 50% of incident infections in 15–19 year olds might take place in school, that workplaces are important places for adult transmission, and that household and public transport might be important sites of transmission between age groups.[95] These inferences are potentially useful but the studies were small in terms of geographical extent and participant numbers. The conclusions might be context specific and assume tuberculosis disease prevalence is the same in each location within an age and sex group, which is probably not the case. Similar studies need to be undertaken on larger scales and in more settings, ideally combined with location specific estimates of tuberculosis prevalence.

Health facilities, particularly in HIV endemic areas, are an important setting in which patients infectious with tuberculosis mix with susceptible people. People with infectious tuberculosis attend health-care facilities before diagnosis, when presenting with tuberculosis symptoms, and during the course of tuberculosis treatment. Delays in recognition, diagnosis, and isolation of infectious cases augment the risk of nosocomial transmission from unsuspected index cases. Overcrowded outpatient clinics and emergency departments congregate people at risk of progressing from infection to disease in settings where the likelihood of exposure to patients with infectious tuberculosis is relatively high. In a study[122] based in an emergency department in Lima, Peru, IGRA conversion was reported in 30% of health-care workers during a 12 month period compared with 0% of hospital security and domestic personnel; remarkably a third of patients identified with tuberculosis in the study were attending the hospital for an apparently unrelated reason (trauma, pregnancy, etc) showing the importance of the unrecognised risk. Nosocomial transmission played an important part in the Tugela Ferry, South Africa, extensively drug-resistant tuberculosis outbreak.[123] Future research should quantify the proportion of *M tuberculosis* transmission in high burden settings that occurs within health-care facilities and the effect of programmatic infection control interventions on this proportion.

Spatial heterogeneity in the incidence of tuberculosis and drug-resistant tuberculosis is evident in analyses of programmatically collected data and has been well documented in many studies,[124–126] raising the possibility that targeted interventions might be effective. At this time, however, the mechanisms driving such heterogeneity are

not completely understood—eg, localised transmission and/or aggregation of individuals sharing risk factors for infection or progression might combine to generate such patterns of disease. An improved understanding of this spatial heterogeneity might usefully inform targeted tuberculosis control interventions. Improved data for the geographical extent of social contacts relevant for tuberculosis transmission would assist in the design of intervention studies. Research addressing this question would be valuable.

Drug resistance and transmissibility

Globally, an estimated 480 000 incident cases of multidrug-resistant tuberculosis occurred in 2013. The proportion of new cases that are infected with multidrug-resistant tuberculosis is about 3·5% and this percentage has not changed noticeably over the period of 2008 to 2013.[3] Multidrug-resistant tuberculosis is disproportionately distributed, with the highest rates reported in central Asia and eastern Europe where, in several countries, a high proportion of multidrug-resistant tuberculosis cases have no previous history of tuberculosis treatment.[3] This finding suggests high levels of multidrug-resistant tuberculosis transmission.

Projections of the future burden of multidrug-resistant tuberculosis depend crucially on estimates of the reproductive potential of drug-resistant strains as compared with drug-sensitive strains.[127,128] This reproductive potential, which can be quantified as the expected number of secondary infectious cases attributable to a single infectious case, is the product of several factors: the duration of infectiousness, the rate at which respiratory exposures occur, the probability that exposure results in transmission, and the probability that infection progresses to infectious active disease.[129] Drug resistance might affect several of these factors. For example, the duration of infectiousness is often longer in individuals with multidrug-resistant tuberculosis because delays in the diagnosis of resistance can lead to delayed initiation of effective treatment. However, the probability of exposure causing a secondary infectious case might be decreased if resistance-conferring mutations have a fitness cost.[130] In-vitro experiments (eg, competitive growth assays that measure biological fitness) and observational studies (eg, contact tracing and molecular clustering studies, which measure effects of biological fitness and differences in duration of infectiousness) suggest a wide variation in the association between drug resistance and transmission.[131] Furthermore, even where clear biological costs are associated with resistance, these costs can be ameliorated by compensatory mutations.[132,133] WGS analyses of clinical strains suggest successful transmission of multidrug-resistant strains in disparate settings, such as South Africa and Russia.[134,135] However, a household contact study[136] suggested circulating multidrug-resistant tuberculosis strains in Peru were less likely to result in disease in household contacts than drug-sensitive strains. The data for transmission of extensively drug-resistant *M tuberculosis*, at least during the time of observation, are also mixed.[134,137] Because the fitness of drug resistant strains might increase over time (as a result of selection or through compensatory mutations), these mixed results might reflect differences in the maturity of drug-resistant tuberculosis epidemics.[138] In view of the importance of reproductive potential to projections of the multidrug-resistant and extensively drug-resistant tuberculosis epidemics, this topic is still a research priority.

HIV and *M tuberculosis* transmission

Our understanding of the effects of HIV on *M tuberculosis* transmission is scarce.[75] Some data suggest that people living with HIV with tuberculosis disease make a small overall contribution to *M tuberculosis* transmission. The arrival of HIV in Tanzania was associated with an increase in tuberculosis incidence but a reduction in annual risk of tuberculosis infection measured in a series of tuberculin surveys.[69] The arrival of HIV in South Africa was associated with an obvious increase in tuberculosis incidence in people living with HIV but not those who were HIV negative.[139] This finding was replicated in a prospective cohort study[140] in business employees in Harare. These studies were completed before antiretroviral therapy was widely available. A household contact study[141] suggests that transmission to household contacts is less frequent when the index case has more advanced HIV-related immunosuppression. Molecular epidemiology suggests that, in a South African township, HIV-negative people are more likely to be the index cases in strain clusters than HIV-positive people.[17,142] However, a study[78] in Malawi with WGS reported no association between HIV status or receipt of antiretroviral therapy (ART) and the probability of being linked to secondary cases. Several potential explanations exist for these observations. People living with HIV are more likely to have smear-negative or extrapulmonary disease, which are less infectious. Short disease duration due to fast progression to death or treatment might limit the opportunity to transmit,[143] as might the reduced social contact as a result of greater morbidity. However, although people in HIV care might have their tuberculosis diagnosed faster, health-care facilities might be important sites of transmission.

After publication of the START[144] and ANRS TEMPRANO[145] trials, WHO guidelines have been updated.[146] ART is now recommended to be provided to all people living with HIV irrespective of CD4 cell count. In response, national policies are likely to be updated to recommend an earlier initiation of ART than recommended at present. ART reduces tuberculosis disease incidence rates in HIV cohorts by about two-thirds.[147–149] Short-term reductions in population level tuberculosis disease rates have also been reported in

Search strategy

We searched all studies published before Oct 1, 2015, in PubMed. We sought articles published in English using the terms "tuberculosis"and "transmission". We also included papers from the reference lists of these papers and all authors suggested papers for inclusion in the Review.

communities in South Africa and Malawi, where ART has been scaled up rapidly.[150,151] This finding might be largely explained by reduced progression from infection to disease rather than by reductions in transmission. The long-term effect of ART on tuberculosis disease burden at the population level and the effect on *M tuberculosis* infection incidence are uncertain. ART certainly affects longevity, levels of contact with health-care services, and susceptibility to tuberculosis disease. ART might also affect tuberculosis disease duration and phenotype, including the presence of cavities, smear positivity,[152,153] and the frequency of extrapulmonary disease, all of which might affect infectiousness. Furthermore, reduced morbidity as a result of ART might result in increased levels of social contact.

For more about the **transmission meeting** see http://tb.lshtm.ac.uk/tb-transmission

An influential modelling study[154] estimating the effect of the roll-out of annual HIV testing and immediate ART on tuberculosis disease incidence in nine African countries predicted a 21% (range 10–31%) reduction in the cumulative AIDS-related tuberculosis disease incidence over the first 5 years, and a 48% (range 37–55%) reduction in the incidence of tuberculosis disease at 5 years. A multimodel analysis[155] for the time period 2014–33, estimated that increasing ART coverage to 80% of those with a CD4 count of 350 cells per µL or less could reduce tuberculosis incidence by 8–14%. Additionally, if ART were provided to all HIV infected individuals (at present levels of access), incidence could be reduced by 6–30%. However, a more recent modelling study[156] suggested that the long-term effect of expanded ART access was less certain than the earlier models suggested. Although tuberculosis incidence should initially reduce, the model predicted that, if good adherence and immunological responses to ART are not sustained and combined with effective HIV preventive interventions, increases in tuberculosis disease incidence might occur despite high levels of ART coverage.[156] After publication of the new WHO guidelines,[146] more people will probably start treatment with ART at higher CD4 counts than previously. In view of the substantial effect—either positive or negative—that this new practice might have on tuberculosis in HIV endemic settings, the effect of HIV and ART on transmission dynamics should be a focus of research.

Conclusions

Addressing *M tuberculosis* transmission is crucial to achieve control of tuberculosis in high burden settings. Repeated surveys measuring tuberculosis infection in the same community, including young adults, offer a feasible measure of tuberculosis transmission. This approach might be used in trials in high burden settings to enable the effect of interventions on transmission to be disaggregated from their effects on rates of progression to disease or disease duration. The coming years will see innovative and exciting research on *M tuberculosis* transmission in high burden settings. Priority should be given to the development and assessment of strategies for tuberculosis control that place minimal additional demands on poor patients and overstretched health-care systems,[112] or that include elements of social protection or health system strengthening.

Contributors

TAY, PYK, GMK, JGT, RGW, TC, FGC, DAJM, and IA drafted sections of the manuscript. All authors commented on and edited the manuscript. TAY, PYK and IA prepared the final draft. All authors approved the final version of the manuscript before submission.

Declaration of interests

We declare no competing interests.

Acknowledgments

This Review was written after the *M tuberculosis* transmission meeting, held in London in November, 2014, which was funded by the TB Modelling and Analysis Consortium (OPP1084276), University College of London Population Health Domain, and the TB Centre at the London School of Hygiene & Tropical Medicine. TAY is funded with a studentship from the Medical Research Council (UK). PYK is funded by the Wellcome Trust clinical research training fellowship. RGW is funded by the Medical Research Council (UK), the Bill and Melinda Gates Foundation, Centres for Disease Control/President's Emergency Plan for AIDS Relief through the Aurum Institute, and the US Agency for International Development/International Union Against Tuberculosis and Lung Disease/The Union North America. TC is funded by a National Institutes of Health Research Project Grant (R01 AI112438–01). IA is funded by the UK Medical Research Council, National Institute for Health Research, and Public Health England.

References

1 Glaziou P, Falzon D, Floyd K, Raviglione M. Global epidemiology of tuberculosis. *Semin Respir Crit Care Med* 2013; **34:** 3–16.
2 WHO. End TB Strategy. Draft global strategy and targets for tuberculosis prevention, care and control after 2015. Documentation for World Health Assembly 67. Geneva: World Health Organization, 2014.
3 WHO. Global tuberculosis control: WHO report 2014. Geneva: World Health Organization, 2014.
4 Dowdy DW, Chaisson RE. The persistence of tuberculosis in the age of DOTS: reassessing the effect of case detection. *Bull World Health Organ* 2009; **87:** 296–304.
5 WHO. Stop TB Strategy. Building on and enhancing DOTS to meet the TB-related Millennium Development Goals. Geneva: World Health Organization, 2006.
6 Dodd PJ, Gardiner E, Coghlan R, Seddon JA. Burden of childhood tuberculosis in 22 high-burden countries: a mathematical modelling study. *Lancet Glob Health* 2014; **2:** e453–59.
7 Jenkins HE, Tolman AW, Yuen CM, et al. Incidence of multidrug-resistant tuberculosis disease in children: systematic review and global estimates. *Lancet* 2014; **383:** 1572–79.
8 Gopi PG, Subramani R, Nataraj T, Narayanan PR. Impact of BCG vaccination on tuberculin surveys to estimate the annual risk of tuberculosis infection in south India. *Indian J Med Res* 2006; **124:** 71–76.
9 Munim A, Rajab Y, Barker A, Daniel M, Williams B. Risk of *Mycobacterium tuberculosis* infection in Somalia: national tuberculin survey 2006. *East Mediterr Health J* 2008; **14:** 518–30.
10 Doocy SC, Todd CS, Llainez YB, Ahmadzai A, Burnham GM. Population-based tuberculin skin testing and prevalence of tuberculosis infection in Afghanistan. *World Health Popul* 2008; **10:** 44–53.

11 Addo KK, van den Hof S, Mensah GI, et al. A tuberculin skin test survey among Ghanaian school children. *BMC Public Health* 2010; **10:** 35.

12 Hoa NB, Cobelens FG, Sy DN, Nhung NV, Borgdorff MW, Tiemersma EW. First national tuberculin survey in Viet Nam: characteristics and association with tuberculosis prevalence. *Int J Tuberc Lung Dis* 2013; **17:** 738–44.

13 Kritzinger FE, den Boon S, Verver S, et al. No decrease in annual risk of tuberculosis infection in endemic area in Cape Town, South Africa. *Trop Med Int Health* 2009; **14:** 136–42.

14 Wood R, Lawn SD, Johnstone-Robertson S, Bekker LG. Tuberculosis control has failed in South Africa—time to reappraise strategy. *S Afr Med J* 2011; **101:** 111–14.

15 Andrews JR, Hatherill M, Mahomed H, et al. The dynamics of QuantiFERON-TB gold in-tube conversion and reversion in a cohort of South African adolescents. *Am J Respir Crit Care Med* 2015; **191:** 584–91.

16 Houben RM, Glynn JR. A systematic review and meta-analysis of molecular epidemiological studies of tuberculosis: development of a new tool to aid interpretation. *Trop Med Int Health* 2009; **14:** 892–909.

17 Middelkoop K, Mathema B, Myer L, et al. Transmission of tuberculosis in a South African community with a high prevalence of HIV infection. *J Infect Dis* 2015; **211:** 53–61.

18 Houben RM, Crampin AC, Ndhlovu R, et al. Human immunodeficiency virus associated tuberculosis more often due to recent infection than reactivation of latent infection. *Int J Tuberc Lung Dis* 2011; **15:** 24–31.

19 Samandari T, Agizew TB, Nyirenda S, et al. 6-month versus 36-month isoniazid preventive treatment for tuberculosis in adults with HIV infection in Botswana: a randomised, double-blind, placebo-controlled trial. *Lancet* 2011; **377:** 1588–98.

20 Churchyard GJ, Fielding KL, Lewis JJ, et al, and the Thibela TB Study Team. A trial of mass isoniazid preventive therapy for tuberculosis control. *N Engl J Med* 2014; **370:** 301–10.

21 Houben RM, Sumner T, Grant AD, White RG. Ability of preventive therapy to cure latent *Mycobacterium tuberculosis* infection in HIV-infected individuals in high-burden settings. *Proc Natl Acad Sci USA* 2014; **111:** 5325–30.

22 Vynnycky E, Sumner T, Fielding KL, et al. Tuberculosis control in South African gold mines: mathematical modeling of a trial of community-wide isoniazid preventive therapy. *Am J Epidemiol* 2015; **181:** 619–32.

23 Hawn TR, Day TA, Scriba TJ, et al. Tuberculosis vaccines and prevention of infection. *Microbiol Mol Biol Rev* 2014; **78:** 650–71.

24 Cohen T, van Helden PD, Wilson D, et al. Mixed-strain *Mycobacterium tuberculosis* infections and the implications for tuberculosis treatment and control. *Clin Microbiol Rev* 2012; **25:** 708–19.

25 Getahun H, Matteelli A, Abubakar I, et al. Management of latent *Mycobacterium tuberculosis* infection: WHO guidelines for low tuberculosis burden countries. *Eur Respir J* 2015; **46:** 1563–76.

26 Rieder HL. Epidemiological basis of tuberculosis control. Paris: International Union Against Tuberculosis and Lung Disease, 1999.

27 Turner RD, Bothamley GH. Cough and the transmission of tuberculosis. *J Infect Dis* 2015; **211:** 1367–72.

28 Loudon RG, Roberts RM. Droplet expulsion from the respiratory tract. *Am Rev Respir Dis* 1967; **95:** 435–42.

29 Loudon RG, Roberts RM. Singing and the dissemination of tuberculosis. *Am Rev Respir Dis* 1968; **98:** 297–300.

30 Wells WF. On air-bourne infection. Study II. Droplets and droplet nuclei. *Am J Hyg* 1934; **20:** 611–18.

31 Fennelly KP, Jones-López EC, Ayakaka I, et al. Variability of infectious aerosols produced during coughing by patients with pulmonary tuberculosis. *Am J Respir Crit Care Med* 2012; **186:** 450–57.

32 Jones-López EC, Namugga O, Mumbowa F, et al. Cough aerosols of *Mycobacterium tuberculosis* predict new infection: a household contact study. *Am J Respir Crit Care Med* 2013; **187:** 1007–15.

33 Sultan L, Nyka W, Mills C, O'Grady F, Wells W, Riley RL. Tuberculosis disseminators. A study of the variability of aerial infectivity of tuberculous patients. *Am Rev Respir Dis* 1960; **82:** 358–69.

34 Escombe AR, Moore DA, Gilman RH, et al. The infectiousness of tuberculosis patients coinfected with HIV. *PLoS Med* 2008; **5:** e188.

35 Godfrey-Faussett P, Sonnenberg P, Shearer SC, et al. Tuberculosis control and molecular epidemiology in a South African gold-mining community. *Lancet* 2000; **356:** 1066–71.

36 Ypma RJ, Altes HK, van Soolingen D, Wallinga J, van Ballegooijen WM. A sign of superspreading in tuberculosis: highly skewed distribution of genotypic cluster sizes. *Epidemiology* 2013; **24:** 395–400.

37 Walker TM, Ip CL, Harrell RH, et al. Whole-genome sequencing to delineate *Mycobacterium tuberculosis* outbreaks: a retrospective observational study. *Lancet Infect Dis* 2013; **13:** 137–46.

38 Wells WF, Ratcliffe HL, Grumb C. On the mechanics of droplet nuclei infection II. Quantitative experimental air-borne tuberculosis in rabbits. *Am J Hyg* 1948; **47:** 11–28.

39 Lurie MB, Heppleston AG, Abramson S, Swartz IB. Evaluation of the method of quantitative airborne infection and its use in the study of the pathogenesis of tuberculosis. *Am Rev Tuberc* 1950; **61:** 765–97.

40 Dharmadhikari AS, Mphahlele M, Stoltz A, et al. Surgical face masks worn by patients with multidrug-resistant tuberculosis: impact on infectivity of air on a hospital ward. *Am J Respir Crit Care Med* 2012; **185:** 1104–09.

41 Fennelly KP, Jones-López EC. Quantity and quality of inhaled dose predicts immunopathology in tuberculosis. *Front Immunol* 2015; **6:** 313.

42 Verrall AJ, Netea MG, Alisjahbana B, Hill PC, van Crevel R. Early clearance of *Mycobacterium tuberculosis*: a new frontier in prevention. *Immunology* 2014; **141:** 506–13.

43 Barry CE 3rd, Boshoff HI, Dartois V, et al. The spectrum of latent tuberculosis: rethinking the biology and intervention strategies. *Nat Rev Microbiol* 2009; **7:** 845–55.

44 Esmail H, Barry CE 3rd, Young DB, Wilkinson RJ. The ongoing challenge of latent tuberculosis. *Philos Trans R Soc Lond B Biol Sci* 2014; **369:** 20130437.

45 National Institute for Health and Care Excellence. Tuberculosis: clinical diagnosis and management of tuberculosis, and measures for its prevention and control (CG117). 2011. https://www.nice.org.uk/guidance/cg117 (accessed Jan 5, 2016).

46 Fitzwater SP, Caviedes L, Gilman RH, et al. Prolonged infectiousness of tuberculosis patients in a directly observed therapy short-course program with standardized therapy. *Clin Infect Dis* 2010; **51:** 371–78.

47 Kamat SR, Dawson JJ, Devadatta S, et al. A controlled study of the influence of segregation of tuberculous patients for one year on the attack rate of tuberculosis in a 5-year period in close family contacts in South India. *Bull World Health Organ* 1966; **34:** 517–32.

48 Gunnels JJ, Bates JH, Swindoll H. Infectivity of sputum-positive tuberculous patients on chemotherapy. *Am Rev Respir Dis* 1974; **109:** 323–30.

49 Menzies D. Effect of treatment on contagiousness of patients with active pulmonary tuberculosis. *Infect Control Hosp Epidemiol* 1997; **18:** 582–86.

50 Dharmadhikari AS, Nardell E. Serial acid fast bacilli smear and culture conversion rates over 26 weeks in a cohort of 93 sputum culture–positive tuberculosis (TB). *Clin Infect Dis* 2011; **52:** 554–56.

51 Dharmadhikari AS, Mphahlele M, Venter K, et al. Rapid impact of effective treatment on transmission of multidrug-resistant tuberculosis. *Int J Tuberc Lung Dis* 2014; **18:** 1019–25.

52 Loudon RG, Spohn SK. Cough frequency and infectivity in patients with pulmonary tuberculosis. *Am Rev Respir Dis* 1969; **99:** 109–11.

53 Kasaie P, Andrews JR, Kelton WD, Dowdy DW. Timing of tuberculosis transmission and the impact of household contact tracing. An agent-based simulation model. *Am J Respir Crit Care Med* 2014; **189:** 845–52.

54 Dye C, Bassili A, Bierrenbach AL, et al. Measuring tuberculosis burden, trends, and the impact of control programmes. *Lancet Infect Dis* 2008; **8:** 233–43.

55 Cobelens F, van Leth F, van 't Hoog A. Design of pragmatic trials of tuberculosis interventions. *Lancet* 2014; **383:** 213–14.

56 Ayles H, Muyoyeta M, Du Toit E, et al. Effect of household and community interventions on the burden of tuberculosis in southern Africa: the ZAMSTAR community-randomised trial. *Lancet* 2013; **382:** 1183–94.

57 Ayles H, Floyd S, Beyers N, Godfrey-Faussett P. Design of pragmatic trials of tuberculosis interventions—authors' reply. *Lancet* 2014; **383:** 214–15.

58 Rieder H. Annual risk of infection with *Mycobacterium tuberculosis*. *Eur Respir J* 2005; **25:** 181–85.

59 Neuenschwander BF. Bayesian mixture analysis for tuberculin induration data. The Union, 2007. http://www.tbrieder.org/research/mixture/mixture_documentation.pdf (accessed March 26, 2015).

60 Fine PE, Bruce J, Ponnighaus JM, Nkhosa P, Harawa A, Vynnycky E. Tuberculin sensitivity: conversions and reversions in a rural African population. *Int J Tuberc Lung Dis* 1999; **3:** 962–75.

61 Aggerbeck H, Giemza R, Joshi P, et al. Randomised clinical trial investigating the specificity of a novel skin test (C-Tb) for diagnosis of *M. tuberculosis* infection. *PLoS One* 2013; **8:** e64215.

62 Berry MP, Graham CM, McNab FW, et al. An interferon-inducible neutrophil-driven blood transcriptional signature in human tuberculosis. *Nature* 2010; **466:** 973–77.

63 Cliff JM, Lee JS, Constantinou N, et al. Distinct phases of blood gene expression pattern through tuberculosis treatment reflect modulation of the humoral immune response. *J Infect Dis* 2013; **207:** 18–29.

64 Kaforou M, Wright VJ, Oni T, et al. Detection of tuberculosis in HIV-infected and -uninfected African adults using whole blood RNA expression signatures: a case-control study. *PLoS Med* 2013; **10:** e1001538.

65 Anderson ST, Kaforou M, Brent AJ, et al, for the ILULU Consortium and the KIDS TB Study Group. Diagnosis of childhood tuberculosis and host RNA expression in Africa. *N Engl J Med* 2014; **370:** 1712–23.

66 van Leth F, van der Werf MJ, Borgdorff MW. Prevalence of tuberculous infection and incidence of tuberculosis: a re-assessment of the Styblo rule. *Bull World Health Organ* 2008; **86:** 20–26.

67 Borgdorff M. Annual risk of infection—time for an update? *Bull World Health Organ* 2002; **80:** 501–02.

68 Johnstone-Robertson SP, Mark D, Morrow C, et al. Social mixing patterns within a South African township community: implications for respiratory disease transmission and control. *Am J Epidemiol* 2011; **174:** 1246–55.

69 Egwaga SM, Cobelens FG, Muwinge H, Verhage C, Kalisvaart N, Borgdorff MW. The impact of the HIV epidemic on tuberculosis transmission in Tanzania. *AIDS* 2006; **20:** 915–21.

70 van Embden JD, Cave MD, Crawford JT, et al. Strain identification of *Mycobacterium tuberculosis* by DNA fingerprinting: recommendations for a standardized methodology. *J Clin Microbiol* 1993; **31:** 406–09.

71 Supply P, Mazars E, Lesjean S, Vincent V, Gicquel B, Locht C. Variable human minisatellite-like regions in the *Mycobacterium tuberculosis* genome. *Mol Microbiol* 2000; **36:** 762–71.

72 Supply P, Lesjean S, Savine E, Kremer K, van Soolingen D, Locht C. Automated high-throughput genotyping for study of global epidemiology of *Mycobacterium tuberculosis* based on mycobacterial interspersed repetitive units. *J Clin Microbiol* 2001; **39:** 3563–71.

73 Borgdorff MW, van Soolingen D. The re-emergence of tuberculosis: what have we learnt from molecular epidemiology? *Clin Microbiol Infect* 2013; **19:** 889–901.

74 du Plessis DG, Warren R, Richardson M, Joubert JJ, van Helden PD. Demonstration of reinfection and reactivation in HIV-negative autopsied cases of secondary tuberculosis: multilesional genotyping of *Mycobacterium tuberculosis* utilizing IS 6110 and other repetitive element-based DNA fingerprinting. *Tuberculosis (Edinb)* 2001; **81:** 211–20.

75 Godfrey-Faussett P. Population-level control of HIV-related TB (oral presentation). 21st Conference of Retroviral and Opportunistic Infections; Boston, Massachusetts; March 3–6, 2014. http://www.croiwebcasts.org/console/player/22244?mediaType=audio& (accessed March 27, 2015).

76 Bryant JM, Schürch AC, van Deutekom H, et al. Inferring patient to patient transmission of *Mycobacterium tuberculosis* from whole genome sequencing data. *BMC Infect Dis* 2013; **13:** 110.

77 Roetzer A, Diel R, Kohl TA, et al. Whole genome sequencing versus traditional genotyping for investigation of a *Mycobacterium tuberculosis* outbreak: a longitudinal molecular epidemiological study. *PLoS Med* 2013; **10:** e1001387.

78 Guerra-Assunção JA, Crampin AC, Houben RM, et al. Large-scale whole genome sequencing of *M tuberculosis* provides insights into transmission in a high prevalence area. *eLife* 2015; **4:** 4.

79 Ford CB, Lin PL, Chase MR, et al. Use of whole genome sequencing to estimate the mutation rate of *Mycobacterium tuberculosis* during latent infection. *Nat Genet* 2011; **43:** 482–86.

80 Colangeli R, Arcus VL, Cursons RT, et al. Whole genome sequencing of *Mycobacterium tuberculosis* reveals slow growth and low mutation rates during latent infections in humans. *PLoS One* 2014; **9:** e91024.

81 Pérez-Lago L, Comas I, Navarro Y, et al. Whole genome sequencing analysis of intrapatient microevolution in *Mycobacterium tuberculosis*: potential impact on the inference of tuberculosis transmission. *J Infect Dis* 2014; **209:** 98–108.

82 Didelot X, Gardy J, Colijn C. Bayesian inference of infectious disease transmission from whole-genome sequence data. *Mol Biol Evol* 2014; **31:** 1869–79.

83 van Soolingen D, Borgdorff MW, de Haas PE, et al. Molecular epidemiology of tuberculosis in the Netherlands: a nationwide study from 1993 through 1997. *J Infect Dis* 1999; **180:** 726–36.

84 Murray M, Alland D. Methodological problems in the molecular epidemiology of tuberculosis. *Am J Epidemiol* 2002; **155:** 565–71.

85 Field N, Cohen T, Struelens MJ, et al. Strengthening the Reporting of Molecular Epidemiology for Infectious Diseases (STROME-ID): an extension of the STROBE statement. *Lancet Infect Dis* 2014; **14:** 341–52.

86 Mastorides SM, Oehler RL, Greene JN, Sinnott JT 4th, Kranik M, Sandin RL. The detection of airborne *Mycobacterium tuberculosis* using micropore membrane air sampling and polymerase chain reaction. *Chest* 1999; **115:** 19–25.

87 Wan GH, Lu SC, Tsai YH. Polymerase chain reaction used for the detection of airborne *Mycobacterium tuberculosis* in health care settings. *Am J Infect Control* 2004; **32:** 17–22.

88 Matuka O, Singh TS, Bryce E, et al. Pilot study to detect airborne *Mycobacterium tuberculosis* exposure in a South African public healthcare facility outpatient clinic. *J Hosp Infect* 2015; **89:** 192–96.

89 Riley EC, Murphy G, Riley RL. Airborne spread of measles in a suburban elementary school. *Am J Epidemiol* 1978; **107:** 421–32.

90 Riley RL, Mills CC, Nyka W, et al. Aerial dissemination of pulmonary tuberculosis: a two-year study of contagion in a tuberculosis ward. *Am J Hyg* 1959; **70.**

91 Riley RL, Mills CC, O'Grady F, Sultan LU, Wittstadt F, Shivpuri DN. Infectiousness of air from a tuberculosis ward. Ultraviolet irradiation of infected air: comparative infectiousness of different patients. *Am Rev Respir Dis* 1962; **85:** 511–25.

92 Riley RL, Nardell EA. Clearing the air. The theory and application of ultraviolet air disinfection. *Am Rev Respir Dis* 1989; **139:** 1286–94.

93 Nardell EA, Keegan J, Cheney SA, Etkind SC. Airborne infection. Theoretical limits of protection achievable by building ventilation. *Am Rev Respir Dis* 1991; **144:** 302–06.

94 Beggs CB, Noakes CJ, Sleigh PA, Fletcher LA, Siddiqi K. The transmission of tuberculosis in confined spaces: an analytical review of alternative epidemiological models. *Int J Tuberc Lung Dis* 2003; **7:** 1015–26.

95 Andrews JR, Morrow C, Walensky RP, Wood R. Integrating social contact and environmental data in evaluating tuberculosis transmission in a South African township. *J Infect Dis* 2014; **210:** 597–603.

96 Rudnick SN, Milton DK. Risk of indoor airborne infection transmission estimated from carbon dioxide concentration. *Indoor Air* 2003; **13:** 237–45.

97 Wood R, Morrow C, Ginsberg S, et al. Quantification of shared air: a social and environmental determinant of airborne disease transmission. *PLoS One* 2014; **9:** e106622.

98 Johnstone-Robertson S, Lawn SD, Welte A, Bekker LG, Wood R. Tuberculosis in a South African prison—a transmission modelling analysis. *S Afr Med J* 2011; **101:** 809–13.

99 Kranzer K, Afnan-Holmes H, Tomlin K, et al. The benefits to communities and individuals of screening for active tuberculosis disease: a systematic review. *Int J Tuberc Lung Dis* 2013; **17:** 432–46.

100 WHO. WHO policy on TB infection control in health-care facilities, congregate settings and households. Geneva: World Health Organization, 2009.

101 Li Y, Leung GM, Tang JW, et al. Role of ventilation in airborne transmission of infectious agents in the built environment—a multidisciplinary systematic review. *Indoor Air* 2007; **17:** 2–18.

102 Menzies D, Fanning A, Yuan L, FitzGerald JM, and the Canadian Collaborative Group in Nosocomial Transmission of TB. Hospital ventilation and risk for tuberculous infection in Canadian health care workers. *Ann Intern Med* 2000; **133:** 779–89.

103 Escombe AR, Oeser CC, Gilman RH, et al. Natural ventilation for the prevention of airborne contagion. *PLoS Med* 2007; **4:** e68.

104 Cox H, Escombe R, McDermid C, et al. Wind-driven roof turbines: a novel way to improve ventilation for TB infection control in health facilities. *PLoS One* 2012; **7:** e29589.

105 Escombe AR, Moore DA, Gilman RH, et al. Upper-room ultraviolet light and negative air ionization to prevent tuberculosis transmission. *PLoS Med* 2009; **6:** e43.

106 Shenoi SV, Escombe AR, Friedland G. Transmission of drug-susceptible and drug-resistant tuberculosis and the critical importance of airborne infection control in the era of HIV infection and highly active antiretroviral therapy rollouts. *Clin Infect Dis* 2010; **50** (suppl 3): S231–37.

107 WHO. WHO guidelines for indoor air quality: dampness and mould. Geneva: World Health Organization, 2009.

108 Khan MAI, Noakes CJ, Toropov VV. Development of a numerical optimisation approach to ventilation system design to control airborne contaminant dispersion and occupant comfort. *Build Simul* 2012; **5:** 39–50.

109 Kim SH, Augenbroe G. Decision support for choosing ventilation operations strategy in hospital isolation rooms: a multi-criterion assessment under uncertainty. *Build Environ* 2013; **60:** 305–18.

110 TB Care II. FAST: a tuberculosis infection control strategy. USAID, 2013. https://drtbnetwork.org/sites/default/files/FAST%20May%202013%20Booklet.pdf (accessed Jan 25, 2015).

111 Nardell EA, Bucher SJ, Brickner PW, et al. Safety of upper-room ultraviolet germicidal air disinfection for room occupants: results from the Tuberculosis Ultraviolet Shelter Study. *Public Health Rep* 2008; **123:** 52–60.

112 Yates TA, Tanser F, Abubakar I. Plan beta for tuberculosis: it's time to think seriously about poorly ventilated congregate settings. *Int J Tuberc Lung Dis* 2016; **20:** 5–10.

113 Verver S, Warren RM, Munch Z, et al. Proportion of tuberculosis transmission that takes place in households in a high-incidence area. *Lancet* 2004; **363:** 212–14.

114 Glynn JR, Guerra-Assunção JA, Houben RM, et al. Whole genome sequencing shows a low proportion of tuberculosis disease is attributable to known close contacts in rural Malawi. *PLoS One* 2015; **10:** e0132840.

115 Buu TN, van Soolingen D, Huyen MN, et al. Tuberculosis acquired outside of households, rural Vietnam. *Emerg Infect Dis* 2010; **16:** 1466–68.

116 Narain R, Nair SS, Rao GR, Chandrasekhar P. Distribution of tuberculous infection and disease among households in a rural community. *Bull World Health Organ* 1966; **34:** 639–54.

117 Brooks-Pollock E, Becerra MC, Goldstein E, Cohen T, Murray MB. Epidemiologic inference from the distribution of tuberculosis cases in households in Lima, Peru. *J Infect Dis* 2011; **203:** 1582–89.

118 Wood R, Johnstone-Robertson S, Uys P, et al. Tuberculosis transmission to young children in a South African community: modeling household and community infection risks. *Clin Infect Dis* 2010; **51:** 401–08.

119 Middelkoop K, Bekker LG, Morrow C, Lee N, Wood R. Decreasing household contribution to TB transmission with age: a retrospective geographic analysis of young people in a South African township. *BMC Infect Dis* 2014; **14:** 221.

120 Zelner JL, Murray MB, Becerra MC, et al. Age-specific risks of tuberculosis infection from household and community exposures and opportunities for interventions in a high-burden setting. *Am J Epidemiol* 2014; **180:** 853–61.

121 Horna-Campos OJ, Consiglio E, Sánchez-Pérez HJ, Navarro A, Caylà JA, Martín-Mateo M. Pulmonary tuberculosis infection among workers in the informal public transport sector in Lima, Peru. *Occup Environ Med* 2011; **68:** 163–65.

122 Escombe AR, Huaroto L, Ticona E, et al. Tuberculosis transmission risk and infection control in a hospital emergency department in Lima, Peru. *Int J Tuberc Lung Dis* 2010; **14:** 1120–26.

123 Gandhi NR, Weissman D, Moodley P, et al. Nosocomial transmission of extensively drug-resistant tuberculosis in a rural hospital in South Africa. *J Infect Dis* 2013; **207:** 9–17.

124 Dowdy DW, Golub JE, Chaisson RE, Saraceni V. Heterogeneity in tuberculosis transmission and the role of geographic hotspots in propagating epidemics. *Proc Natl Acad Sci USA* 2012; **109:** 9557–62.

125 Jenkins HE, Gegia M, Furin J, et al. Geographical heterogeneity of multidrug-resistant tuberculosis in Georgia, January 2009 to June 2011. *Euro Surveill* 2014; **19:** pii: 20743.

126 Zelner JL, Murray MB, Becerra MC, et al. Identifying hotspots of multidrug-resistant tuberculosis transmission using spatial and molecular genetic data. *J Infect Dis* 2016; **213:** 287–94.

127 Blower SM, Gerberding JL. Understanding, predicting and controlling the emergence of drug-resistant tuberculosis: a theoretical framework. *J Mol Med (Berl)* 1998; **76:** 624–36.

128 Dye C, Espinal MA. Will tuberculosis become resistant to all antibiotics? *Proc Biol Sci* 2001; **268:** 45–52.

129 Anderson RM, May RM. Infectious diseases of humans: dynamics and control. Oxford: Oxford University Press, 1991.

130 Cohen T, Dye C, Colijn C, Williams B, Murray M. Mathematical models of the epidemiology and control of drug-resistant TB. *Expert Rev Respir Med* 2009; **3:** 67–79.

131 Cohen T, Sommers B, Murray M. The effect of drug resistance on the fitness of *Mycobacterium tuberculosis*. *Lancet Infect Dis* 2003; **3:** 13–21.

132 Comas I, Borrell S, Roetzer A, et al. Whole-genome sequencing of rifampicin-resistant *Mycobacterium tuberculosis* strains identifies compensatory mutations in RNA polymerase genes. *Nat Genet* 2012; **44:** 106–10.

133 de Vos M, Müller B, Borrell S, et al. Putative compensatory mutations in the rpoC gene of rifampin-resistant *Mycobacterium tuberculosis* are associated with ongoing transmission. *Antimicrob Agents Chemother* 2013; **57:** 827–32.

134 Ioerger TR, Feng Y, Chen X, et al. The non-clonality of drug resistance in Beijing-genotype isolates of *Mycobacterium tuberculosis* from the Western Cape of South Africa. *BMC Genomics* 2010; **11:** 670.

135 Casali N, Nikolayevskyy V, Balabanova Y, et al. Evolution and transmission of drug-resistant tuberculosis in a Russian population. *Nat Genet* 2014; **46:** 279–86.

136 Grandjean L, Gilman RH, Martin L, et al. Transmission of multidrug-resistant and drug-susceptible tuberculosis within households: a prospective cohort study. *PLoS Med* 2015; **12:** e1001843.

137 Shah S. Majority of XDR TB Cases are due to transmission in a high HIV prevalence setting. 22nd Conference of Retroviral and Opportunistic Infections; Seattle, WA, USA; Feb 23–26, 2015.

138 Knight GM, Colijn C, Shrestha S, et al. The distribution of fitness costs of resistance-conferring mutations is a key determinant for the future burden of drug-resistant tuberculosis: a model-based analysis. *Clin Infect Dis* 2015; **61** (suppl 3): S147–54.

139 Corbett EL, Charalambous S, Fielding K, et al. Stable incidence rates of tuberculosis (TB) among human immunodeficiency virus (HIV)-negative South African gold miners during a decade of epidemic HIV-associated TB. *J Infect Dis* 2003; **188:** 1156–63.

140 Corbett EL, Bandason T, Cheung YB, et al. Epidemiology of tuberculosis in a high HIV prevalence population provided with enhanced diagnosis of symptomatic disease. *PLoS Med* 2007; **4:** e22.

141 Huang CC, Tchetgen ET, Becerra MC, et al. The effect of HIV-related immunosuppression on the risk of tuberculosis transmission to household contacts. *Clin Infect Dis* 2014; **58:** 765–74.

142 Yates TA, Abubakar I, Tanser F. HIV infection and the transmission of tuberculosis. *J Infect Dis* 2015; **211:** 1510.

143 Corbett EL, Charalambous S, Moloi VM, et al. Human immunodeficiency virus and the prevalence of undiagnosed tuberculosis in African gold miners. *Am J Respir Crit Care Med* 2004; **170:** 673–79.

144 Lundgren JD, Babiker AG, Gordin F, et al. Initiation of antiretroviral therapy in early asymptomatic HIV infection. *N Engl J Med* 2015; **373:** 795–807.

145 Danel C, Moh R, Gabillard D, et al. A trial of early antiretrovirals and isoniazid preventive therapy in Africa. *N Engl J Med* 2015; **373:** 808–22.

146 WHO. Guideline on when to start antiretroviral therapy and pre-exposure prophylaxis for HIV. Geneva: World Health Organization, 2015.

147 Lawn SD, Wood R, De Cock KM, Kranzer K, Lewis JJ, Churchyard GJ. Antiretrovirals and isoniazid preventive therapy in the prevention of HIV-associated tuberculosis in settings with limited health-care resources. *Lancet Infect Dis* 2010; **10:** 489–98.

148 Lawn SD, Harries AD, Williams BG, et al. Antiretroviral therapy and the control of HIV-associated tuberculosis. Will ART do it? *Int J Tuberc Lung Dis* 2011; **15:** 571–81.

149 Suthar AB, Lawn SD, del Amo J, et al. Antiretroviral therapy for prevention of tuberculosis in adults with HIV: a systematic review and meta-analysis. *PLoS Med* 2012; **9:** e1001270.

150 Middelkoop K, Wood R, Bekker LG. The impact of antiretroviral treatment programs on tuberculosis notification rates. *Int J Tuberc Lung Dis* 2011; **15:** 1714–15.

151 Zachariah R, Bemelmans M, Akesson A, et al. Reduced tuberculosis case notification associated with scaling up antiretroviral treatment in rural Malawi. *Int J Tuberc Lung Dis* 2011; **15:** 933–37.

152 Munthali L, Khan PY, Mwaungulu NJ, et al. The effect of HIV and antiretroviral therapy on characteristics of pulmonary tuberculosis in northern Malawi: a cross-sectional study. *BMC Infect Dis* 2014; **14:** 107.

153 van Halsema CL, Fielding KL, Chihota VN, et al. Brief report: the effect of antiretroviral therapy and CD4 count on markers of infectiousness in HIV-associated tuberculosis. *J Acquir Immune Defic Syndr* 2015; **70:** 104–08.

154 Williams BG, Granich R, De Cock KM, Glaziou P, Sharma A, Dye C. Antiretroviral therapy for tuberculosis control in nine African countries. *Proc Natl Acad Sci USA* 2010; **107:** 19485–89.

155 Pretorius C, Menzies NA, Chindelevitch L, et al. The potential effects of changing HIV treatment policy on tuberculosis outcomes in South Africa: results from three tuberculosis-HIV transmission models. *AIDS* 2014; **28** (suppl 1): S25–34.

156 Dodd PJ, Knight GM, Lawn SD, Corbett EL, White RG. Predicting the long-term impact of antiretroviral therapy scale-up on population incidence of tuberculosis. *PLoS One* 2013; **8:** e75466.

3 Measuring the annual risk of *Mycobacterium tuberculosis* infection in young children

3.1 Introduction

The annual risk of *M.tb* infection (ARTI) is a calculated average from an observed prevalence of *M.tb* infection in a specified population. By approximating the incidence of *M.tb* infection, it has the potential to provide information about the extent of community *M.tb* transmission. However, accuracy of the estimation of this metric is a function of, and therefore, dependent on the measure of observed *M.tb* infection prevalence in the chosen population.

This next chapter is a methodological research paper which highlights the difficulties in estimating the ARTI in children, especially in a setting where the vast majority of children have been BCG-vaccinated and the prevalence of infections due to non-tuberculous mycobacteria are non-negligible.

In the paper, a method first described more than 40 years ago in the *American Journal of Epidemiology* was used. I came across this paper whilst reading the paper by Rieder[14] on the methodological issues in the estimation of the burden of tuberculosis (TB) from tuberculin surveys. Rieder mentions that Rust and Thomas developed a model based on tuberculin skin test (TST) data in United States Navy recruits by asking about known history of exposure to a TB case and separated people into contacts and non-contacts, which then allowed a separation of the influences of TST reactions caused by environmental mycobacteria from those resulting from infection with *M.tb.*

As part of my PhD research, I had TST data from children aged under 5 from the demographic surveillance site in Karonga and contemporaneous TST data from

children aged under 5 with known household exposure to a smear-positive TB case, so decided to analyse my data using the Rust and Thomas method.

I think that method has been under-utilised and not taken up by other researchers because it requires TST data to be available on those with known household exposure to a TB case and those who have not been exposed in the same population, which is not widely available. I do think that it is regrettable that in some countries where large tuberculin surveys were being undertaken in school children in the late 1970s that a concurrent tuberculin study in child household contacts of TB was not undertaken. It may have not even been considered. However, there are also major cost and resource implications of conducting such an additional study on a large scale.

3.2 Research paper II

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT
www.lshtm.ac.uk

Registry
T: +44(0)20 7299 4646
F: +44(0)20 7299 4656
E: registry@lshtm.ac.uk

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

RESEARCH PAPER COVER SHEET

PLEASE NOTE THAT A COVER SHEET MUST BE COMPLETED FOR EACH RESEARCH PAPER INCLUDED IN A THESIS.

SECTION A – Student Details

Student	Palwasha Yousafzai Khan
Principal Supervisor	Professor Judith Glynn
Thesis Title	Investigating Mycobacterium tuberculosis transmission in rural Malawi

If the Research Paper has previously been published please complete Section B, if not please move to Section C

SECTION B – Paper already published

Where was the work published?	American Journal of Epidemiology		
When was the work published?	Online ahead of print on 19 May 2017		
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion	N/A		
Have you retained the copyright for the work?*	**No**	Was the work subject to academic peer review?	**Yes**

**If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.*

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	
Please list the paper's authors in the intended authorship order:	
Stage of publication	Choose an item.

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	I conceived of the idea for the analysis, led on all aspects of the fieldwork, data collection and analysis. I wrote the first and final drafts of the paper.

Student Signature: [redacted] **Date:** 30/6/2017

Supervisor Signature: [redacted] **Date:** 30/6/2017

Original contribution

Challenges in the estimation of the annual risk of *Mycobacterium tuberculosis* infection in children aged under 5 years

Khan PY, Glynn JR, Mzembe T, Mulawa D, Chiumya R, Crampin AC, Kranzer K, Fielding KL

Corresponding author:

Dr Palwasha Yousafzai Khan

MBBCh (Hons), MRCS, MRCP, DTM&H, MSc

Department of Infectious Disease Epidemiology

London School of Hygiene and Tropical Medicine

Keppel Street, London, WD1E 7HT, United Kingdom

Mobile: +44 (0) 7811 902 455

Fax: +44 (0) 207 636 8739

Email: palwasha.khan@lshtm.ac.uk

Running head: Challenges in the estimation of ARTI in children

ABSTRACT

Accurate estimates of *Mycobacterium tuberculosis (M.tb)* infection in young children provide a critical indicator of on-going community *M.tb* transmission. Cross reactions due to infection with environmental mycobacteria and/or bacille Calmette-Guérin (BCG) vaccination compromise the estimates derived from population-level tuberculin skin test surveys using traditional cut-off methods. Newer statistical approaches are prone to failure of model convergence, especially in settings where the prevalence of *M.tb* infection is low and environmental sensitisation is high. We conducted a tuberculin skin test survey in 5119 pre-school children in the general population and among household contacts of tuberculosis cases in 2012-2014 in the same district in northern Malawi, where sensitisation to environmental mycobacteria is common and almost all children are BCG-vaccinated. We compared different proposed methods of estimating *M.tb* prevalence, including a method described by Rust and Thomas more than 40 years ago. With the different methods, estimated prevalence in the general population was 0.7%-11.9% at age <2 years and 0.8-3.3% at age 2-4 years. The Rust and Thomas method was the only method to give a lower estimate in the younger age group (0.7% vs 0.8%), suggesting it was the only method which adjusted appropriately for the marked effect of BCG-attributable induration in the very young.

Word count: 200

Keywords: *Mycobacterium tuberculosis;* infection prevalence; tuberculin skin test; annual risk of tuberculous infection (ARTI); children; BCG; Malawi

Abbreviations:

M.tb	*Mycobacterium tuberculosis*
TST	tuberculin skin test
ARTI	annual risk of *M.tb* infection
BCG	bacille Calmette-Guérin
KHDSS	Karonga Health and Demographic Surveillance Site
IGRA	interferon-gamma release assay
HIV	human immunodeficiency virus

Introduction

Childhood tuberculosis has not been considered a priority in high burden settings until recent years.[1] Children have pauci-bacillary disease and are unlikely to contribute to onward transmission of *Mycobacterium tuberculosis* (*M.tb*).[2] This has led to significant under-reporting of paediatric tuberculosis.[1] However childhood tuberculosis and *M.tb* infection in the very young necessarily result from recent transmission, so accurate estimates could provide a critical indicator of the effectiveness of prevention programmes to curtail on-going community *M.tb* transmission.[3,4]

Historically, measurements of the global burden of tuberculosis, including the incidence of tuberculosis disease have been, in part, inferred from estimates of the annual risk of *M.tb* infection (ARTI) as derived from *M.tb* infection prevalence data obtained from tuberculin skin test (TST) surveys in school-age children.[5-7] Direct estimates of tuberculosis disease incidence would require prohibitively large longitudinal cohorts, even in areas where the burden of disease is high.[8] Hence the comparatively cheaper and logistically simpler TST surveys were undertaken on a global programmatic scale. The inference of tuberculosis disease incidence from ARTI was based on Styblo's rule, where a 1% ARTI risk corresponds to 50 incident tuberculosis cases per 100,000 population per year.[5] It is now recognised that accurate estimates of incidence of tuberculosis cases using the 'Styblo' method are not valid,[9] though trend estimates of ARTI based on tuberculin surveys can be useful.[10-13]

The TST measures the immunological response to a previously acquired infection with a mycobacterium that shares antigens with those in tuberculin. The challenge is to disentangle reactions due to *M.tb* infection from reactions due to exposure to environmental mycobacteria and bacille Calmette-Guérin (BCG) vaccination.[12] Despite the lack of specificity of the TST,[14] and because of the cost and logistical issues (need for venepuncture, skilled personnel and laboratory equipment)[15] and the lack of clarity around the conversion and reversion phenomena associated with serial testing of the more specific interferon-gamma release assays (IGRAs),[16] serial population-wide tuberculin surveys undertaken in young children in high burden countries remain as one of the few ways to assess the impact of tuberculosis control strategies over time. However, this assessment relies on the need for a consistent estimate of *M.tb* infection prevalence which is not always possible with the traditional cut-off methods, especially in settings where cross-reactivity with environmental mycobacteria and BCG-attributable reactions are common.[11,17]

Despite the advent of sophisticated statistical techniques, such as latent variable modelling,[18] ascertainment of the prevalence of *M.tb* infection using tuberculin data is not always possible; failures of the model to converge are frequent, especially in areas where there is a moderate to strong influence of cross-reactions and low prevalence of *M.tb* infection.[12] An alternative method to estimate the prevalence of *M.tb* infection was published by Rust and Thomas forty years ago,[19] using tuberculin data from US Navy recruits. The authors stated that their proposed approach should "become even more preferable in the years to come" as the prevalence of *M.tb* infection continued to decrease compared

with the prevalence of infection with environmental mycobacteria, which is likely to remain constant.

We aim to determine the prevalence of *M.tb* infection and the ARTI in recently BCG-vaccinated pre-school children in rural Malawi using the model proposed by Rust and Thomas. We compare these estimates with those derived using the classical TST cut-off methods (≥10mm and ≥15mm), fixed mirror,[6,20] and mixture analysis.[21-24]

Methods

Study setting

Karonga district, northern Malawi is predominantly rural with an adult human immunodeficiency virus (HIV) prevalence around 9%, and incidence of new smear-positive tuberculosis of 87/100,000 adults per year.[25] BCG vaccination is administered to children on first health system contact (usually birth) as part of the Expanded Programme on Immunisation. The whole population (~39,000 people) in an area in the south of the district is under demographic surveillance in the Karonga Health and Demographic Surveillance Site (KHDSS).[26]

Study participants

Population at low risk of M.tb infection

We conducted a population-wide TST survey in pre-school children in 2012, nested in the KHDSS. All children aged 3 months to 4 years old resident in the KHDSS area at the time of household recruitment were eligible to take part in the study.

Population at high risk of M.tb infection

We also conducted a district-wide cross-sectional tuberculosis household case-contact study from January 2013 to December 2014. Household contacts, including children < 5 years, of an adult with smear-positive pulmonary tuberculosis were tuberculin tested.

Study procedures

Field staff were trained in the placement and reading of skin tests according to standard international guidelines.[27] Two international units of tuberculin purified protein derivative RT23 (Statens Serum Institut, Copenhagen, Denmark) were injected into the volar surface of the forearm, and induration was measured 48-72 hours later. The transverse and longitudinal diameter of the induration was recorded to the nearest millimetre, and an average calculated.[20]

Children with TST≥10mm were assessed for tuberculosis-related symptoms by field staff, and the results were recorded in the child's health passport. Any child with symptoms suggestive of tuberculosis (fever, weight loss, failure to thrive, night sweats or cough) was reviewed by a clinician and referred to the district hospital. All children with TST≥15mm were commenced on 6-month isoniazid preventive treatment (10mg/kg once daily) after active disease had been excluded.

Demographic data on sex, household size, household socioeconomic status (using a composite score based on quality of dwelling place, number of assets, employment of head of household, food security and availability of soap) and

maternal HIV status were collected for all study participants. HIV status of the child was not ascertained unless clinically indicated. Written informed consent was obtained from a parent or guardian of each participating child.

Ethics approval

Both studies were approved by the Malawi National Health Sciences Research Committee (study protocols #944 and #1049) and the London School of Hygiene and Tropical Medicine Ethics committee (study protocols #6065 and #6285).

Statistical analyses

The frequency distributions of induration for (i) *population at **higher** risk of M.tb infection:* children aged <5 years resident in the household of an adult with smear-positive pulmonary tuberculosis, and (ii) *population at **lower** risk of M.tb infection:* children aged <5 years resident in the KHDSS, excluding 20 children who had known direct household contact with tuberculosis, were tabulated using 2mm categories to minimise the number of categories with no data.

We then used four methods to estimate *M.tb* infection prevalence in both populations:

1. Rust and Thomas method

This method is based on the distribution of induration size (mm) rather than a classification system of defining individuals as positive, negative or doubtful. The technique was originally applied to a well-defined population of white male US Navy recruits aged 17-21 years old, who had been lifetime US residents. This

population was then divided into two groups, those with known household exposure to a tuberculosis case based on self-report, defined as 'high risk' and those without such exposure, defined as 'low risk'.

The Rust and Thomas method is built on a simple mathematical model. The underlying assumptions are that:

- the population can be divided into two groups which differ only in the prevalence of the infection
- there is an identifiable category in which no individual has *M.tb* infection (TST = 0mm)
- there is an identifiable category in which all individuals have *M.tb* infection (TST ≥ *n* mm)

The rationale of the Rust and Thomas model is as follows: in a hypothetical population without any *M.tb* infection, the majority will have a TST of zero induration. If sensitisation to environmental mycobacteria and/or recent BCG vaccination is prevalent, reactions of moderate size will also occur. The distribution of this population is called the "non-infected" distribution, referring to the absence of infection with *M.tb.* Comparably, in a hypothetical population in which everyone has been infected with *M.tb*, all but a few individuals will have a fairly large reaction size, and a bell-shaped "infected" distribution will be observed. In an existent population, the observed distribution will be a combination of "infected" and "non-infected" distributions. Thus, the observed 'higher risk' and 'lower risk' groups are each a mixture of overlapping

distributions of "infected" and "non-infected". If one observed the "non-infected" population alone, there would be a very large proportion with zero induration, and the proportion of "non-infected" with small- to medium-sized reactions would depend on the prevalence of nontuberculous mycobacterial infection and/or BCG-attributable induration, i.e. the distribution of reaction size depends upon sensitisation to environmental mycobacteria or recent BCG vaccination, but not upon contact status. Expressed in a different way, the ratio of the proportion with zero induration to the proportion "non-infected" is constant (Equation 1).[19]

Similarly, the distribution of reaction sizes among those who are truly infected is independent of contact status. Assuming that there is a reaction size (n) above which all individuals are infected, the ratio of the proportion in category n to the proportion "infected" is constant (Equation 2).

Equation 1 for "non-infected" population

$$\frac{f_0}{1-P} = \frac{f'_0}{1-P'}$$

Equation 2 for "infected" population

$$\frac{f_n}{P} = \frac{f'_n}{P'}$$

where f_0 = proportion of 'higher risk' population with zero induration

f'_0 = proportion of 'lower risk' population with zero induration

f_n = proportion of 'higher risk' population in induration category n

f'_n = proportion of 'lower risk' population in induration category n

P = prevalence of *M.tb* infection in the 'higher risk' population

P' = prevalence of *M.tb* infection in the 'lower risk' population

Equations 1 and 2 can then be solved for P and P'; the prevalence of *M.tb* infection in the 'higher risk' and 'lower risk' respectively.

$$P = \frac{fn\,(f'0 - f0)}{fnf'0 - f'nf0}$$

$$P' = \frac{f'n\,(f'0 - f0)}{fnf'0 - f'nf0}$$

A TST reaction size of ≥ 20mm was chosen as the category *n* in which all individuals were assumed to have *M.tb* infection. This category was chosen following examination of the prevalence of infection calculated for different values of *n*. The optimal choice being that category *n* in which the computed prevalence is approximately the same as that for higher values of *n*.[19] (Web table 1 and Web appendix 1 to see why a reaction size of 20mm was chosen).

Bias-corrected 95% confidence intervals were calculated using a bootstrapping approach in Stata version 14.1 (Stata Corporation, College Station, TX, USA).

2. Fixed cut-off points at 10mm or 15mm

M.tb infection prevalence was calculated as the proportion of children with a "positive" TST defined by these cut-offs, divided by the total number of children with a TST result

3. Fixed mirror method (17mm)

The fixed mirror method assumes that in individuals with *M.tb* infection the distribution of induration size is symmetric around a fixed mode of 17mm, and

that no non-specific reactions, such as BCG-attributable induration, reach 17mm.[6,20] Therefore all reactions of 17mm were counted once and indurations of >17mm were counted twice and summed to obtain the estimated number of *M.tb* infections. Prevalence was calculated as the count of '*M.tb* infections' divided by the total number of children with TST results.[24]

4. Mixture analysis

Mixture analysis of the tuberculin survey data, which is a form of latent variable modelling,[18] was based on implementation of the Bayesian Markov Chain Monte Carlo approach in R (R Foundation for Statistical Computing, Vienna).[28] Three parametric models (Weibull, lognormal and gamma distributions) were tested to determine the best fit model using the maximum log likelihood function as a guide. The quality of the fit was assessed by comparing predicted and observed frequencies via posterior predictive model checks.[24,28]

Sensitivity analyses

The effect of neonatal BCG vaccination on TST induration, which is most pronounced in the first few months after vaccination, is thought to wane rapidly.[29,30] The analysis was repeated, stratifying children into under 2 years and ≥ 2 years to assess the effect of BCG-attributable induration on estimates of *M.tb* prevalence.

Annual risk of M.tb infection (ARTI)

The ARTI, the probability of being infected in any one year, was calculated using the formula:[31]

$$\text{ARTI} \approx 1 - (1 - \text{P})^{1/\text{a}}$$

where P= prevalence of *M.tb* infection, a= the mean age of children. The ARTI was only calculated for the children resident in the KHDSS, which was assumed to be representative of the ARTI in children aged <5 years in the district.

Results

The frequency distribution of tuberculin data from the 'lower risk' and 'higher risk' study populations are shown in Table 1 and the Web figures 1-4. Among all children <5 years, 85% of the 'lower risk' population had zero induration compared to 56% of the 'higher risk' population (p<0.001). When stratifying by age, the proportion with zero induration in the 'lower risk' group was 92% in those aged ≥2 years compared to 73% in those aged <2 years. In the 'higher risk' group the proportion with zero induration was 54% in those aged ≥2 years and 60% in those <2 years. There was no evidence that distribution of induration size was affected by HIV exposure status of the child (chi-squared test: 'lower' risk group p=0.9; 'higher' risk group p=0.8). Web figure 4 illustrates the percentage distribution of induration category stratified by HIV exposure status and contact status.

Prevalence of *M.tb* infection

Table 2 shows the estimated prevalence of *M.tb* infection using the different methods. In the 'lower risk' group the estimates of infection prevalence were consistently higher in the under-2s compared to those aged ≥2 years in all methods *except* for the Rust and Thomas model. For the under-2s the estimates

ranged from 0.7% to 11.9%; the mixture model and the cut-off TST ≥10mm method estimated the highest infection prevalence (11 – 12%). Although the fixed mirror method estimated a similar infection prevalence for ≥ 2years to the Rust and Thomas method, in the under-2s the infection prevalence was nearly 3 times that of the Rust and Thomas method.

In the 'higher risk' group the estimates were higher in the children aged ≥ 2 years compared to the youngest age group for all methods. The estimates for the older age group in the 'higher risk' group were similar for all methods, ranging from 39.9 % to 42.0%, except for the TST cut-off ≥15mm method, which estimated a prevalence of *M.tb* infection of 31.9%. The bias-corrected 95% confidence interval of Rust and Thomas model for the 'higher risk' children aged < 2 years includes 0. This is likely to be a result of the small sample size, n=52, in this group.

ARTI

The estimates of the ARTI ranged from 0.3% (95% CI: 0.1 – 0.9%) to 2.6% (95% CI: 1.8 – 2.7) depending on the method used to estimate the prevalence of *M.tb* infection. The Rust and Thomas model estimate was the most conservative at 0.3% (95% CI: 0.1 – 0.9%). It was also the method which demonstrated the least difference in ARTI estimates when stratified by age (see Table 3).

Discussion

Our findings highlight the challenges of using tuberculin surveys to estimate the risk of *M.tb* infection in young BCG-vaccinated children. ARTI estimates varied 5-

fold depending on the method used to estimate *M.tb* infection prevalence. The Rust and Thomas method generated a consistent estimate of *M.tb* infection prevalence and ARTI, irrespective of age, in a setting where sensitisation to environmental mycobacteria is known to be high [22], and over 90% of children are BCG-vaccinated within 3 months of birth. It was the only method which appeared to appropriately adjust for the marked effect of BCG-attributable induration in the very young (aged < 2 years).

The Rust and Thomas method protects against changes in prevalence estimates caused by differences in strength of tuberculin used or the use of different equipment and/or techniques, thus making it possible to compare *M.tb* infection prevalence found by different investigators at varying times and place.[19] Because the Rust and Thomas method relies on the distribution of induration in those known to have been exposed to *M.tb* and the distribution of induration in those at 'lower' risk at the same point in time, as long as the same tuberculin and technique is used in both populations, the prevalence estimates over time are much more likely to be comparable, despite differences in geographical settings, climate zones, changing BCG vaccination policies and introduction of new vaccines. In addition, the Rust and Thomas model can be used to generate the probability of *M.tb* infection at each induration size, therefore making it possible to calculate sensitivity and specificity, area under Receiver Operating Characteristic curve and the positive predictive value of the TST in a given population.[14] Another advantage compared to traditional cut-off methods is that prevalence estimates are less sensitive to alterations in the chosen critical point. For the Rust and Thomas method this is the reaction size category in which all

individuals are assumed to have *M.tb* infection. As long as this reaction size exceeds the maximum reaction size occurring among the 'non-infected', the calculated prevalence will only be subject to random fluctuations. However, if the reaction size is too small, the basic assumption that all individuals with reactions of that size or larger have been infected with *M.tb* will not be fulfilled and the estimated infection prevalence will therefore overestimate the true prevalence.[19]

One of the reasons that the method has been apparently forgotten may be the requirement of tuberculin data from 'low risk' and 'high risk' groups. The US Navy recruits study, used in the original study, used self-report of household contact with a tuberculosis case.[32] In our study, we combined tuberculin data from a TST survey conducted in a demographic surveillance area and data from a concurrent tuberculosis case-contact household study in the whole district. The demographic surveillance area may not be representative of the whole district: research has been conducted in KHDSS for the last 12 years, which may have influenced health-seeking behaviour which in turn may affect *M.tb* transmission dynamics in the area. One of the major assumptions of the Rust and Thomas model is that the 'high risk' and 'low risk' populations only differ with respect to contact status and therefore prevalence of infection. Reassuringly in our study the two groups did not differ significantly with regards to age, sex, household size, household socioeconomic status and maternal HIV status (see Web appendix 2 and Web table 2).

In the 'lower risk' group the proportion of children with a TST≥20mm (our chosen *n*th category) was larger in the under-2s (0.3%) than in the over-2s (0.2%) in our study. If these are not due to *M.tb* infection it would violate the assumption that only those truly infected are included in the *n*th category and would therefore **over-estimate** the infection prevalence. Similarly for the fixed mirror method, any induration size ≥ 17mm due to BCG-attributable induration rather than true *M.tb* infection prevalence would over-estimate infection prevalence. Very large induration secondary to BCG vaccination is more likely to occur in the under-2s who have been more recently vaccinated. Interestingly, a Taiwanese study which proposed age-specific cut-offs to detect *M.tb* infection in children, suggested a cut-off of 21mm for infants aged less than 2 years.[33] Data in the larger induration (>20mm) categories were sparse and misclassification of a small number has a large effect on the resultant proportion in the *n*th category, which is a limitation of the data. The 95% confidence interval of the prevalence of *M.tb* infection in the 'high risk' children aged less than 2 years using the Rust and Thomas method included 0, also underscoring the importance of an adequate sample size. A similar study among older children, adolescents and young adults, who are likely to be at greater risk for *M.tb* infection compared to young children,[20,34,35] would be useful to assess the robustness of the Rust and Thomas method. It would require a household contact study as well as a "low risk" population survey which would have cost implications.

Our findings present evidence that the Rust and Thomas method appears to address the effect of recent BCG vaccination in the under-2s. In the over-2s, the results of the different methods vary less and are all likely to be plausible but

because the Rust and Thomas method performed more appropriately in dealing with the cross-reactions due to BCG in the younger age group, we can have confidence that it is dealing appropriately with cross-reactivity in the older age group as well.

One may ask the question, why should we continue to advocate the use of tuberculin in an era of more specific diagnostics such as IGRAs and newer skin tests, such as the C-Tb skin test, a novel skin test containing ESAT-6 and CFP-10, antigens which are specific to *M.tb*?[36] The latest skin tests, for which there is currently limited data, do appear to offer higher specificity than tuberculin, but this might come at the cost of reduced sensitivity.[37] The cost, technical complexity and the requirement of a laboratory infrastructure in order to undertake large IGRA surveys usually precludes population-level studies. However, IGRA sub-studies nested within tuberculin surveys could *potentially* be used to refine estimates of *M.tb* infection prevalence [12,38,39], although it is not clear how discrepancies between TST and IGRA should be interpreted. In longitudinal studies, IGRA and TST responses seem to convert and revert at different rates, so the two tests are unlikely to give the same assessment of infection in any population.[40,41]

We wanted to estimate the ARTI in pre-school children based on the rationale that determination of the average ARTI in the very young provides a critical indicator of the extent of recent *M.tb* transmission. It is important to bear in mind that risk of *M.tb* infection is not independent of age,[20,34,42] and is most likely related to *M.tb* exposure through age-assortative social mixing.[43] Thus the

average ARTI in the youngest within a population is unlikely to be representative of the ARTI in those that are older, but it does provide the most contemporary marker of recent *M.tb* transmission. Repeated tuberculin surveys in the youngest generation could potentially be used to assess whether implementation of tuberculosis control strategies within the community have resulted in a decrease of recent *M.tb* transmission.[11]

Conclusion

There is unequivocally a need for more accurate epidemiological indicators of *M.tb* transmission and *M.tb* infection prevalence estimates in order to understand the dynamics of the tuberculosis epidemic in varying settings [44]. In our study, the Rust and Thomas method was the only method to find a lower estimate in the youngest age group, suggesting that it accounted appropriately for the cross-reactivity due to BCG vaccination.

Word count: 3702

Acknowledgements

Authors affiliations:

Khan PY (Department of Infectious Disease Epidemiology, Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, United Kingdom; Karonga Prevention Study/MEIRU, Karonga district, Malawi); Glynn JR (Department of Infectious Disease Epidemiology, Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, United Kingdom); Mzemba T (Karonga Prevention Study/MEIRU, Karonga district, Malawi); Mulawa D (Karonga Prevention

Study/MEIRU,_Karonga district, Malawi); Chiumya R (Karonga Prevention Study/MEIRU, Karonga district, Malawi); Crampin AC (Department of Infectious Disease Epidemiology, Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, United Kingdom; Karonga Prevention Study/MEIRU, Karonga district, Malawi); Kranzer K (National and Supranational Mycobacterium Reference Laboratory, Forschungszentrum Borstel, Germany; Department of Infectious Disease Epidemiology, Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, United Kingdom); Fielding KL (Department of Infectious Disease Epidemiology, Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, United Kingdom)

Conflicts of interest:

There are no conflicts of interest (including financial and other relationships) for the authors.

Funding:

This work was supported by a Wellcome Trust Strategic Award for the Karonga Prevention Study [grant number 098610/Z/12/Z] and a Wellcome Trust clinical research training fellowship [PYK: grant number 100137/Z/12/Z].

The authors would like to thank the field team in Karonga (Mr Taniel Njawala; Ms Ephrida Mwiba; Mr Tota Mwafulirwa; Mr Glyn Msiska; Mr Khalanga Mwenechanya; Mr Taniel Njawala; Mr Kasimba Simplex and Mr Michael

Mwenibabu) and the data/IT team in London (Mr Keith Branson and Ms Jacky Saul), without their input this research would not have been possible.

REFERENCES

1. Dodd PJ, Gardiner E, Coghlan R, Seddon JA. Burden of childhood tuberculosis in 22 high-burden countries: a mathematical modelling study. The Lancet Global health. 2014;2(8):e453-459. doi: 10.1016/S2214-109X(14)70245-1. PubMed PMID: 25103518.
2. Marais BJ, Gie RP, Schaaf HS, Beyers N, Donald PR, Starke JR. Childhood pulmonary tuberculosis: old wisdom and new challenges. Am J Respir Crit Care Med. 2006;173(10):1078-1090. doi: 10.1164/rccm.200511-1809SO. PubMed PMID: 16484674.
3. Bloch AB, Snider DE, Jr. How much tuberculosis in children must we accept? Am J Public Health. 1986;76(1):14-15. PubMed PMID: 3940448; PubMed Central PMCID: PMC1646406.
4. Marais BJ, Obihara CC, Warren RM, Schaaf HS, Gie RP, Donald PR. The burden of childhood tuberculosis: a public health perspective. Int J Tuberc Lung Dis. 2005;9(12):1305-1313. Epub 2006/02/10. PubMed PMID: 16466051.
5. Styblo K. The relationship between the risk of tuberculous infection and the risk of developing infectious tuberculosis. Bull Int Union Tuberc. 1985;60(3):117-119.
6. Bleiker MA, Sutherland I, Styblo K, ten Dam HG, Misljenovic O. Guidelines for estimating the risks of tuberculous infection from tuberculin test results in a representative sample of children. Bull Int Union Tuberc Lung Dis. 1989;64(2):7-12. PubMed PMID: 2790288.
7. Cauthen GM, Pio A, ten Dam HG. Annual risk of tuberculous infection. 1988. Bull World Health Organ. 2002;80(6):503-511; discussion 1-2. PubMed PMID: 12132011; PubMed Central PMCID: PMC2567543.
8. Yates TA, Khan PY, Knight GM, Taylor JG, McHugh TD, Lipman M, et al. The transmission of Mycobacterium tuberculosis in high burden settings. Lancet Infect Dis. 2016;16(2):227-238. doi: 10.1016/S1473-3099(15)00499-5. PubMed PMID: 26867464.
9. van Leth F, van der Werf MJ, Borgdorff MW. Prevalence of tuberculous infection and incidence of tuberculosis: a re-assessment of the Styblo rule. Bull World Health Organ. 2008;86(1):20-26. Epub 2008/02/01. doi: S0042-96862008000100011 [pii]. PubMed PMID: 18235886; PubMed Central PMCID: PMC2647347.
10. Borgdorff M. Annual risk of infection- time for an update? Bull World Health Organ. 2002;80(6):501-502.
11. Rieder H. Annual risk of infection with Mycobacterium tuberculosis. Eur Respir J. 2005;25(1):181-185. Epub 2005/01/11. doi: 25/1/181 [pii] 10.1183/09031936.04.00103804. PubMed PMID: 15640340.
12. Rieder HL, Chadha VK, Nagelkerke NJ, van Leth F, van der Werf MJ, Foundation KT. Guidelines for conducting tuberculin skin test surveys in high-prevalence countries. Int J Tuberc Lung Dis. 2011;15 Suppl 1:S1-S25. PubMed PMID: 21276325.

13. Dye C. Breaking a law: tuberculosis disobeys Styblo's rule [editorial]. Bull World Health Organ. 2008;86(1):4. PubMed PMID: 18235879; PubMed Central PMCID: PMC2647350.
14. Rieder HL. Methodological issues in the estimation of the tuberculosis problem from tuberculin surveys. Tuber Lung Dis. 1995;76(2):114-121. Epub 1995/04/01. doi: 0962-8479(95)90552-9 [pii]. PubMed PMID: 7780092.
15. World Health Organization. Use of tuberculosis interferon-gamma release assays (IGRAs) in low- and middle-income countries: policy statement (WHO/HTM/TB/2011.18). Geneva: World Health Organisation; 2011. http://www.who.int/tb/features_archive/policy_statement_igra_oct2011.pdf. Accessed December 1, 2016
16. Trajman A, Steffen RE, Menzies D. Interferon-Gamma Release Assays versus Tuberculin Skin Testing for the Diagnosis of Latent Tuberculosis Infection: An Overview of the Evidence. Pulmonary medicine. 2013;2013:601737. doi: 10.1155/2013/601737. PubMed PMID: 23476763; PubMed Central PMCID: PMC3582085.
17. Neuenschwander BE, Zwahlen M, Kim SJ, Lee EG, Rieder HL. Determination of the prevalence of infection with Mycobacterium tuberculosis among persons vaccinated against Bacillus Calmette-Guerin in South Korea. Am J Epidemiol. 2002;155(7):654-663. PubMed PMID: 11914193.
18. Sismanidis C, Williams B. Detecting LTBI: old techniques, new approaches. Int J Tuberc Lung Dis. 2008;12(8):867-868. PubMed PMID: 18647444.
19. Rust P, Thomas J. A method for estimating the prevalence of tuberculosis infection. Am J Epidemiol. 1975;101(4):311-322. PubMed PMID: 1124761.
20. Fine PE, Bruce J, Ponnighaus JM, Nkhosa P, Harawa A, Vynnycky E. Tuberculin sensitivity: conversions and reversions in a rural African population. Int J Tuberc Lung Dis. 1999;3(11):962-975. Epub 1999/12/10. PubMed PMID: 10587318.
21. Neuenschwander BE, Zwahlen M, Kim SJ, Engel RR, Rieder HL. Trends in the prevalence of infection with mycobacterium tuberculosis in Korea from 1965 to 1995: an analysis of seven surveys by mixture models. Int J Tuberc Lung Dis. 2000;4(8):719-729. Epub 2000/08/19. PubMed PMID: 10949323.
22. Davies GR, Fine PE, Vynnycky E. Mixture analysis of tuberculin survey data from northern Malawi and critique of the method. Int J Tuberc Lung Dis. 2006;10(9):1023-1029. Epub 2006/09/13. PubMed PMID: 16964795.
23. Villate JI, Ibanez B, Cabriada V, Pijoan JI, Taboada J, Urkaregi A. Analysis of latent tuberculosis and mycobacterium avium infection data using mixture models. Bmc Public Health. 2006;6. doi: 240 10.1186/1471-2458-6-240.
24. Shanaube K, Sismanidis C, Ayles H, Beyers N, Schaap A, Lawrence KA, et al. Annual risk of tuberculous infection using different methods in communities with a high prevalence of TB and HIV in Zambia and South Africa. PLoS One. 2009;4(11):e7749. Epub 2009/11/17. doi: 10.1371/journal.pone.0007749. PubMed PMID: 19915666; PubMed Central PMCID: PMC2771909.

25. Mboma SM, Houben RM, Glynn JR, Sichali L, Drobniewski F, Mpunga J, et al. Control of (Multi)Drug Resistance and Tuberculosis Incidence over 23 Years in the Context of a Well-Supported Tuberculosis Programme in Rural Malawi. PLoS One. 2013;8(3):e58192. Epub 2013/03/14. doi: 10.1371/journal.pone.0058192. PubMed PMID: 23483994; PubMed Central PMCID: PMCPMC3590148.
26. Crampin AC, Dube A, Mboma S, Price A, Chihana M, Jahn A, et al. Profile: the Karonga Health and Demographic Surveillance System. Int J Epidemiol. 2012;41(3):676-685. Epub 2012/06/26. doi: 10.1093/ije/dys088. PubMed PMID: 22729235; PubMed Central PMCID: PMCPMC3396313.
27. Arnadottir T, Rieder HL, Trebucq A, Waaler HT. Guidelines for conducting tuberculin skin test surveys in high prevalence countries. Tuber Lung Dis. 1996;77 Suppl 1:1-19. PubMed PMID: 8759471.
28. Neuenschwander BF. Bayesian Mixture Analysis for Tuberculin Induration Data. http://www.tbrieder.org/research/mixture/mixture_documentation.pdf. Published July 2007. Accessed December 1, 2016.
29. Menzies D. What does tuberculin reactivity after bacille Calmette-Guerin vaccination tell us? Clin Infect Dis. 2000;31 Suppl 3:S71-S74. Epub 2000/09/30. doi: 10.1086/314075. PubMed PMID: 11010826.
30. Floyd S, Ponnighaus JM, Bliss L, Nkhosa P, Sichali L, Msiska G, et al. Kinetics of delayed-type hypersensitivity to tuberculin induced by bacille Calmette-Guerin vaccination in northern Malawi. J Infect Dis. 2002;186(6):807-814. doi: 10.1086/342416. PubMed PMID: 12198615.
31. Rieder HL. Epidemiological basis of tuberculosis control. Paris: International Union Against Tuberculosis and Lung Disease; 1999.
32. Edwards LB, Palmer CE. Tuberculosis. II. Tuberculous Infection. Lowell AM, editor. Cambridge, Mass.: Harvard University Press; 1969.
33. Chan PC, Chang LY, Wu YC, Lu CY, Kuo HS, Lee CY, et al. Age-specific cut-offs for the tuberculin skin test to detect latent tuberculosis in BCG-vaccinated children. Int J Tuberc Lung Dis. 2008;12(12):1401-1406. Epub 2008/11/20. PubMed PMID: 19017449.
34. Narain R, Nair SS, Chandrasekhar P, Rao GR. Problems connected with estimating the incidence of tuberculosis infection. Bull World Health Organ. 1966;34(4):605-622. PubMed PMID: 5296384; PubMed Central PMCID: PMC2475992.
35. Middelkoop K, Bekker LG, Morrow C, Lee N, Wood R. Decreasing household contribution to TB transmission with age: a retrospective geographic analysis of young people in a South African township. BMC Infect Dis. 2014;14(1):221. doi: 10.1186/1471-2334-14-221. PubMed PMID: 24758715; PubMed Central PMCID: PMC4012060.
36. Hoff ST, Peter JG, Theron G, Pascoe M, Tingskov PN, Aggerbeck H, et al. Sensitivity of C-Tb: a novel RD-1-specific skin test for the diagnosis of tuberculosis infection. Eur Respir J. 2016;47(3):919-928. doi: 10.1183/13993003.01464-2015. PubMed PMID: 26677940.

37. Pai M, Sotgiu G. Diagnostics for latent TB infection: incremental, not transformative progress. Eur Respir J. 2016;47(3):704-706. doi: 10.1183/13993003.01910-2015. PubMed PMID: 26929311.

38. Pai M, Dendukuri N, Wang L, Joshi R, Kalantri S, Rieder HL. Improving the estimation of tuberculosis infection prevalence using T-cell-based assay and mixture models. International Journal of Tuberculosis and Lung Disease. 2008;12(8):895-902.

39. Dodd PJ, Millington KA, Ghani AC, Mutsvangwa J, Butterworth AE, Lalvani A, et al. Interpreting tuberculin skin tests in a population with a high prevalence of HIV, tuberculosis, and nonspecific tuberculin sensitivity. Am J Epidemiol. 2010;171(9):1037-1045. Epub 2010/04/13. doi: kwq017 [pii] 10.1093/aje/kwq017. PubMed PMID: 20382638; PubMed Central PMCID: PMC2858871.

40. Dye C, Bassili A, Bierrenbach AL, Broekmans JF, Chadha VK, Glaziou P, et al. Measuring tuberculosis burden, trends, and the impact of control programmes. Lancet Infect Dis. 2008;8(4):233-243. doi: 10.1016/S1473-3099(07)70291-8. PubMed PMID: 18201929.

41. Pai M, Denkinger CM, Kik SV, Rangaka MX, Zwerling A, Oxlade O, et al. Gamma Interferon Release Assays for Detection of Mycobacterium tuberculosis Infection. Clin Microbiol Rev. 2014;27(1):3-20. doi: 10.1128/CMR.00034-13. PubMed PMID: 24396134; PubMed Central PMCID: PMC3910908.

42. Wood R, Liang H, Wu H, Middelkoop K, Oni T, Rangaka MX, et al. Changing prevalence of tuberculosis infection with increasing age in high-burden townships in South Africa. Int J Tuberc Lung Dis. 2010;14(4):406-412. Epub 2010/03/06. PubMed PMID: 20202297; PubMed Central PMCID: PMC2837545.

43. Dodd PJ, Looker C, Plumb ID, Bond V, Schaap A, Shanaube K, et al. Age- and Sex-Specific Social Contact Patterns and Incidence of Mycobacterium tuberculosis Infection. Am J Epidemiol. 2016;183(2):156-166. doi: 10.1093/aje/kwv160. PubMed PMID: 26646292; PubMed Central PMCID: PMCPMC4706676.

44. Rieder HL. The dynamics of tuberculosis epidemiology. The Indian journal of tuberculosis. 2014;61(1):19-29. PubMed PMID: 24640341.

Table 1. Frequency Distribution of Tuberculin Data in Children Aged Under 5 years in Rural Malawi in 2012-2014 Stratified by Risk Group and Age

Induration size (mm)	Frequency distribution of tuberculin data					
	Lower risk group			Higher risk group		
	All < 5 years N=4,947	< 2 years N=1,797	≥2 years N=3,150	All < 5 years N=152	< 2 years N=52	≥2 years N=100
	n %	n %	n %	n %	n %	n %
0	4,187 84.6	1,301 72.5	2,886 91.6	85 55.9	31 59.6	54 54.0
2 – 3	26 0.5	10 0.6	16 0.5	3 2.0	1 1.9	2 2.0
4 – 5	72 1.4	37 2.0	35 1.1	2 1.3	2 3.8	0 0.0
6 – 7	160 3.2	105 5.8	55 1.8	4 2.6	3 5.8	1 1.0
8 – 9	191 3.9	137 7.6	54 1.7	2 1.3	1 1.9	1 1.0
10 – 11	99 2.0	60 3.3	39 1.2	4 2.6	1 1.9	3 3.0
12 – 13	88 1.8	62 3.5	26 0.8	8 5.3	3 5.8	5 5.0
14 – 15	64 1.3	49 2.7	15 0.5	9 5.9	2 3.8	7 7.0
16 – 17	32 0.6	22 1.2	10 0.3	11 7.2	2 3.8	9 9.0
18 – 19	19 0.4	9 0.5	10 0.3	13 8.6	2 3.8	11 11.0
20 – 21	3 0.1	1 0.1	2 0.1	8 5.3	3 5.8	5 5.0
22+	6 0.1	4 0.2	2 0.1	3 2.0	1 1.9	2 2.0

Table 2. ***Mycobacterium tuberculosis*** **Infection Prevalence Estimates and 95% Confidence Intervals Using Different Methods in Children Aged Under 5 years in 2012-2014 in Rural Malawi**

Method to estimate *M.tb* infection prevalence	LOWER RISK GROUP						HIGHER RISK GROUP					
	ALL <5 YEARS		< 2 YEARS		≥2 YEARS		ALL < 5 YEARS		< 2 YEARS		≥2 YEARS	
	N=4,947		N=1,797		N=3,150		N=152		N=52		N=100	
	%	95 % CI	%	95 % CI	%	95 % CI	%	95 % CI	%	95 % CI	%	95 % CI
TST cut-off												
10 mm	6.3	5.6, 7.0	11.5	10.1, 13.1	3.3	2.7, 4.0	38.6	30.3, 47.5	26.9	15.6, 41.0	41.8	31.5, 52.6
15mm	1.9	1.6, 2.4	3.4	2.6, 4.3	1.1	0.8, 1.5	28.8	21.2, 37.3	17.3	8.2, 30.3	31.9	22.5, 42.5
Fixed mirror	1.3	1.0, 1.7	2.0	1.4, 2.7	1.0	0.7, 1.4	36.8	29.2, 45.0	26.9	15.6, 41.0	42.0	32.2, 53.3
Mixture model	6.4	3.4, 9.9	12.4	4.4, 19.3	1.9	0.03, 4.7	33.4	24.2, 43.1	27.6	4.4, 47.8	39.9	24.4, 49.9
Rust & Thomas model	0.9	0.3, 2.4	0.7	0.1, 4.6	0.8	0.2, 2.6	34.5	25.6, 43.8	18.2	0.0, 35.8	41.5	30.2, 51.8

TST tuberculin skin test; CI confidence interval

Lower risk group is composed of children aged under 5 years resident in the demographic surveillance site

Higher risk group is composed of children aged under 5 years who are resident in the household of a smear-positive pulmonary TB index case

Table 3. Annual Risk of *Mycobacterium tuberculosis* infection (ARTI) Estimates Based on *Mycobacterium tuberculosis* Infection Prevalence Estimates in the Children Under 5 years in 2012-2014 Resident in the Demographic Surveillance Area in Rural Malawi

Method used to estimate prevalence of M.tb infection	**ARTI**					
	ALL <5 YEARS		**< 2 YEARS**		**≥ 2 YEARS**	
	%	95 % CI	%	95 % CI	%	95 % CI
TST cut-off						
10 mm	2.4	1.8, 2.7	10.2	6.3, 11.6	1.0	0.7, 1.2
15mm	0.7	0.5, 0.8	3.0	2.3, 3.8	0.3	0.2, 0.4
Fixed mirror	0.5	0.4, 0.7	1.8	1.2, 2.4	0.3	0.2, 0.4
Mixture model	2.6	1.4, 4.0	10.5	3.2, 17.1	0.6	0.1, 1.4
Rust & Thomas model	0.3	0.1, 0.9	0.6	0.1, 4.1	0.2	0.1, 0.8

TST tuberculin skin test; ARTI annual risk of *M.tb* infection; CI Confidence Interval

Web Material

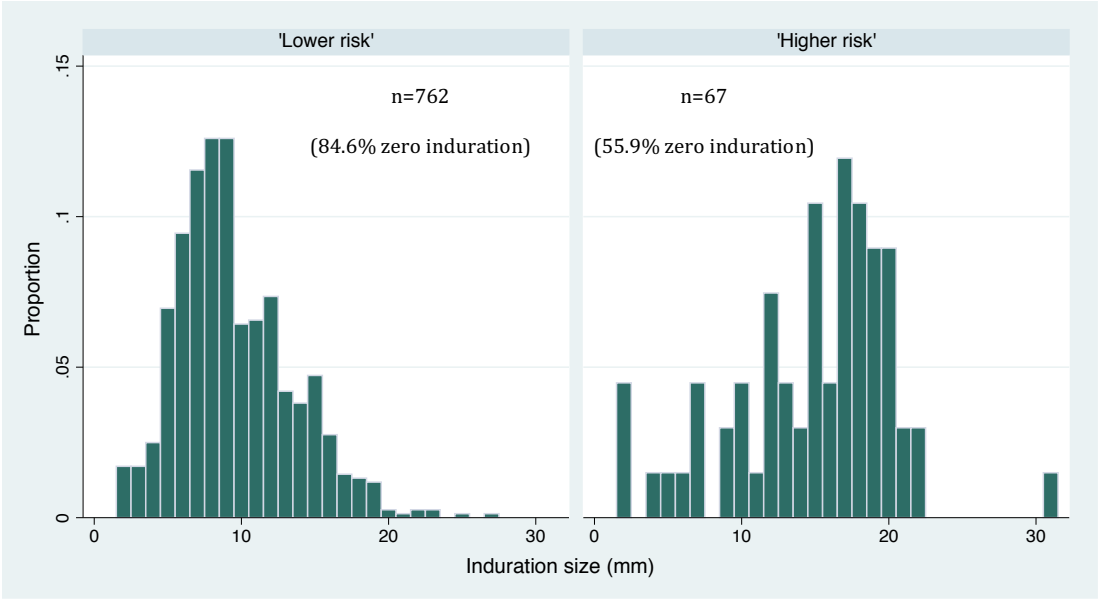
Web table 1. *M.tb* Infection Prevalence (%) Using Different Induration Categories for *n* and stratified by risk group and different age categories

		Induration (mm) considered to identify *M.tb* infection:						
		22+	**20+**	**18+**	**16+**	**14+**	**12+**	**10+**
'Higher risk' group (years)	All < 5	35.9	34.5	34.8	35.2	36.0	37.0	38.3
	< 2	19.5	18.2	18.7	19.8	22.1	24.1	27.2
	≥ 2	43.4	41.5	41.7	41.8	42.0	42.4	43.1
'Lower risk' group (years)	All < 5	2.9	0.9	1.2	1.8	3.1	4.6	6.5
	< 2	2.3	0.7	1.3	2.6	5.5	7.9	11.7
	≥2	3.9	0.8	1.0	1.2	1.6	2.3	3.4
	2.0 - 2.9	3.9	1.6	0.6	0.6	0.8	1.3	2.6
	3.0 – 3.9	-	-	1.2	1.2	1.4	1.8	3.0
	4.0 – 4.9	-	1.2	1.2	1.8	2.6	3.9	4.4

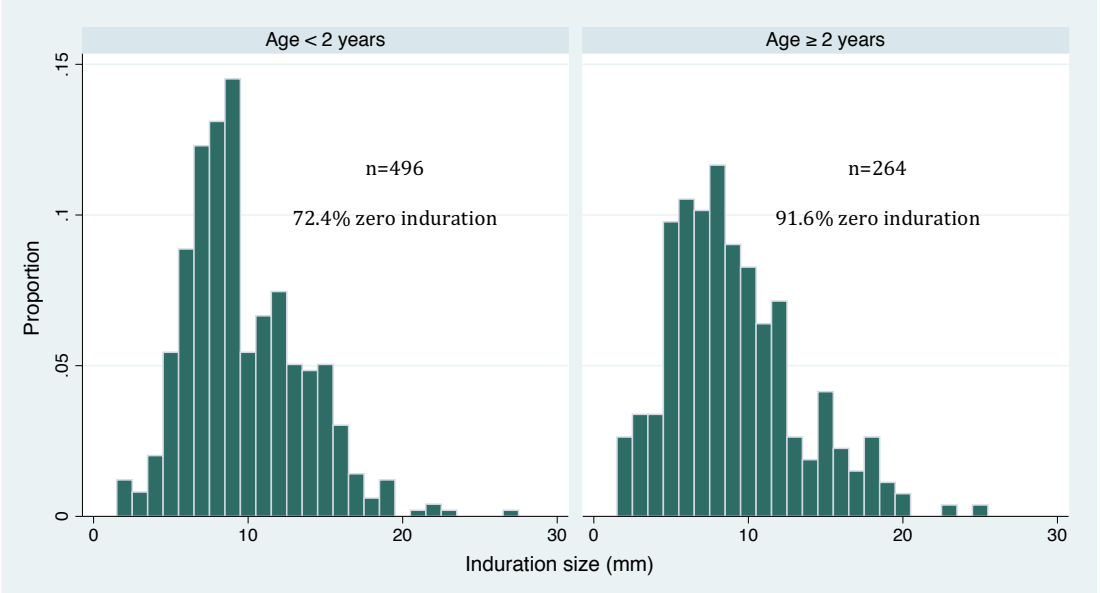
Web appendix 1

The prevalence of *M.tb* infection for each group was calculated for different reaction size categories using the equations for P and P' (see Methods) and are shown in table above (Web table 1). Infection prevalence estimates start to become 'approximately constant', as deduced from visual inspection, from 16+ category in all groups apart from the under-2s in the 'lower risk' group. This suggests that the 16+ category in the 'lower risk' under-2s includes individuals who are not truly 'infected'. In order to fulfil the assumption that all individuals in category *n* are 'infected', one must choose the largest induration size allowed by the data. In this dataset, reaction size of 20mm was chosen to represent *n* category as there was obvious instability of estimates for the 22+ category seen across all groups due to the small numbers in this category.

Web figure 1. Histogram illustrating the distribution of non-zero TST data in all children aged < 5 years stratified by risk group ('lower risk' N=4947; 'higher risk' N=152)

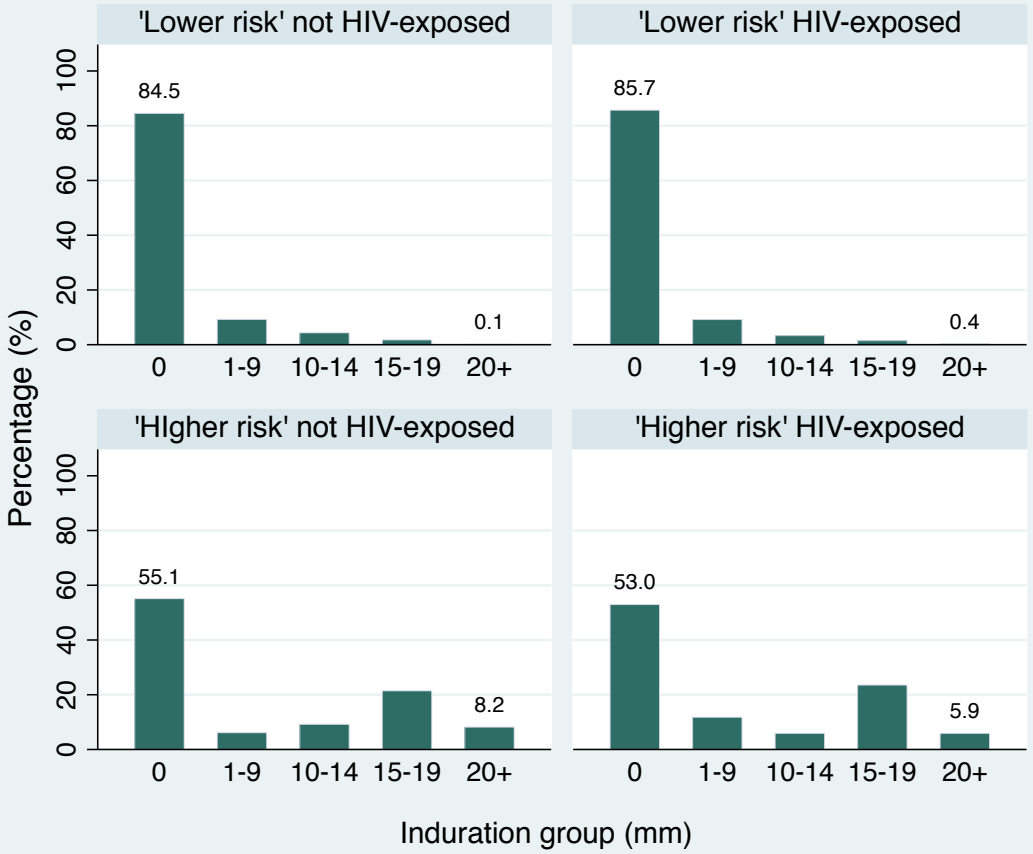


Web figure 2. Histogram illustrating the distribution of non-zero TST data in children in the 'lower risk' group stratified by age (<2 years N=1797; ≥2 years N=3150)



Web figure 3. Histogram illustrating the distribution of non-zero TST data in children in the 'higher risk' group stratified by age (<2 years N=52; ≥2 years N=100)

Web figure 4. Percentage distribution of induration category stratified by HIV exposure status and contact status of children ('Lower' risk not HIV-exposed n=4453; HIV-exposed n=272; 'higher' risk not HIV-exposed n=98; HIV-exposed n=17)



Web appendix 2

The Rust and Thomas method depends on the assumption that the two groups only differ with respect to contact status and prevalence of M.tb infection. Table (2) shows the demographic characteristics of the 'lower risk' and 'higher risk' group.

Web table 2. Demographic Characteristics of the 'Lower Risk' and 'Higher Risk' Groups

	'Lower risk'	'Higher risk'
Age in years (mean (sd))	2.6 (1.4)	2.7 (1.3)
Male (%)	50.0	48.5
Mother HIV positive (%)	5.5%	7.9%
	(95% CI 4.9 – 6.2)	(95% CI 4.1 – 13.4)
Median no. of adults in the HH (IQR)	2 (2 - 3)	2 (1 - 2)
Proportion in the lowest SES category	28.4%	23.4%

sd standard deviation; HH household; IQR interquartile range; CI confidence interval SES socioeconomic status

4 Prevalent *Mycobacterium tuberculosis* infection in children under 5 years

4.1 Research paper III

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT
www.lshtm.ac.uk

Registry
T: +44(0)20 7299 4646
F: +44(0)20 7299 4656
E: registry@lshtm.ac.uk

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

RESEARCH PAPER COVER SHEET

PLEASE NOTE THAT A COVER SHEET MUST BE COMPLETED FOR EACH RESEARCH PAPER INCLUDED IN A THESIS.

SECTION A – Student Details

Student	Palwasha Yousafzai Khan
Principal Supervisor	Professor Judith Glynn
Thesis Title	Investigating Mycobacterium tuberculosis transmission in rural Malawi

If the Research Paper has previously been published please complete Section B, if not please move to Section C

SECTION B – Paper already published

Where was the work published?	International Journal of Tuberculosis and Lung Disease		
When was the work published?	March 2016		
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion	N/A		
Have you retained the copyright for the work?*	**No**	Was the work subject to academic peer review?	**Yes**

**If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.*

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	
Please list the paper's authors in the intended authorship order:	
Stage of publication	Choose an item.

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	I wrote the study protocol and led on all aspects of the fieldwork, data collection and analysis. I wrote the first and final drafts of the paper.

Student Signature: [redacted] **Date:** 30/6/2017

Supervisor Signature: [redacted] **Date:** 30/6/2017

INT J TUBERC LUNG DIS 20(3):342–349

http://dx.doi.org/10.5588/ijtld.15.0672

Risk factors for *Mycobacterium tuberculosis* infection in 2–4 year olds in a rural HIV-prevalent setting

P. Y. Khan,*† J. R. Glynn,* K. L. Fielding,* T. Mzembe,† D. Mulawa,† R. Chiumya,† P. E. M. Fine,* O. Koole,*† K. Kranzer,‡ A. C. Crampin*†

*Department of Infectious Disease Epidemiology, London School of Hygiene & Tropical Medicine, London, UK; †Karonga Prevention Study, Chilumba, Malawi; ‡National and Supranational Mycobacterium Reference Laboratory, Forschungszentrum Borstel, Borstel, Germany

SUMMARY

BACKGROUND: *Mycobacterium tuberculosis* infection in children acts as a sentinel for infectious tuberculosis.
OBJECTIVE: To assess risk factors associated with tuberculous infection in pre-school children.
METHOD: We conducted a population-wide tuberculin skin test (TST) survey from January to December 2012 in Malawi. All children aged 2–4 years residing in a demographic surveillance area were eligible. Detailed demographic data, including adult human immunodeficiency virus (HIV) status, and clinical and sociodemographic data on all diagnosed tuberculosis (TB) patients were available.
RESULTS: The prevalence of *M. tuberculosis* infection was 1.1% using a TST induration cut-off of 15 mm (estimated annual risk of infection of 0.3%). The main identifiable risk factors were maternal HIV infection at birth (adjusted OR [aOR] 3.6, 95%CI 1.1–12.2), having three or more adult members in the household over a lifetime (aOR 2.4, 95%CI 1.2–4.8) and living in close proximity to a known case of infectious TB (aOR 1.6, 95%CI 1.1–2.4), modelled as a linear variable across categories (>200 m, 100–200 m, <100 m, within household). Less than 20% of the infected children lived within 200 m of a known diagnosed case.
CONCLUSION: Household and community risk factors identified do not explain the majority of *M. tuberculosis* infections in children in our setting.
KEY WORDS: *M. tuberculosis* infection; risk factors; children; community; household; HIV

MYCOBACTERIUM TUBERCULOSIS infection in children aged <5 years indicates recent transmission and acts as a sentinel for infectious (typically adult) tuberculosis (TB).[1–4] Accurately identifying *M. tuberculosis* infection in children can highlight recent failures in control measures in the community,[1] and identify a population at high risk of progression to active disease.[5] Untreated *M. tuberculosis* infection in children will ultimately form the reservoir from which future infectious cases will arise.[2,6] Understanding and managing this reservoir is critical to achieving the ambitious Stop TB Partnership target of TB elimination by 2050.[7]

Most *M. tuberculosis* transmission to young children is thought to result from household contact.[1] While household contacts of TB cases are at high risk for infection and disease,[8] in endemic areas most adult TB appears to occur from transmission outside the household.[9–11] This may also be true for paediatric TB, but there is little evidence to date,[12] and results are conflicting. Wide variations in estimates of the proportion of *M. tuberculosis* infection in children attributable to contact with infectious adults within the household have been reported, ranging from 75% in Cape Town[13] to <1% in a peri-urban shanty town in Peru.[14] Differences in settings, such as community human immunodeficiency virus (HIV) and TB prevalence, population density and housing, and differences in age groups, with varying degrees of susceptibility to infection[15] and disparate social contact patterns,[16] contribute to differences in the estimates reported.

Within the context of a well-implemented TB control programme, we aimed to assess household and community risk factors associated with *M. tuberculosis* infection in pre-school children to help elucidate factors driving *M. tuberculosis* transmission that are not being addressed by current prevention strategies.

STUDY POPULATION AND METHODS

Study setting

Karonga District, northern Malawi, is predominantly rural, with an adult HIV prevalence of around 9%,

Correspondence to: Palwasha Y Khan, Department of Infectious Disease Epidemiology, London School of Hygiene & Tropical Medicine, Keppel Street, London WD1E 7HT, UK. e-mail: palwasha.khan@lshtm.ac.uk
Article submitted 10 August 2015. Final version accepted 5 October 2015.

[A version in French of this article is available from the Editorial Office in Paris and from the Union website www.theunion.org]

and an incidence of new smear-positive TB of 87 per 100 000 adults per year.[17] Bacille Calmette-Guérin (BCG) vaccination is administered to children on first health system contact (usually birth) as part of the Expanded Programme on Immunisation.

The present study was conducted in the Karonga demographic surveillance site (DSS), with a population of approximately 36 000, which has been described in detail elsewhere.[18] The entire population is seen annually for re-census and related surveys, and 97% of all births and 99% of all deaths are reported by key informants.[18] Key informants are responsible community individuals who are trained to report vital events at a monthly meeting at which they are given a nominal sum (of about US$3) as compensation.[19] The demographic data include detailed information on family relationships, household socio-economic status, global positioning system (GPS) co-ordinates, dwelling structure, screening for chronic cough ($\geqslant$2 weeks among those seen) and vaccination history of children aged $\leqslant$5 years. Adult HIV status (age $\geqslant$15 years) was collected in HIV prevalence surveys from 2007 to 2011 and other studies.

Participants

We conducted a population-wide tuberculin skin test (TST) survey in 2012 among all children aged 2–4 years residing within the DSS at the time of household recruitment. Children whose mothers were known to be HIV-positive at the time of delivery were defined as HIV-exposed, while children whose mothers were HIV-negative after or up to 1 year before their birth were defined as non-HIV-exposed.

A composite dwelling index was generated based on the building materials used and the presence or absence of glass windows.[20] A composite asset index was constructed by summing the ownership of household items, weighted by the monetary value of each item, to create an asset score.[21] The lowest 10–15% scores were coded as '1', the highest 15% as '4' and the middle groups were divided into '2' and '3' to create the respective socio-economic indices.

Diagnosed tuberculosis cases

The Karonga Prevention Study (KPS) collaborates with the district National Tuberculosis Programme to support core TB prevention and care activities. Bacteriological (including smear and culture status), demographic (including GPS co-ordinates of TB case household/s) and clinical (including HIV status) data were available from an ongoing prospective cohort of all patients starting treatment for TB in the district.[17] A total of 108 adult (aged $\geqslant$15 years) residents of the DSS were diagnosed with smear-positive pulmonary TB during 2007–2012.

The average annual smear-positive notification rate for each predefined residential area of approximately 450 households was calculated as the number of smear-positive TB cases reported per year per 100 000 population for the period 2007–2012.[22] Distance to the nearest smear-positive pulmonary TB case was calculated using ArcGIS 10® software (Environmental Systems Research Institute, Redlands, CA, USA). A categorical variable was generated taking into account the variation in distance to the closest neighbour within the DSS. Preparatory work with parents identified that children aged $<$5 years did not generally venture further than 'calling' distance from home, which was estimated to be about 200 m. The variable was therefore categorised to include household contact, neighbourhood contact (resident within 100 m; resident within 100–200 m) and children who lived $>$200 m from a known smear-positive TB case (baseline category).

Study procedures

Field staff were trained in TST placement and reading according to standard international guidelines,[23] using 2 international units of RT23 (Statens Serum Institute, Copenhagen, Denmark) and measuring the induration 48–72 h later. Periodic blinded comparison readings were done against the same reference reader to ensure consistency of TST reading.

Children with TST $\geqslant$10 mm were assessed for TB-related symptoms by field staff, and the results were recorded in the child's health passport. Any child with symptoms suggestive of TB (fever, weight loss, failure to thrive, night sweats or cough) was reviewed by a clinician and referred to the district hospital. The HIV status of the children was not determined unless clinically indicated at the hospital. All children with TST $\geqslant$15 mm were commenced on 6-month isoniazid preventive therapy (IPT) (10 mg/kg once daily) after active disease had been excluded.

Case definition

A positive TST was defined as an induration of $\geqslant$15 mm, based on a previous mixture analysis of a tuberculin survey conducted in Karonga in 1980–1984 to estimate infection prevalence where exposure to environmental mycobacteria is common.[24] To check the suitability of this cut-off, the analysis was repeated in the new data set.

Statistical analysis

Mixture analysis of tuberculin data was based on implementation of the Expectation Maximization (EM) algorithm in *R* (R Foundation for Statistical Computing, Vienna, Austria) by fitting a two-component model to the observed profiles of the non-zero indurations,[24,25]

Annual risk of infection was calculated using the formula:[2]

$$R \approx (1 - P)^{1/a}$$

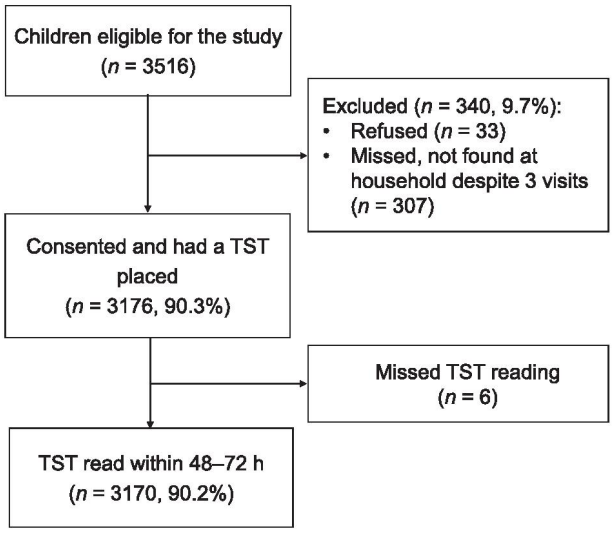


Figure 1 Flow diagram showing study participant flow at each stage from eligibility through to analysis, including non-participation. TST = tuberculin skin test.

where R = annual risk of infection, P = proportion of children with 'positive' TST and a = the mean age of the children.

A random effects logistic regression model taking into account clustering within the residential area was used to assess the relationship between risk factors and TST positivity. The likelihood ratio test was used to assess the overall significance of risk factors, tests for trend and departures from linearity, unless otherwise specified. To prevent over-parameterisation,[26] only those risk factors most strongly associated with the outcome were tested in the multivariable analysis ($P < 0.1$) and limited to a maximum of 4–5 parameters in the fully adjusted model. The population attributable fraction (PAF) was calculated using the formula:[27,28]

$$\text{PAF} = P'\left(\frac{\theta - 1}{\theta}\right)$$

where P' = proportion of cases exposed and θ = the adjusted odds ratio (OR) of the association of *M. tuberculosis* infection in children with the risk factor.

Analyses were performed using Stata v. 13.1 for Mac (Stata Corporation, College Station, TX, USA).

Ethics approval

The study was approved by the Malawi National Health Sciences Research Committee, Lilongwe, Malawi, and the London School of Hygiene & Tropical Medicine Ethics Committee, London, UK. Written informed consent was provided by a parent or guardian of each participating child.

RESULTS

A total of 3516 children aged 2–4 years were resident in the DSS and eligible to take part, 3170 (90.2%) of whom underwent a TST (read within 72 h) (Figure 1). Twenty-six children (0.8%) had lived in the same household as an adult with diagnosed infectious TB, and altogether 606 children (19.1%) had lived within 200 m of an adult with diagnosed infectious TB during their lifetime; 87 children (2.7%) were born to HIV-positive mothers, and altogether 470 (14.8%) children had lived in a household with at least one HIV-positive household member.

Tuberculin data

An inter-rater reliability analysis on a sample of 215 TSTs of reading induration size to within 1 mm was found to be excellent ($\kappa = 0.85$, 95% conference interval [CI] 0.78–0.92), although this does not preclude digit bias; 91.6% of children aged 2–4 years, the majority ($n = 2340$, 81%) of whom had been BCG-vaccinated within the first year of life, had no induration in response to TST. The frequency distribution of non-zero indurations, shown in Figure 2, shows a bimodal distribution with modes at 8 mm and 15 mm, and some evidence of digit preference (at 5 mm, 8 mm and 15 mm).

Mixture analysis of non-zero induration data in this study (Figure 3A) generated a two-component

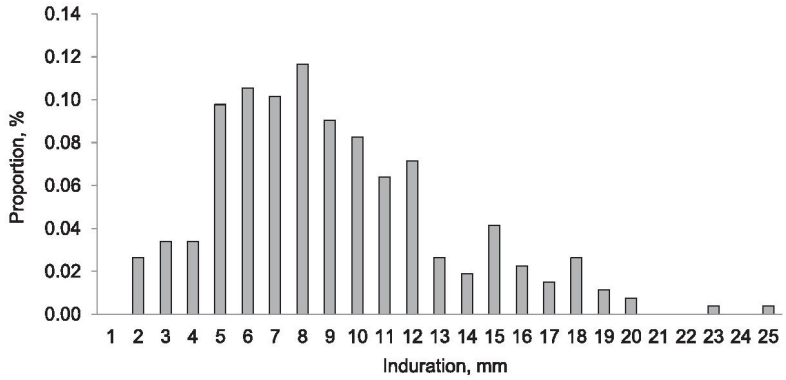


Figure 2 Observed frequency distribution of non-zero TST induration in children aged 2–4 years ($n = 266$). TST = tuberculin skin test.

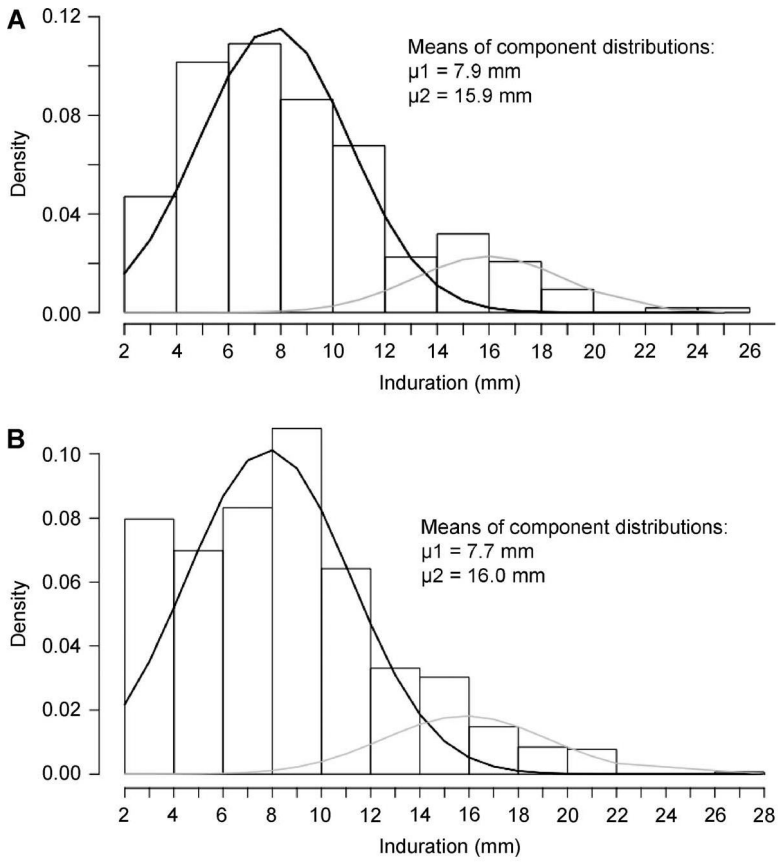


Figure 3 Mixture analysis of tuberculin data **A)** from 2012 in DSS (current study) and **B)** from 1980 to 1984 in Karonga District. DSS = demographic surveillance site.

distribution, with a mean at 7.9 mm representing sensitisation to environmental mycobacteria and/or BCG vaccination, and a mean of 15.9 mm representing *M. tuberculosis* infection. Mixture analysis of non-zero induration data from Karonga District in 1980–1984, restricted to those in the same age group with a BCG scar, revealed a strikingly similar distribution, with means at 7.7 mm and 16.0 mm (Figure 3B).

Of 2782 non-HIV-exposed children, 2555 (91.8%) had zero indurations, compared to 89.7% (78/87) of HIV-exposed children (χ^2, $P = 0.5$). The median and interquartile range (IQR) of the non-zero indurations were similar in HIV-exposed ($n = 9$, median 8 mm, IQR 8–15) and non-HIV-exposed children ($n = 227$, median 8 mm, IQR 6–11).

Prevalence and estimated annual risk of M. tuberculosis *infection*

Of the 3170 children, 35 had a TST induration of $\geqslant$15 mm, giving an estimated *M. tuberculosis* infection prevalence of 1.1% (95%CI 0.8–1.6), adjusted for clustering by residential area. The mean age of children included in the analysis was 3.5 years, giving an estimated average annual risk of *M. tuberculosis* infection (ARTI) of 0.3% (95%CI 0.2–0.5).

Risk factors associated with prevalent M. tuberculosis *infection*

Distance from the nearest known TB case during the child's lifetime, being resident in an area with an average smear-positive notification rate of >30/100 000 population/year (community *M. tuberculosis* exposure), being born to an HIV-positive mother, cohabiting with $\geqslant$3 adult (age $\geqslant$15 years) household members over the child's lifetime, dwelling score and age at BCG vaccination were associated with a positive TST on univariable analysis (Table 1). Adjusting for age and sex did not change the effect estimate for any of the risk factors examined. There was evidence of a dose-response effect of decreasing distance from the nearest known infectious TB case (P trend across categories = 0.02), with the highest ORs in those living within the same household, compared to children who lived >200 m from the nearest infectious case.

In a multivariable model, distance from the nearest known infectious TB case, cumulative number of adult household members and maternal HIV status at birth were the only risk factors that remained strongly associated with the risk of a positive TST ($P < 0.05$; Table 2). The adjusted ORs for these three risk factors were similar to the crude ORs.

The odds of *M. tuberculosis* infection was two-fold

Table 1 Risk factors associated with a positive TST (⩾15 mm) ($n = 3170$)

Risk factor	TST ⩾15 mm *n/N* (row %)	Univariable* OR (95%CI)	*P* value	aOR (95%CI)†	*P* value
Child-related					
Age, years					
2.0–2.9	8/1053 (0.8)	1	0.31	—	—
3.0–3.9	12/1088 (1.1)	1.5 (0.6–3.6)	(test for		
4.0–4.9	15/1029 (1.5)	1.9 (0.8–4.6)	trend 0.13)		
Sex					
Female	16/1579 (1.0)	1	0.63	—	—
Male	19/1591 (1.2)	1.2 (0.6–2.3)			
Time since BCG, months					
<23	2/77 (2.6)	1	0.44	1	0.13
24–35	7/902 (0.8)	0.3 (0.1–1.5)		0.3 (0.06–1.4)	
36–47	9/895 (1.0)	0.4 (0.1–1.8)		0.2 (0.02–2.1)	
⩾48	10/714 (1.4)	0.6 (0.1–2.5)		0.2 (0.01–5.2)	
Missing	7/582 (1.2)				
Age at BCG, days					
<7	16/1005 (1.6)	1	0.05	1	0.05
⩾7	12/1586 (0.8)	0.5 (0.2–1.0)		0.5 (0.2–1.0)	
Missing	7/582 (1.2)				
Household-related					
Dwelling score					
(Worst) 1	5/629 (0.5)	1	0.05	1	0.05
2	15/1037 (1.5)	3.2 (1.0–11.3)		3.2 (0.9–11.1)	
3	4/809 (0.6)	1.3 (0.3–5.5)		1.3 (0.3–5.4)	
(Best) 4	9/560 (1.6)	3.4 (0.9–12.7)		3.4 (0.9–12.6)	
Missing	2/135 (1.5)				
Household asset score					
(Lowest) 1	7/706 (1.0)	1	0.40	1	0.41
2	5/712 (0.7)	0.7 (0.2–2.2)		0.7 (0.2–2.2)	
3	14/892 (1.6)	1.6 (0.6–4.0)		1.6 (0.6–3.9)	
(Highest) 4	7/715 (1.0)	1.0 (0.3–2.8)		1.0 (0.3–2.8)	
Missing	2/145 (1.4)				
Total cumulative number of adult household members during child's lifetime					
⩽3	23/2613 (0.9)	1	0.01	1	0.02
>3	12/557 (2.2)	2.5 (1.2–5.0)		2.4 (1.2–4.9)	
Resident within 50 m of tarmac or main dirt road					
No	29/2847 (1.0)	1	0.21	1	0.21
Yes	6/323 (1.9)	1.8 (0.8–4.5)		1.9 (0.8–4.5)	
Population density in residential area, person/km^2					
<250	12/1027 (1.2)	1	0.79	1	0.81
250–1000	17/1452 (1.2)	1.0 (0.5–2.1)		1.0 (0.5–2.1)	
⩾1000	6/691 (0.9)	0.7 (0.3–2.0)		0.8 (0.3–2.0)	
TB-related					
Known household infectious TB contact during child's lifetime					
No	34/3150 (1.0)	1	0.22	1	0.21
Yes	1/20 (5.0)	4.8 (0.6–37.1)		5.0 (0.7–38.5)	
Distance from nearest known infectious TB case during child's lifetime					
>200 m	23/2564 (0.9)	1	0.14	1	0.16
100–200 m	6/359 (1.7)	1.9 (0.8– 4.6)	(test for	1.9 (0.8– 4.6)	(test for
<100 m	5/227 (2.2)	2.5 (0.9–6.6)	trend 0.02)‡	2.5 (0.9–6.5)	trend 0.02)‡
In same household	1/20 (5.0)	4.4 (0.8–45.3)		4.6 (0.8–46.8)	
		1.6 (1.1–2.4)‡		1.7 (1.1–2.5)‡	
Community *M. tuberculosis* exposure (average smear-positive TB notification rate/100 000/residential area)§					
⩽30	10/1476 (0.7)	1	0.03	1	0.03
>30	25/1694 (1.5)	2.2 (1.1–4.6)		2.2 (1.1–4.6)	
HIV-related					
Maternal HIV status at birth					
Negative	25/2782 (0.9)	1	0.03	1	0.03
Positive	3/87 (3.5)	3.9 (1.2–13.3)		4.1 (1.2–13.9)	
Unknown	7/301 (2.3)	2.6 (1.1–6.1)		2.6 (1.1–6.1)	

Table 1 (continued)

Risk factor	TST ⩾15 mm *n/N* (row %)	Univariable* OR (95%CI)	*P* value	aOR (95%CI)†	*P* value
Total number of known HIV-positive household members during child's lifetime					
0	27/2700 (1.0)	1	0.26	1	0.33
1	4/319 (1.3)	1.3 (0.4–3.6)	(test for	1.2 (0.4–3.5)	(test for
2	4/151 (2.7)	3.0 (0.9–7.8)	trend 0.12)	2.5 (0.9–7.3)	trend 0.14)

* Adjusted for clustering by residential area.
† Adjusted for age and sex.
‡ Assuming linear trend across categories (coded 1 = >200 m from nearest TB case, 2 = 100–200 m, 3 = <100 m and 4 = within household).
§ 21 geographically-defined residential areas.
TST = tuberculin skin test; OR = odds ratio; CI = confidence interval; aOR = adjusted OR; BCG = bacille Calmette-Guérin; TB = tuberculosis; HIV = human immunodeficiency virus.

higher in children living within 200 m of an infectious adult TB case (excluding household contact) compared to children living >200 m from an infectious adult TB case (aOR 2.1, 95%CI 1.0–4.5, $P = 0.05$, adjusted for HIV exposure status and number of adult household members). The odds of *M. tuberculosis* infection in children with known household contact was 3.6 times higher than among children with no known household contact (aOR 3.6, 95%CI 0.5–28.6, $P = 0.3$, adjusted for HIV exposure status and number of adult household members). The proportion of *M. tuberculosis* infection attributable to non-household contact with diagnosed infectious TB within 200 m (the PAF) was estimated at 17.0%, and the PAF for household contact was 2.3%.

DISCUSSION

Despite the known lack of specificity of TST, tuberculin surveys remain a valuable epidemiological tool and provide important information on trends in tuberculous infection.[29] These surveys are usually undertaken in school-aged children (6–11 years), and there is thus a paucity of data on the risk of recent *M. tuberculosis* infection as inferred by TST positivity in the very young.[30,31] In our study, the prevalence of *M. tuberculosis* infection in children aged 2–4 years was approximately 1.1%, using a TST cut-off of 15 mm, yielding an estimated ARTI of 0.3% (95%CI 0.2–0.5).

Table 2 Multivariable analysis of risk factors associated with TST ⩾15 mm ($n = 3170$)

Risk factor	Multivariable model*† OR (95%CI)	*P* value
Total cumulative number of adult household members during child's lifetime		
⩽3	1	0.02
>3	2.4 (1.2–4.8)	
Distance from known TB case during child's lifetime	1.6 (1.1–2.4)‡	0.03
Maternal HIV status at birth		
Negative	1	0.05
Positive	3.6 (1.1–12.2)	
Unknown	2.4 (1.0–5.6)	

* Adjusted for clustering by residential area.
† Adjusted for all risk factors in the model.
‡ Assuming linear trend across categories (coded 1 = >200 m from nearest TB case, 2 = 100–200 m, 3 = <100 m and 4 = within household).
TST = tuberculin skin test; OR = odds ratio; CI = confidence interval; TB = tuberculosis; HIV = human immunodeficiency virus.

Using the tuberculin data from Karonga District in 1980–1984, and restricting this study to a comparable population (aged 2–4 years with BCG scar) and the same methods, the estimated prevalence of *M. tuberculosis* infection was 2.3% (95%CI 1.8–2.8) and the ARTI was 0.7% (95%CI 0.5–0.8). Replication of surveys, with identical antigens and procedures, in the same district and segments of the population (as attempted in this study) is probably the most accurate way to assess changes in the risk of *M. tuberculosis* infection.[32] Mixture analysis of tuberculin data from this study and from the earlier survey demonstrated strikingly similar two-component distributions, supporting the use of the 15 mm cut-off. The results suggest a reduction in the risk of infection in this age group over time, which is consistent with observed trends in incidence of smear-positive adult TB in the district.[17] Our ARTI estimate of 0.3% may not reflect the ARTI in older children and adults in this setting, due to different social mixing patterns,[33,34] although the determinants of the ARTI in this study population of very young children may in fact reflect the social mixing patterns and transmission dynamics of the mother or main carer.[35]

The association with neighbourhood TB was expected. A study in a township in Cape Town, South Africa, found a similar doubling in the odds of TST positivity (using TST ⩾10 mm) in children who lived on the same plot as an adult with smear-positive TB, compared to children with no plot-related exposure.[3] In a setting with a low-to-moderate TB burden such as Karonga, the relative importance of household transmission might be expected to be greater than in Cape Town, where the very high TB incidence might make community transmission more important. It should be noted that despite the association with distance to the nearest infectious

TB case, the PAF in our setting was very low, at 2.3% for household contact and 17.0% for neighbourhood (within 200 m), excluding household contact. This suggests that the maximum public health impact of treating all neighbourhood child contacts (aged <5 years) of known adult infectious TB cases with IPT in this population would not be substantial.

Children born to HIV-positive mothers had a 3.6-fold increased odds of *M. tuberculosis* infection, after adjusting for distance to known TB cases and the number of adult household members. An increased risk of *M. tuberculosis* infection in HIV-exposed, non-infected children compared to non-HIV-exposed children in early childhood was also seen in a Ugandan study, which used both TST and an interferon-gamma release assay to diagnose *M. tuberculosis* infection status.[36] We did not know the HIV status of the children or the maternal *M. tuberculosis* infection status. It is unclear in our study whether the increased risk of *M. tuberculosis* infection seen in HIV-exposed children is due to increased susceptibility to *M. tuberculosis* infection following exposure and/or increased exposure, via either occult infectious TB within the household or unmeasured lifestyle factors associated with maternal HIV, such as increased attendance at health facilities.

Interestingly, no association between *M. tuberculosis* infection and population density was found. A plausible explanation may be that *M. tuberculosis* infections in young children in this community occur as a result of social mixing, which is unrelated to the population density of the area in which the child resides. There may be more undiagnosed infectious TB cases among females (the usual care givers of young children in rural Malawi), particularly in areas with lower population density, which are further from primary health care facilities than areas with the highest population density. A study is currently ongoing in the DSS to see if it is possible to identify the source of incident *M. tuberculosis* infection in children. Incident *M. tuberculosis* infection is identified by a repeat TST 1 year later in children with a previously negative TST. Household, neighbourhood and family contacts of children identified with incident *M. tuberculosis* infection are then screened for TB. It is hoped that this type of targeted case finding may identify a subset of undiagnosed cases that are responsible for onward *M. tuberculosis* transmission to the youngest members within the community.

The 1% of young *M. tuberculosis*-infected children in this population act as a sentinel for infectious TB in the community. In this setting, the majority of transmission events take place outside the household, either from unidentifiable casual contacts (with known or unknown TB) or from undiagnosed close contacts. This has major implications for strategies aiming to interrupt transmission. This study highlights the need to identify people with undiagnosed infectious TB in the community, the need to identify congregate transmission settings and the need to develop strategies to improve community infection control, irrespective of the presence of people with known TB. IPT for household contacts of smear-positive TB cases, although justified given the high risk to individual children, is predicted to have little effect on the population reservoir of infection, even with effective implementation in settings such as this.

CONCLUSION

The majority of *M. tuberculosis* infections in children in our setting, which has a well-implemented TB programme, occur either from casual contact with infectious TB in the community or from undiagnosed infectious TB in close contacts. A better understanding of the loci of *M. tuberculosis* transmission to young children is needed to effectively target prevention interventions to mitigate future TB-related morbidity and mortality in children and ultimately improve longer-term TB control.

Acknowledgements

This work was supported by a Wellcome Trust (London, UK) Strategic Award for the Karonga Prevention Study [grant number 098610/Z/12/Z] and a Wellcome Trust clinical research training fellowship [PYK: grant number 100137/Z/12/Z].

Some of the preliminary findings of this study have been presented in part at the: Conference on Retroviruses and Opportunistic Infections (CROI) 2015, 23–26 February 2015, Seattle, WA, USA. [Abstract number 885], and the Tuberculosis Surveillance Research Unit (TSRU) meeting at the World Health Organization, 15–17 April 2015, Geneva, Switzerland. [Oral presentation]

Conflicts of interest: none declared.

References

1 Bloch A B, Snider D E, Jr. How much tuberculosis in children must we accept? Am J Public Health 1986; 76: 14–15.
2 Rieder H L. Epidemiological basis of tuberculosis control. Paris, France: International Union Against Tuberculosis and Lung Disease, 1999: p 162.
3 Middelkoop K, Bekker L G, Morrow C, Zwane E, Wood R. Childhood tuberculosis infection and disease: a spatial and temporal transmission analysis in a South African township. S Afr Med J 2009; 99: 738–743.
4 Marais B J, Obihara C C, Warren R M, Schaaf H S, Gie R P, Donald P R. The burden of childhood tuberculosis: a public health perspective. Int J Tuberc Lung Dis 2005; 9: 1305–1313.
5 Marais B J, Gie R P, Schaaf H S, Beyers N, Donald P R, Starke J R. Childhood pulmonary tuberculosis: old wisdom and new challenges. Am J Respir Crit Care Med 2006; 173: 1078–1090.
6 Diel R, Loddenkemper R, Zellweger J P, et al. Old ideas to innovate TB control: preventive treatment to achieve elimination. Eur Respir J 2013 2013; 42: 785–801.
7 Esmail H, Barry C E, 3rd, Young D B, Wilkinson R J. The ongoing challenge of latent tuberculosis. Philos Trans R Soc Lond B Biol Sci 2014; 369: 20130437.

8 Morrison J, Pai M, Hopewell P C. Tuberculosis and latent tuberculosis infection in close contacts of people with pulmonary tuberculosis in low-income and middle-income countries: a systematic review and meta-analysis. Lancet Infect Dis 2008; 8: 359–368.
9 Verver S, Warren R M, Munch Z, et al. Proportion of tuberculosis transmission that takes place in households in a high-incidence area. Lancet 2004; 363: 212–214.
10 Crampin A C, Glynn J R, Traore H, et al. Tuberculosis transmission attributable to close contacts and HIV status, Malawi. Emerg Infect Dis 2006; 12: 729–735.
11 Buu T N, van Soolingen D, Huyen M N, et al. Tuberculosis acquired outside of households, rural Vietnam. Emerg Infect Dis 2010; 16: 1466–1468.
12 Schaaf H S, Michaelis I A, Richardson M, et al. Adult-to-child transmission of tuberculosis: household or community contact? Int J Tuberc Lung Dis 2003; 7: 426–431.
13 Wood R, Johnstone-Robertson S, Uys P, et al. Tuberculosis transmission to young children in a South African community: modeling household and community infection risks. Clin Infect Dis 2010; 51: 401–408.
14 Madico G, Gilman R H, Checkley W, et al. Community infection ratio as an indicator for tuberculosis control. Lancet 1995; 345: 416–419.
15 Zelner J L, Murray M B, Becerra M C, et al. Age-specific risks of tuberculosis infection from household and community exposures and opportunities for interventions in a high-burden setting. Am J Epidemiol 2014; 180: 853–861.
16 Middelkoop K, Bekker L G, Morrow C, Lee N, Wood R. Decreasing household contribution to TB transmission with age: a retrospective geographic analysis of young people in a South African township. BMC Infect Dis 2014; 14: 221.
17 Mboma S M, Houben R M, Glynn J R, et al. Control of (multi)drug resistance and tuberculosis incidence over 23 years in the context of a well-supported tuberculosis programme in rural Malawi. PLOS ONE 2013; 8: e58192.
18 Jahn A, Crampin A, Glynn J, Mwinuka V, et al. Evaluation of a village-informant driven demographic surveillance system. Demographic Res 2007; 16: 219–248.
19 Crampin A C, Dube A, Mboma S, et al. Profile: the Karonga Health and Demographic Surveillance System. Int J Epidemiol 2012; 41: 676–685.
20 Boccia D, Hargreaves J, Ayles H, Fielding K, Simwinga M, Godfrey-Faussett P. Tuberculosis infection in Zambia: the association with relative wealth. Am J Trop Med Hyg 2009; 80: 1004–1011.
21 Odone A, Crampin A C, Mwinuka V, et al. Association between socioeconomic position and tuberculosis in a large population-based study in rural Malawi. PLOS ONE 2013; 8: e77740.
22 Dye C, Scheele S, Dolin P, Pathania V, Raviglione M C. Consensus statement. Global burden of tuberculosis: estimated incidence, prevalence, and mortality by country. WHO Global Surveillance and Monitoring Project. JAMA 1999; 282: 677–686.
23 Arnadottir T, Rieder H L, Trébucq A, Waaler H T. Guidelines for conducting tuberculin skin test surveys in high prevalence countries. Tubercle Lung Dis 1996; 77 (Suppl 1): 1–19.
24 Davies G R, Fine P E, Vynnycky E. Mixture analysis of tuberculin survey data from northern Malawi and critique of the method. Int J Tuberc Lung Dis 2006; 10: 1023–1029.
25 Benaglia T, Chauveau D, Hunter D R, Young D. mixtools: an R package for analyzing finite mixture models. J Stat Software 2009; 32: 1–29.
26 Peduzzi P, Concato J, Kemper E, Holford T R, Feinstein A R. A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol 1996; 49: 1373–1379.
27 Greenland S, Drescher K. Maximum likelihood estimation of the attributable fraction from logistic models. Biometrics 1993; 49: 865–872.
28 Rockhill B, Newman B, Weinberg C. Use and misuse of population attributable fractions. Am J Public Health 1998; 88: 15–19.
29 Borgdorff M. Annual risk of infection—time for an update? Bull World Health Organ 2002; 80: 501–502.
30 Joos T J, Miller W C, Murdoch D M. Tuberculin reactivity in bacille Calmette-Guerin vaccinated populations: a compilation of international data. Int J Tuberc Lung Dis 2006; 10: 883–891.
31 Reid J K, Ward H, Marciniuk D, Hudson S, Smith P, Hoeppner V. The effect of neonatal bacille Calmette-Guerin vaccination on purified protein derivative skin test results in Canadian aboriginal children. Chest 2007; 131: 1806–1810.
32 Fine P E, Bruce J, Ponnighaus J M, Nkhosa P, Harawa A, Vynnycky E. Tuberculin sensitivity: conversions and reversions in a rural African population. Int J Tuberc Lung Dis 1999; 3: 962–975.
33 Buu T N, Quy H T, Qui N C, Lan N T, Sy D N, Cobelens F G. Decrease in risk of tuberculosis infection despite increase in tuberculosis among young adults in urban Viet Nam. Int J Tuberc Lung Dis 2010; 14: 289–295.
34 Wood R, Liang H, Wu H, et al. Changing prevalence of tuberculosis infection with increasing age in high-burden townships in South Africa. Int J Tuberc Lung Dis 2010; 14: 406–412.
35 Mossong J, Hens N, Jit M, et al. Social contacts and mixing patterns relevant to the spread of infectious diseases. PLOS MED 2008; 5: e74.
36 Marquez C, Chamie G, Achan J, et al., eds. Tuberculosis infection in early childhood in Uganda and the influence of HIV exposure. 21st Conference of Retroviral and Opportunistic Infections, 3–6 March, 2014; Boston, MA, USA. http://www.croiwebcasts.org/console/player/22177?mediaType=slidevideo&

RESUME

CONTEXTE : L'infection à *Mycobacterium tuberculosis* chez l'enfant se comporte comme une sentinelle de la tuberculose (TB) contagieuse. Le but de l'étude a été d'évaluer les facteurs de risque associés à l'infection chez des enfants d'âge préscolaire.
MÉTHODE : Nous avons réalisé une enquête par test cutané à la tuberculine (TST) dans toute la population de janvier à décembre 2012 au Malawi. Tous les enfants âgés de 2 à 4 ans résidant dans une zone de surveillance démographique ont été éligibles. Des données démographiques détaillées, notamment le statut par rapport au virus de l'immunodéficience humaine (VIH) de l'adulte et les données cliniques et sociodémographiques concernant tous les patients ayant eu un diagnostic de TB ont été disponibles.
RÉSULTATS : La prévalence de l'infection à *M. tuberculosis* a été de 1,1% en utilisant comme seuil une induration du TST de 15 mm (risque annuel estimé d'infection de 0,3%). Les principaux facteurs de risque identifiables ont été l'infection à VIH maternelle à la naissance (OR ajusté [ORa] 3,6 ; IC95% 1,1–12,2), la présence de ⩾3 adultes dans le même foyer pendant toute leur vie (ORa 2.4 ; IC95% 1,2–4,8) et le fait de vivre à proximité d'un cas de TB contagieuse connu (ORa 1,6 ; IC95% 1,1–2,4), modélisé comme une variable linéaire à travers les différentes catégories (>200 m, 100–200 m, <100 m, au sein du foyer). Moins de 20% des enfants infectés vivaient dans un rayon de 200 m d'un cas diagnostiqué connu.
CONCLUSION : Les facteurs de risque identifiés au niveau des foyers et des communautés n'expliquent pas la majorité des infections à *M. tuberculosis* des enfants dans notre contexte.

RESUMEN

MARCO DE REFERENCIA: La infección por *Mycobacterium tuberculosis* en los niños cumple una función centinela de la tuberculosis (TB) contagiosa. El presente estudio tuvo por objeto evaluar los factores de riesgo de contraer la infección en los niños de edad prescolar.
MÉTODO: Se llevó a cabo una encuesta tuberculínica (TST) a escala poblacional de enero a diciembre del 2012 en Malawi. Todos los niños de edad de 2–4 años residentes en una zona de vigilancia demográfica eran aptos para participar en la encuesta. Todos los pacientes con diagnóstico de TB contaban con datos demográficos exhaustivos, incluida la situación frente al virus de la inmunodeficiencia humana (VIH), los datos clínicos y sociodemográficos.
RESULTADOS: La prevalencia de infección tuberculosa fue 1,1%, al utilizar un umbral discriminatorio de 15 mm de induración (un riesgo anual de infección de 0,3%). Los principales factores de riesgo de contraer la infección que se detectaron fueron la infección materna por el VIH en el momento del parto (OR ajustado [ORa] 3,6; IC95% 1,1–12,2); haber cohabitado con ⩾3 adultos miembros del hogar durante toda la vida (ORa 2,4; IC95% 1,2–4,8); y el hecho de haber vivido próximo a un caso conocido de TB contagiosa (ORa 1,6; IC95% 1,1–2,4), modelado como una variable lineal en todas las categorías (a >200 m, de 100 m a 200 m, <100 m y en el domicilio). Menos de 20% de los niños infectados había vivido a <200 m de un caso diagnosticado conocido.
CONCLUSIÓN: Los factores de riesgo domiciliarios y comunitarios de contraer la infección tuberculosa que se detectaron no explican la mayor parte de estas infecciones en los niños de este entorno.

- Erratum of the formula to calculate annual risk of *M.tb* infection on page 343 of published paper. *Journal was informed on 4th May 2016.*
  Correct formula:

$$R \approx 1 - (1 - P)^{1/a}$$

  where P= prevalence of *M.tb* infection, a= the mean age of children

- Erratum of PAF of household TB contact on page 347 of published paper: should be 2.1% not 2.3%

5 Incident *Mycobacterium tuberculosis* infection in children under 6 years

5.1 Research paper IV

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT
www.lshtm.ac.uk

Registry
T: +44(0)20 7299 4646
F: +44(0)20 7299 4656
E: registry@lshtm.ac.uk

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

RESEARCH PAPER COVER SHEET

PLEASE NOTE THAT A COVER SHEET MUST BE COMPLETED FOR EACH RESEARCH PAPER INCLUDED IN A THESIS.

SECTION A – Student Details

Student	Palwasha Yousafzai Khan
Principal Supervisor	Professor Judith Glynn
Thesis Title	Investigating Mycobacterium tuberculosis transmission in rural Malawi

If the Research Paper has previously been published please complete Section B, if not please move to Section C

SECTION B – Paper already published

Where was the work published?			
When was the work published?			
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion			
Have you retained the copyright for the work?*	Choose an item.	Was the work subject to academic peer review?	Choose an item.

**If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.*

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	American Journal of Respiratory & Critical Care Medicine
Please list the paper's authors in the intended authorship order:	Khan PY, Glynn JR, Fielding KL, Mzembe T, Mulawa D, Chiumya R, Kranzer K, Fine PEM, Crampin AC
Stage of publication	**Not yet submitted**

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	AC/JG conceived of the idea for the analysis. I led on all aspects of the fieldwork, data collection and analysis. I wrote the first and final drafts of the paper

Student Signature: [redacted] **Date:** 2/8/2017

Supervisor Signature: [redacted] **Date:** 2/8/2017

Incident *Mycobacterium tuberculosis* infections in children identify risk factors for community *M.tb* transmission in Malawi

Authors: Palwasha Y Khan,[1] Judith R Glynn,[1] Katherine L Fielding,[1] Themba Mzembe,[2] Dominic Mulawa,[2] Regina Chiumya,[2] Katharina Kranzer,[3] Paul EM Fine,[1] Amelia C Crampin,[1,2]

Affiliations:

[1] Department of Infectious Disease Epidemiology, London School of Hygiene & Tropical Medicine, London, UK

[2] Malawi Epidemiology and Intervention Research Unit, Lilongwe, Malawi

[3] National and Supranational Mycobacterium Reference Laboratory, Forschungszentrum Borstel, Germany

Corresponding author:

Dr Palwasha Yousafzai Khan

MBBCh (Hons), MRCS, MRCP, DTM&H, MSc

Department of Infectious Disease Epidemiology

London School of Hygiene and Tropical Medicine

Keppel Street, London, WD1E 7HT, United Kingdom

Mobile: +44 (0) 7811 902 455

Fax: +44 (0) 207 636 8739

Email: palwasha.khan@lshtm.ac.uk

Authors' contributions: Conceived of the study: ACC and JRG; fieldwork and data collection: PYK (lead), TM, DM and RC; performed and interpreted data analysis: PYK and KLF; contributed to writing of manuscript: PYK (lead) and JRG; final approval of version submitted for publication: *not yet taken place*

Funding: This work was supported by a Wellcome Trust clinical research training fellowship [PYK: grant number 100137/Z/12/Z] and a Wellcome Trust Strategic Award for Malawi Epidemiology and Intervention Research Unit (formerly known as Karonga Prevention Study) [grant number 098610/Z/12/Z].

Running head: Risk factors for incident *M.tb* infection

Subject of manuscript: 11.2 Epidemiology of Tuberculosis

Word count for body of manuscript: 4410

Word count for abstract: 251

At a Glance Commentary

Scientific knowledge on the Subject: Important gaps exist in our knowledge of Mycobacterium tuberculosis (*M.tb)* transmission, especially in high HIV prevalence settings, and there is significant uncertainty about where most *M.tb* transmission takes place in the community. Young children with incident *M.tb* infection are sentinels of recent community transmission and may highlight new areas for tuberculosis (TB) control programmes to target.

What this Study Adds to the Field: In this population-based longitudinal tuberculin skin test (TST) study in children in rural Malawi, increasing age, having an HIV-positive father, living within a residential area with a higher tuberculosis (TB) notification rate, attendance at church, and travel on mini-buses were all independent risk factors for incident *M.tb* infection, as inferred from TST conversion. Identifying risk factors for incident *M.tb* infection in young children provides insight into drivers of community *M.tb* transmission which are not being addressed by current TB control strategies.

This article has an online data supplement, which is accessible from this issue's table of content online at www.atsjournals.org

Abstract

Rationale: Incident *Mycobacterium tuberculosis (M.tb)* infection in young children is a critical indicator of recent community transmission, highlighting failures in TB control measures

Objectives: To investigate risk factors associated with incident *M.tb* infection in children living in an HIV-prevalent setting.

Methods: Children aged under 6 years of age resident within a demographic surveillance site in Malawi were recruited. Tuberculin skin testing (TST) was performed at baseline and repeated after 1-2 years. Incident *M.tb* infection was defined as TST conversion from <10mm to ≥10mm with an increment of ≥13mm (based on mixture analysis). Multivariate analyses used random-effects Poisson regression.

Measurements and Main Results: Among the 3066 children who underwent serial TST, 91 children TST-converted (3.0%), giving an incidence of 1.6 per 100 person-years (95% CI: 1.3 – 2.0 per 100py). Age (incidence rate ratio (IRR) 1.4: 95% CI 1.1 - 3.0 for each year increase in age), having an HIV-positive father (IRR 2.3: 95% CI 1.2- 4.3), living within a residential area with a smear-positive TB notification rate > 30 per 100,000 (IRR 2.5: 95% CI 1.4 – 4.4), church attendance (IRR 3.4: 95% 0.8 – 14.1) and travel on mini-buses (IRR 1.8: 95% CI 1.1 – 3.0) were all associated with incident *M.tb* infection. Sixty-eight percent of incident infections were attributable to church attendance.

Conclusions: This research highlights the need for better infection control practices in congregate settings such as churches and mini-buses and supports the call for the implementation of TB-HIV interventions at the household level to further reduce the burden of TB.

Introduction

Tuberculosis (TB) remains one of the leading causes of death from an infectious disease globally.[1] A hugely ambitious target of 'ending the global TB epidemic' within the next 20 years, defined as less than 10 new cases per 100,000 population per year, has been set by the World Health Organisation (WHO) as part of the End TB Strategy.[2] Achieving this overarching goal would require sustained reductions in TB incidence of up to 5-10 fold higher than the current average decrease of 1.5% per annum.[3] This required acceleration in the reduction of TB incidence will only be realised if there is increased emphasis on minimising transmission.[4] Unfortunately, we do not yet know how best to break the *M.tb* transmission cycle, as significant gaps remain in our understanding of the dynamics of *M.tb* transmission, especially in high HIV prevalence settings.[5]

Childhood TB usually follows transmission from an infectious adult;[6-8] and because in young children infection must be recent, TB incidence in children provides a measure of the recent performance of TB control programmes,[9] and is potentially an indicator of *Mycobacterium tuberculosis (M.tb)* transmission 'hot spots' in a community.[10] The incidence of *M.tb* infection in children should be a more precise measure of contemporaneous *M.tb* transmission in a community than the incidence of TB disease, as it is more timely and not all *M.tb*-infected children progress to active disease. Knowledge of the distribution of *M.tb* infection in children may assist evaluation of the level of ongoing transmission and help to guide control strategies.[11]

We conducted a population-wide tuberculin skin test (TST) cohort study in young children resident in a demographic surveillance site in Karonga district, Malawi, to estimate the incidence of *M.tb* infection and identify risk factors associated with recent *M.tb* transmission to children in the community.

Methods

Study setting and study population

The study setting and study population in the Karonga demographic surveillance area (DSS) has been described in detail.[12,13] In brief, the Karonga DSS has a population of about 39,000; adult HIV prevalence is around 9% and incidence of new smear-positive TB is approximately 85 per 100,000 population per year.[12,13]

Study design

The initial baseline survey was conducted in January 2012 to April 2013 among all children aged from 3 months to 4 years old resident within the Karonga DSS. The plan was to repeat a TST one year later in all children whose initial TST induration was less than 10mm at baseline. Unfortunately, due to a global shortage of RT23 tuberculin and difficulties in procurement,[14] there was a delay in some children receiving the 2nd TST. Only those children aged less than 6 years of age by the time of the 2nd round of TST (January 2013 to August 2015) were eligible for inclusion in the cohort study. The median time between TSTs was 1.7 years, with an interquartile range of 1.1 to 2.6 years.

Case definition

TST conversion was defined as any child with an increase in induration size ≥ 13mm. An increase of 13mm was based on mixture analysis (described below), rather than the traditional criteria proposed by the American Thoracic Society (ATS) of an increase ≥10mm.[15] Children who had a TST of less than 10mm at baseline were assumed to be uninfected at baseline.

Study procedures

The movement and contacts of each child were recorded through structured guardian interviews prior to the placing of the second TST, to reduce recall bias. Information on history of TB within the family and close contacts within the last year were ascertained using a standardised questionnaire. An adaptation to the questionnaire to record the number of times the child had attended any crowded community gathering places, such as funerals, churches, healthcare facilities, markets and travel on minibuses within the last year was made part-way through the 2nd round of TST. Therefore, data on attendance at gathering places was only available on 2041 children (66.6% of the total cohort).

All children identified with incident *M.tb* infection were evaluated for TB-related symptoms, examined for evidence of BCG scar and scar size, and height, weight and mid upper arm circumference were recorded. Any child with symptoms suggestive of TB (fever, weight loss, failure to thrive, night sweats or cough) was reviewed by a clinician and referred to the district hospital. The HIV status of the children was not determined unless clinically indicated at the hospital. All children with incident *M.tb* infection were commenced on 6-month isoniazid

preventive therapy (IPT) (10 mg/kg once daily) after active disease had been excluded.

Demographic data

Demographic data available include detailed information on family relationships, household socio-economic status (SES), global positioning system (GPS) co-ordinates, dwelling structure, and vaccination history. Adult HIV status (age ≥15 years) was collected in HIV sero-prevalence surveys from 2007 to 2011 and other studies nested within the DSS in 2012 to 2014.

A composite dwelling index was generated based on the building materials used and the presence or absence of glass windows.[16] A composite score for household socioeconomic status was created using head of household employment, number of assets, and availability of soap.[17] The lowest 10– 15% scores were coded as '1', the highest 15% as '4' and the middle groups were divided into '2' and '3' to create the respective socio-economic indices. Food insecurity was defined as a binary variable based on whether there had been a time in the last year when there was not enough food for the household to have its normal meals (described as either fewer meals per day, and/or smaller meals, and/or less variety of foods).

Diagnosed TB cases in DSS

Bacteriological (including smear and culture status), demographic (including GPS co-ordinates of TB case household/s) and clinical (including HIV status) data were available from an ongoing prospective cohort of all patients starting

treatment for TB in the district.[18] A total of 38 adults (aged ≥15 years) resident in the DSS were diagnosed with smear-positive pulmonary TB during 2013–2015. The average annual smear-positive TB notification rate for each predefined residential area of approximately 450 households was calculated as the number of smear-positive TB cases reported per year per 100,000 population for the period 2013–2015. Distance to the nearest smear-positive pulmonary TB case, distance to the nearest clinic and distance to the main tarmac road were calculated using ArcGIS 10® software (Environmental Systems Research Institute, Redlands, CA, USA).

Statistical analysis

Mixture analysis of the tuberculin data was based on implementation of the Expectation Maximisation (EM) algorithm in R (R Foundation for Statistical Computing, Vienna) by fitting a finite mixture model to the observed profiles of the non-zero increment in induration size in mm.[13,19-21] The density of the bimodal distribution of the tuberculin data (see results) is defined as:

$$\int (x|\lambda, \mu_1, \sigma_1, \mu_2, \sigma_2) = \lambda \times D_1(x|\mu_1, \sigma_1) + (1 - \lambda) \times D_2(x|\mu_2, \sigma_2)$$

where D_1 and D_2 are the distributions accounting for the first (i.e. lower increment in induration size: mean and standard deviation, μ_1, σ_1) and second (i.e. the higher increment in induration size: mean and standard deviation, μ_2, σ_2) peaks respectively, and λ and $(1 - \lambda)$ are the weights for the D_1 and D_2 distributions, respectively.[22] The fitted finite mixture model was used to identify a cut-off value that discriminated the two modes of the dataset. This was done by calculating

the probability, *P,* to belong to the lower induration size peak as $P = D_1/(D_1 + D_2)$ to be less than 0.05 (to maximise specificity of the definition of TST conversion). This was then used to derive the cut-off value. The confidence interval of the cut-off value was computed using Monte Carlo simulations.[21,22]

A random effects Poisson regression model, to account for clustering within residential area, was used to assess the relationship between risk factors and TST conversion. The likelihood ratio test was used to assess the overall significance of risk factors, tests for trend and departures from linearity, unless otherwise specified. Only those risk factors most strongly associated with the outcome were tested in the multivariable analysis ($p<0.2$). Sex, household SES and dwelling structure were included as *a priori* confounders in the multivariate model. The baseline category for each of the variables relating to gathering place attendance was chosen *a priori* as those unexposed. The data were collected as categorical variables grouped as 0, 1-3, 4-12, 12-23, ≥24 visits in the last year. Very few children had not attended church in the last year (4%) so this category was combined with the next category (1-3 visits). To prevent over-parameterisation of the multivariate model, risk factors were limited to a maximum of 8 parameters. There was no evidence of a dose response in number of visits and risk of *M.tb* infection for any of the gathering places on univariate analysis and so these data were categorised as binary variables.

Sensitivity analyses were undertaken using (i) the standard ATS criteria of TST conversion ≥10mm; (ii) a more stringent cut-off of ≥15mm to define TST

conversion; and (iii) using TST conversion ≥13mm and excluding children with a TST conversion of between 9 to 12mm from the analysis.

The population attributable fraction (PAF) was calculated using the formula: [23]

$$\text{PAF} = \text{p'}\left(\frac{\text{RR}-1}{\text{RR}}\right)$$

where p' = proportion of cases exposed; RR = adjusted rate ratio of the association of TST conversion in children with the risk factor. Analysis was performed using Stata 13.1 for Mac (Stata Corporation, College Station, TX).

Ethics approval

The study was approved by the Malawi National Health Sciences Research Committee, Lilongwe, Malawi and the London School of Hygiene & Tropical Medicine Ethics Committee, London, UK. Written informed consent was provided by a parent or guardian of each participating child.

Results

A total of 3357 children aged under 6 years were eligible for inclusion in the cohort, 3066 (91.3%) of whom underwent a 2nd TST which was read within 72 hours. The study flowchart is shown in Figure 1.

Tuberculin data

The frequency distribution of all TST conversions (i.e. increments between TSTs > 0 mm, n=350) is shown in Figure 2a. Of those children, who had zero induration

at baseline, 90 per cent (2477/2745) did not exhibit any induration at the 2nd TST. Of those children who had some evidence of tuberculin sensitivity (induration between 2mm and 9mm) at baseline, 73 per cent (235/321) had a decrease in the size of induration at the 2nd TST; most of whom reverted to zero induration (220/235).

Mixture analysis of all TST conversions (range 2 to 26mm) generated a two-component distribution, as shown in Figure 2b. The mean (and standard deviation) were 6.5 (2.3), and 14.7 (3.1) for the 1st and 2nd peak respectively. The derived cut-off point using the finite mixture model was 12.8 (95% CI 12.6 – 13.0) as shown in Figure 2b. A TST conversion of 13mm was chosen to define incident *M.tb* infection to minimise false positives.

Of the 2588 children with no known HIV-exposure *in utero*, 2128 (82.2%) exhibited no change in induration size between TSTs, compared to 77.8% (238/306) of HIV-exposed children (chi-squared, P=0.32). The median and interquartile range (IQR) of the increment in induration size between TSTs were similar in HIV-exposed (n=42, median 7 mm, IQR 6–13) and non-HIV-exposed children (n=292, median 8 mm, IQR 6–13).

Incidence of M.tb infection

Of the 3066 children who underwent serial skin testing, 91 TST-converted (3.0%). The 3066 children contributed 5580 person-years (py), giving an *M.tb* infection incidence of 1.6 per 100py (95% CI: 1.3 – 2.0 per 100py). Using the ATS criteria of TST conversion ≥10mm, *M.tb* infection incidence was 2.5 per 100py

(95%CI 2.1 – 2.9 per 100py) and using the more stringent TST conversion ≥15mm, infection incidence was 1.1 per 100py (95% CI: 0.8 – 1.4).

Risk factors associated with incident M.tb infection

The incidence of TST-conversion (incident *M.tb* infection) was higher at older ages, among those without documented BCG vaccination, those with known household contact with an infectious TB case and those resident in areas with a higher community *M.tb* exposure. The incidence was lower in those living in the worst dwellings but there was no association with other measures of SES. The incidence was higher in those with an HIV positive father but with only a weak association with HIV-positive mothers or HIV prevalence in the area (Table 1).

Data on attendance at gathering places within the last year were only available on 2041 children. A table showing the characteristics of children in whom data on attendance at gathering places were collected compared to children in whom these data were not available is included in the online supplement (Table E1). Although a higher percentage of children had mothers who were HIV-positive and a higher percentage of children lived in areas with the higher community *M.tb* exposure (based on average notification rate of smear-positive pulmonary TB), *M.tb* incidence rate in these children was 1.5 per 100py (95% CI: 1.2 – 1.90), which was similar to the incidence rate seen in the total study population. In those with data available, there was weak evidence for an association with travel on a mini-bus and school attendance ($p<0.1$) and even weaker evidence ($p<0.2$) for an association with church attendance and going to market on univariate analysis (Table 2).

In the multivariable model, age at the time of the 1st skin test, known HIV-positive father, higher community *M.tb* exposure, attendance at church and travel on mini-buses were the only risk factors which were associated with incident *M.tb* infection ($p<0.05$; Table 3). A one year increase in age at 1st TST was associated with a 1.4 increase in the incidence rate of *M.tb* infection (aOR 1.4, 95% CI 1.1 – 3.0; p=0.01). Attendance at church was the strongest risk factor for *M.tb* infection, with children who had attended church at least 4 times in the last year, having a 3.4-fold increase in the rate of incident infection compared to children who had attended church 0-3 times in the last year, although the confidence interval was wide (aOR 3.4, 95% CI 0.8 – 14.1; p=0.04). Having an HIV-positive father was associated with a 2.3 times increase in incident *M.tb* infection (aOR 2.3, 95% CI 1.2 – 4.2; p=0.02) and community *M.tb* exposure was associated with a 2.5-fold increase in incidence rate (aOR 2.5, 95% CI 1.4 – 4.4, p=0.002). Children who had travelled on a mini-bus within the last year had a 1.8 times higher incidence rate than children who had not been on a mini-bus (aOR 1.8, 95% CI: 1.1 – 3.0, p=0.03).

The proportion of incident *M.tb* infection attributable to having an HIV-positive father, being resident in an area with higher community *M.tb* exposure, church attendance and travel on mini-bus was 6.9%, 28.7%, 67.8% and 27.7%, respectively (Table 3). The confidence interval around the population attributable fraction (PAF) for church attendance includes 0, reflecting the wide confidence interval of the effect estimate in the multivariable model.

In the sensitivity analyses, using the ATS criteria for TST conversion (TST increase ≥10mm) as the definition of incident *M.tb* infection reduced the strength of the association for all risk factors in the multivariable model, whilst using the more stringent definition (TST increase≥15mm) reduced the strength of the association with church attendance (aOR 2.5: 95% CI 0.6 – 10.4) and increased the strength of the association with having a known HIV-positive father (aOR 3.0: 95% CI 1.2 – 7.6) but made little difference to the other risk factors. Excluding children with TST increases between 9 and 12 mm made little difference to the original multivariable model using TST increase≥13mm (Table 3).

Discussion

Despite a well-functioning TB control programme in Karonga,[18] the risk of *M.tb* infection in children is estimated to be 1.6% (95% CI 1.3 – 2.0%) per annum in our study. Using children with incident *M.tb* infection as 'sentinels' of recent transmission, we have identified a number of factors which are potentially driving community *M.tb* transmission. Exposure to crowding in poorly ventilated indoor congregate settings, such as churches and mini-buses, having a HIV-positive father, and community *M.tb* exposure as estimated from TB notification data are all associated with an increased risk of incident *M.tb* infection in young children.

Travel on minibus was associated with a nearly 2-fold increase in the rate of incident *M.tb* transmission in young children, and church attendance was associated with a 3-fold increase in the rate of incident infection. Although the

confidence interval for church attendance in the multivariable model was wide, the direction of the effect was consistent in each of the sensitivity analyses and the p value from the likelihood ratio test for inclusion of this risk factor in each of the models was consistently less than 0.05. A higher level of community *M.tb* exposure was associated with a 2-fold increase in the rate of incident infection in children. However, only 10/91 (11%) of TST conversions occurred in children known to be resident within 200 metres of a diagnosed smear-positive TB case. This suggests that the majority of *M.tb* transmission to young children is occurring as a result of casual contact with undiagnosed infectious TB in the community rather than neighbourhood contact with known smear-positive TB cases. There has been an accumulation of evidence over the years, that in high burden settings, despite some evidence for clustering of *M.tb* infection and disease within households at a population-level,[11,24,25] the majority of *M.tb* transmission, occurs from community contact rather than from within the household or close contacts.[26-29]

It is unsurprising that travel on mini-bus is associated with an increased risk of incidence infection in our study. Public transport in low- and middle-income countries, which is often densely crowded and poorly ventilated, has been shown to pose a significant risk of *M.tb* transmission in other settings.[30-32] Interestingly, we found no evidence of an association between incident *M.tb* infection and attendance at healthcare facilities in our setting. Outpatient waiting areas in healthcare clinics in rural Malawi tend to be outdoor spaces with seating areas and are therefore well-ventilated, reducing the risk of transmission,[33] whereas churches are indoors. Social contact pattern data from Zambia, which is a similar

setting to ours, found that the highest adult/youth and child contact hours occurred in churches.[34] Churches have been long identified as locations of *M.tb* transmission;[35-37] and these locales do provide ideal conditions for community 'outbreaks',[38] with effective aerolisation of 'infectious' droplet nuclei generated by singing,[39] within a confined poorly-ventilated space which is filled with individuals of all ages in close proximity to each other. In our study, 67.8% of incident *M.tb* infections in young children were attributable to church attendance, which implies that potential impact from better infection control practices, such as opening windows[40] or installing roof-driven turbines[41] in churches may be substantial in reducing *M.tb* infection risk in this community.

Although *M.tb* incidence was not associated with HIV prevalence at the community level, HIV in fathers was associated with *M.tb* infection in children in our setting. A higher prevalence of *M.tb* infection in children born to HIV-positive women was found on analysis of the baseline TST prevalence data; children who were exposed to HIV *in utero* had a 3-fold increase in odds of TST-positivity compared to children born to HIV-negative mothers,[13] a finding replicated in a number of studies in sub-Saharan Africa.[42-44] However, maternal HIV infection was not associated with incident *M.tb* infection in this analysis. This may have been because children who had a TST≥10mm at baseline (prevalent *M.tb* infection) were not eligible for a repeat skin test thereby illustrating one of the drawbacks of using a TST cohort design to estimate infection risk, which is the exclusion at baseline of individuals at highest risk, i.e. those at highest risk are usually already infected at baseline.[4]

The increased risk of *M.tb* infection seen in children of men with HIV infection may be due to undiagnosed TB in these fathers. Of note, no HIV-positive fathers had been diagnosed with any form of TB during the study period. A recently conducted meta-analysis of sex differences in TB burden in low- and middle-income countries found that the male-to-female prevalence ratio for smear-positive TB was 2.5 (95% CI 2.1 – 3.0).[45] Prolonged duration of infectiousness of TB appears to be associated with male sex, even in countries with a generalized HIV epidemic where females are disproportionately more affected by HIV.[45] The authors also highlighted that men were less likely to access and remain in HIV care, and would therefore be less likely to be assessed and treated for TB. It is assumed that because women are usually the main carers of young children, they are the most likely source of *M.tb* infection to children. But a modelling study by Dodd *et al.* using adult TB disease prevalence, *M.tb* infection incidence in children and social contact pattern data collected in a high HIV/TB burden setting (South Africa and Zambia) found that more than 50% of infections in men, women and children were due to contact with adult men.[5]

Another finding of note from this study is the difference in estimate of the risk of *M.tb* infection per annum in Karonga when using longitudinal data compared to cross-sectional data. Incidence of *M.tb* infection, estimated to be 1.6 per 100 person-years in children under 6 years in this study is much higher than the average annual risk of infection (ARTI) of 0.3% (95% CI: 0.1 – 0.9%) derived from *M.tb* infection prevalence data collected during the baseline TST round of this cohort study.[46] Even using the most stringent cut-off of TST conversion ≥15mm to define *M.tb* infection, infection incidence is estimated to be 1.1 per 100py (95%

CI: 0.8 – 1.4 per 100py), which remains higher than the ARTI estimate although the 95% confidence intervals do cross.

Methodological shortcomings of the use of TST in longitudinal studies have been widely cited,[9] with 'boosting' of pre-sensitisation to mycobacterial antigens from repeat skin testing complicating interpretation.[47,48] 'Boosting' is a phenomenon best described as an effect of the initial tuberculin test which, although it elicits minimal or no reactivity itself, acts as an antigenic stimulus able to recall waned or reverted pre-existing mycobacterial sensitivity, i.e. from remote exposure to *M.tb* but also from remote exposure to other mycobacterial antigens such as BCG vaccination and/or non-tuberculous mycobacteria (NTM) infection.[47] Thus resulting in a level of reactivity observed at the subsequent test which may be indistinguishable from *M.tb* infection, even if the individual has never been infected with *M.tb*.[48,49] From a mixture analysis of the tuberculin data a bimodal distribution of the increment in induration size between TSTs was apparent. The 1st peak most likely representing "boosting" of the TST response to previous BCG vaccination and/or NTM sensitisation with a mean of 6.5mm, and the 2nd peak of the bimodal distribution, with a mean of 14.7mm, most likely representing *M.tb* infection. Using the probability distributions generated by the finite-mixture model we were able to identify a cut-off of TST conversion ≥13mm which minimised the inclusion of false positives (probability < 0.05) whilst maintaining adequate power for the risk factor analysis.

Estimates of risk of infection derived from *M.tb* prevalence data have long been recognized as potentially under-estimating the ARTI.[49,50,51] Waning of tuberculin

sensitivity was noted as early as 1948 in the Prophit survey, which was one of the largest studies of the natural history of *M.tb* infection undertaken in the pre-chemotherapy and pre-BCG vaccination era. At the time the authors highlighted that TST reversions were important enough to considerably modify the incidence of tuberculin sensitization in a given community.[50] However, the authors of a large longitudinal study conducted in Bangalore, India over a decade later concluded that the estimates of infection risk from longitudinal tuberculin data were a 'gross over-estimation'.[52] The study design for this TST cohort study in India included an additional skin test using a tuberculin which had a strength of PPD RT23 20 times stronger than the tuberculin used in the initial skin test. This additional test was only undertaken in those individuals found to have induration less than 13mm at the first skin test. This second skin test was done 1 week after the first skin test as part of the **first round** of skin testing. This is likely to have contributed to an even greater degree of 'boosting' which became apparent on the second round of TST one year later, leading to significant misclassification of *M.tb* infection status.

More recently, controversies have arisen around how best to estimate the incidence of *M.tb* infection as a measure of the *M.tb* transmission in population-level intervention studies; with tuberculin reactivity being deemed an unsatisfactory surrogate.[53] Although interpretation of induration from repeat TST is complicated, similar issues with conversions and reversions have been observed with serial interferon-gamma release assays (IGRA) testing and there are no data to date to suggest that IGRAs are better at identifying the incidence of new *M.tb* infection than the TST.[54] Estimates from longitudinal studies using

IGRA in high burden settings have similarly found much higher estimates of the risk of infection compared to estimates derived from cross-sectional data; with a study in South African adolescents estimating annual risk of 14.0% using longitudinal data compared to 7.3% from cross-sectional data. The fact remains that we currently do not fully understand what TSTs and IGRAs are actually measuring;[55] whether it is a measure of *M.tb* exposure and/or cleared or established *M.tb* infection. The derived metric is mostly likely a composite effect of the intensity of *M.tb* exposure and the susceptibility of the individual to *M.tb* infection,[56] and is almost certainly likely to differ in varying populations.

Incidence of *M.tb* infection provides the most robust proxy measure of the force of infection in the community; if it can indeed be estimated accurately. However, the force of infection in this population, estimated to be around 1.6% per annum, may not be representative of the adult population. The intensity of contact between 'infectious' individuals and children under 6 years of age may not reflect the intensity of contact between 'infectious' adults and susceptible adolescents/adults in the community. Children, adolescents and adults have different social contact patterns,[5,57] and therefore incidence of *M.tb* infection in children most likely does not reflect the incidence in adults.[9] Risk factors for *M.tb* infection in children also differ from those in adults, as shown by the increased risk of household TB contact in young children compared to adults.[58,59] The increase in rate of *M.tb* infection seen with a one year increase in age, even in this young population reflects increasing *M.tb* exposure as children become more mobile and begin to mix with individuals outside of the household. Thus confirming that the assumption used to derive the ARTI from *M.tb* infection

prevalence, that the risk of *M.tb* infection is constant within age categories, is not valid.[9]

Findings from this study in a low-to-moderate burden setting *suggest* that well-established risk factors are driving community *M.tb* transmission, such as the presence of undiagnosed 'infectious' individuals in the community and inadequate ventilation in congregate settings.[60] Simple infection control practices, such as opening windows or even holding congregations in outdoor spaces may go some way to mitigating the risk of transmission. HIV infection in the family also appears to be a risk factor for incident and prevalent *M.tb* infection in young children in Malawi, and a more family-orientated approach in HIV/TB programmes may reduce the future burden of TB morbidity and mortality in HIV-prevalent settings.

Tables

Table 1. Risk factors for incident *M.tb* infection

		N	TST-conversion (≥13mm)	Exposure time (100 py)	*M.tb* infection incidence per 100py (95% CI)		Crude rate ratio* (95% CI)		P value
Overall incidence rate		3066	91	55.80	1.63	1.33 – 2.00	-	-	-
Risk factors									
Child-related									
Age at time of 1st skin test (years)	<1.0	477	8	9.56	0.84	0.42 – 1.67	1		0.004
	1.0 – 1.9	846	17	16.79	1.01	0.63 – 1.63	1.2	0.5 – 2.8	*<0.001 test for trend
	2.0 – 2.9	835	31	16.25	1.91	1.34 – 2.71	2.3	1.1 – 5.0	
	3.0 – 3.9	605	23	9.52	2.42	1.61 – 3.64	2.8	1.2 – 6.3	
	4.0 – 4.9	303	12	3.69	3.25	1.85 – 5.73	3.8	1.6 – 9.5	
Sex	Female	1521	40	27.74	1.44	0.97 – 1.97	1		0.26
	Male	1545	51	28.07	1.82	1.38 – 2.39	1.3	0.8 – 1.9	
BCG status	Not documented	377	20	6.14	3.26	2.10 – 5.05	1		0.002
	Documented	2689	71	49.66	1.43	1.13 – 1.80	0.5	0.3 – 0.7	
Household-related									
Dwelling score	1 (Worst)	581	10	11.89	0.84	0.45 – 1.56	1		0.02
	2	994	41	17.81	2.30	1.70 – 3.13	2.7	1.4 – 5.5	
	3	643	16	12.24	1.31	0.80 – 2.13	1.5	0.7 – 3.4	
	4 (Best)	530	15	8.62	1.74	1.05 – 2.89	2.0	0.9 – 4.5	
	Missing	*318*	*9*						

Continued…

Table 1 (Continued). Risk factors for incident *M.tb* infection

		N	TST-conversion (≥13mm)	Exposure time (100 py)	*M.tb* infection incidence per 100py (95% CI)		Crude rate ratio* (95% CI)		P value
Household SES score	1 (Lowest)	199	7	3.82	1.83	0.87 – 3.85	1		0.56
	2	671	17	12.53	1.35	0.84 – 2.18	0.7	0.3 – 1.7	
	3	1520	51	27.46	1.86	1.41 – 2.44	1.0	0.4 – 2.2	
	4 (Highest)	601	15	10.94	1.37	0.82 – 2.28	0.7	0.3 – 1.8	
	Missing	*75*	*0*						
Food insecurity	No	1227	33	22.59	1.46	1.04 – 2.06	1		0.41
	Yes	1610	52	29.75	1.75	1.3 – 2.29	1.2	0.8 – 1.9	
	Missing	*229*	6						
No. of adult household members	1	62	1	0.90	1.11	0.16 – 7.86	1		0.62
	2 – 3	2241	70	40.48	1.73	1.37 – 2.19	1.6	0.2 – 11.4	
	≥ 4	763	20	14.42	1.39	0.90 – 2.15	1.3	0.2 – 9.4	
Distance from tarmac road (km)	< 0.5	920	26	13.89	1.87	1.28 – 2.75	1		0.17
	0.5 – 2.4	968	26	15.31	1.70	1.16 – 2.50	0.9	0.5 – 1.6	
	2.5 – 5.0	503	9	11.04	0.82	0.42 – 1.57	0.4	0.2 – 1.0	
	> 5.0	675	30	15.54	1.93	1.35 – 2.76	1.0	0.6 – 1.8	
Distance to nearest clinic (km)	< 0.5	117	5	1.90	2.64	1.10 – 6.34	1		0.13
	0.5 – 2.4	948	34	15.53	2.19	1.57 – 3.07	0.9	0.3 – 2.6	
	2.5 – 5.0	1128	30	19.48	1.54	1.08 – 2.20	0.6	0.2 – 1.8	
	> 5.0	873	22	18.90	1.16	0.77 – 1.77	0.5	0.2 – 1.4	
Population density in residential area, person/km^2	< 200	765	24	14.62	1.64	1.10 – 2.45	1		0.86
	200 – 399	758	23	12.76	1.80	1.20 – 2.71	1.2	0.6 – 2.2	
	400 - 999	826	25	17.06	1.47	0.99 – 2.17	0.9	0.5 – 1.7	
	≥1000	717	19	11.37	1.67	1.07 – 2.62	1.0	0.5 – 2.0	

Continued…

Table 1 (Continued). Risk factors for incident *M.tb* infection

		N	TST-conversion (≥13mm)	Exposure time (100 py)	*M.tb* infection incidence per 100py (95% CI)		Crude rate ratio* (95% CI)		P value
TB-related									
Known household infectious TB contact	No	3055	89	55.63	1.60	1.30 – 1.97	1		0.04
	Yes	11	2	0.17	11.70	2.93 – 46.78	6.9	1.6 – 29.3	
Distance from nearest known infectious TB case during the study period (metres)	>200	2783	81	51.42	1.58	1.27 – 1.96	1		0.06
	100 – 200	146	2	2.33	0.86	0.21 – 3.43	0.5	0.1 – 2.2	
	<100	128	6	1.88	3.20	1.44 – 7.11	2.0	0.8 – 4.7	
	Within HH	11	2	0.17	11.70	2.93 – 46.78	7.1	1.7 – 30.1	
Community *M. tuberculosis* exposure measure:									
Av. **smear-positive TB** notification rate (per 100,000/area/yr)	≤30	1783	42	35.53	1.18	0.87 – 1.60	1		0.002
	>30	1283	49	20.28	2.42	1.83 – 3.20	2.1	1.3 – 3.2	
HIV-related									
HIV exposure *in utero*	No	2588	77	47.33	1.63	1.30 – 2.03	1		0.42
	Yes	306	11	5.06	2.18	1.21 – 3.93	1.6	0.7 – 2.4	
	Unknown	172	3	3.41	0.88	0.28 – 2.72	0.6	0.2 – 1.8	
HIV-positive mother	No	2609	75	47.35	1.58	1.26 – 1.99	1		0.16
	Yes	340	14	5.54	2.53	1.50 – 4.27	1.5	0.9 – 2.7	
	Unknown	117	2	2.91	0.69	0.17 – 2.75	0.5	0.1 – 1.9	
HIV-positive father	No	2043	50	37.54	1.33	1.01 – 1.76	1		0.04
	Yes	196	10	3.40	2.95	1.59 – 5.47	2.2	1.1 – 4.4	
	Unknown	827	31	14.87	2.09	1.47 – 2.97	1.6	1.0 – 2.5	
HIV prevalence in residential area (%)	<4.0	680	28	14.84	1.89	1.30 – 2.73	1		0.15
	4.0 – 7.9	1837	54	37.93	1.42	1.09 – 1.86	0.8	0.5 – 1.2	
	≥8.0%	549	9	3.04	2.96	1.54 – 5.69	1.5	0.7 – 3.5	

*Adjusted for clustering in residential area

Table 2. Attendance at gathering places as risk factors for incident *M.tb* infection (n=2041)

		N (%)	No. of *M.tb* infections	Exposure time (100 py)	*M.tb* infection incidence (per 100py) 95% CI		Crude rate ratio*	95% CI	P value
Overall study population incidence rate		2041	66	44.82	1.50	1.18 – 1.90	-	-	-
Attendance at gathering place (no. of visits per year)									
Church	<4	223	3	4.23	0.71	0.23 – 2.20	1		0.15
	≥4	1796	63	40.21	2.36	1.22 – 2.01	2.4	0.7 – 7.7	
Health-care facility	None	319	8	6.51	1.23	0.61 – 2.46	1		0.47
	At least one visit	1701	58	37.95	1.53	1.18 – 1.98	1.3	0.6 – 2.8	
Travel on mini-bus	None	1008	26	22.24	1.17	0.80 – 1.72	1		0.07
	At least one trip	1012	40	22.20	1.80	1.32 – 2.46	1.6	1.0 – 2.6	
Market	None	1133	44	24.22	1.82	1.35 – 2.44	1		0.13
	At least one visit	902	22	20.48	1.08	0.71 – 1.63	0.7	0.4 – 1.2	
Funeral	None	1663	54	36.80	1.47	1.12 – 1.92	1		0.92
	At least one attendance	372	12	7.89	1.52	0.86 – 2.68	1.0	0.5 – 1.9	
School attendance	No	1281	36	28.00	1.29	0.93 – 1.78	1		0.09
	Yes	755	30	16.71	1.80	1.26 – 2.57	1.6	0.9 – 2.7	

*Adjusted for clustering in residential area
py person-year CI confidence interval

Table 3. Multivariable analysis of risk factors for incident *M.tb* infection

		***M.tb* infection defined as TST conversion ≥13mm**				**Sensitivity analyses**								
						***M.tb* infection defined as TST conversion ≥10mm**			***M.tb* infection defined as TST conversion ≥15mm**			**TST conversion ≥13mm excluding children with conversions 9-12mm**		
		Multivariable model 1* (n=1826)			PAF (95% CI)	Multivariable model 2* (n=1826)			Multivariable model 3* (n=1826)			Multivariable model 4* (n=1778)		
		RR	**95% CI**	**p value**		**RR**	**95% CI**	**p value**	**RR**	**95% CI**	**p value**	**RR**	**95% CI**	**p value**
Age at time of 1st skin test (years)		1.4	1.1 – 3.0	0.01	-	1.3	1.1 – 1.5	0.02	1.4	1.1 -1.8	0.01	1.4	1.1 – 3.1	0.01
HIV-positive father	No	1		0.04		1		0.12	1		0.04	1		0.04
	Yes	2.5	1.1 – 5.8		6.9% (1.1 – 9.5)	1.8	0.8 – 3.7		3.0	1.2 – 7.6		2.6	1.1 – 5.9	
	Unknown	1.7	1.0 – 3.0			1.5	1.0 – 2.4		1.7	0.9 – 3.5		1.8	1.0 – 3.1	
Known community M.tb exposure:														
Av. smear-positive TB notification rate (per 100,000/ area/yr)	≤30	1		0.001		1		0.002	1		0.02	1		0.001
	>30	2.0	1.3 – 3.1		28.7% (13.2 – 38.9)	1.7	1.2 – 2.4		2.0	1.2 – 3.4		2.1	1.3 – 3.1	
Attendance at church (visits per year)	<4	1		0.04		1		0.04	1		0.15	1		0.04
	≥4	3.4	0.8 – 14.0		67.8% (-24.1 – 90.0)	2.6	0.9 – 7.1		2.5	0.6 – 10.4		3.5	0.8 – 14.3	
Travel on mini-bus	None	1		0.03		1		0.05	1		0.05	1		0.02
	Any trip	1.8	1.1 – 3.1		27.7% (5.7 – 42.2)	1.5	1.0 – 2.3		1.8	1.0 – 3.4		1.8	1.1 – 1.7	

*Adjusted for clustering in residential and all risk factors in the model in addition to sex, household SES and dwelling structure

TST tuberculin skin test RR rate ratio CI confidence interval PAF population attributable fraction SES socioeconomic status

Figures

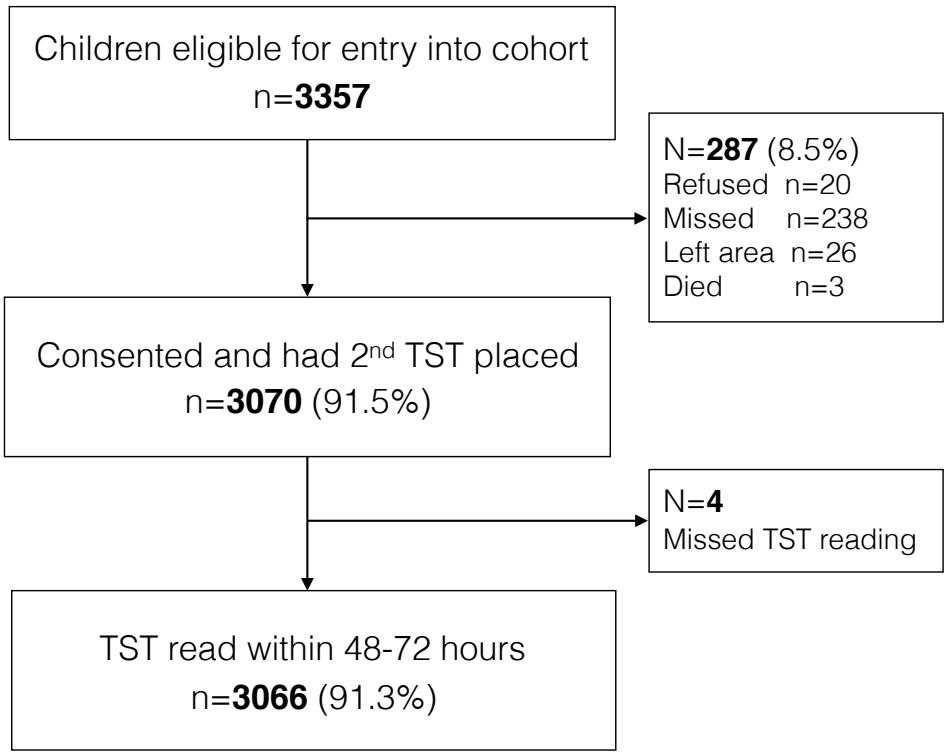


Figure 1. Flow diagram illustrating study participant flow at each stage from eligibility through to analysis. TST=tuberculin skin test

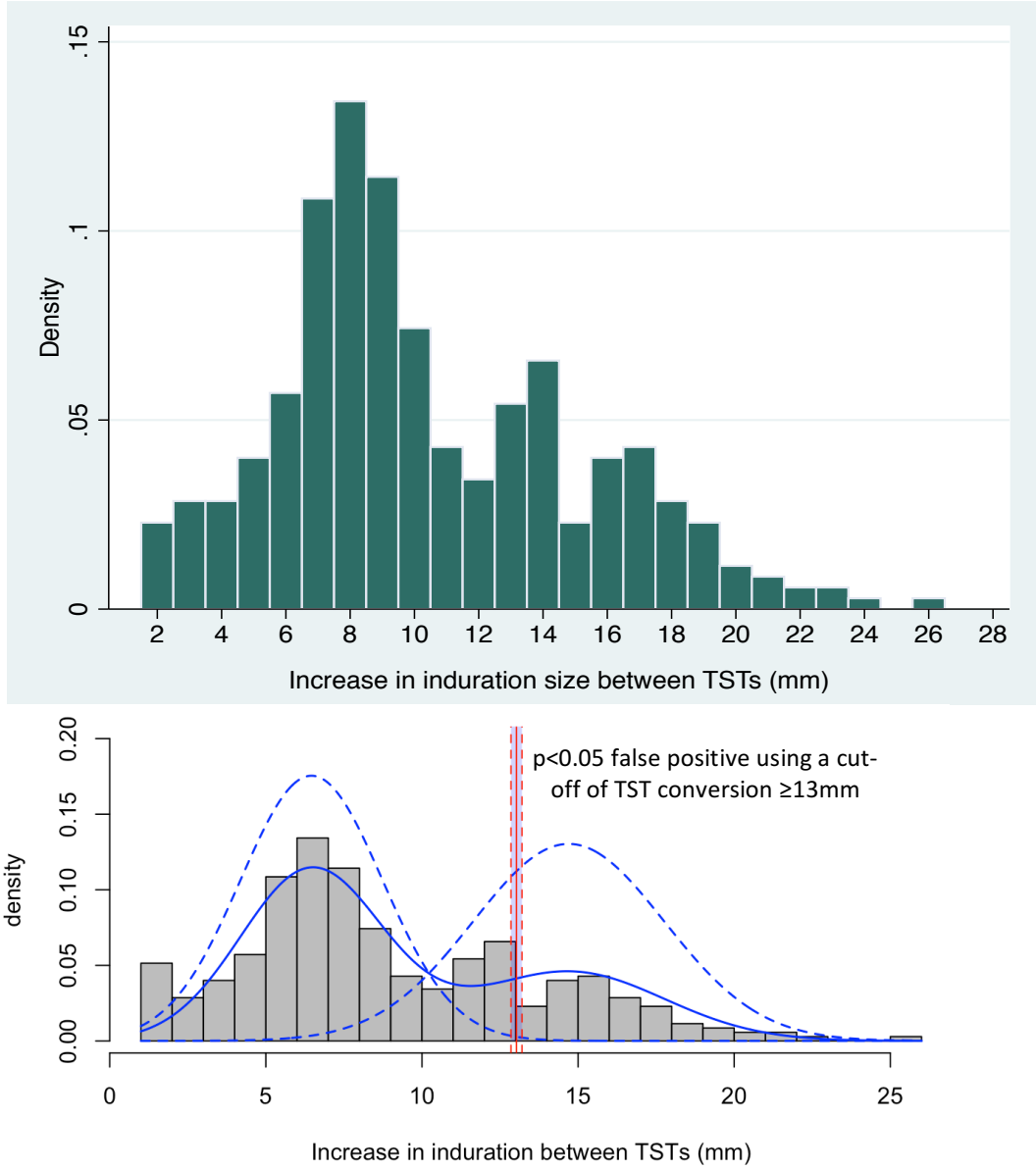


Figure 2. (a) Histogram of increase in induration size between 1st and 2nd TST [n=350] (b) 2-component mixture model based on implementation of Expectation Maximization (EM) algorithm to derive the cut-off value for TST conversion (KEY: grey bars=histogram bars; blue solid line=kernel density estimation of the distribution; blue dashed line=finite-mixture model; red solid line=cut-off value; red dashed lines=95% CI around cut-off value)

Online supplement

Table S1. Characteristics of study participants missing data on gathering places compared with study participants who had data collected

	Data available on gathering places (n=2041)	Missing data on gathering places (n=1025)
Age in years (mean (sd))	2.1 (1.0)	2.7 (1.3)
Male (%)	50.2	50.8
Proportion in the lowest SES category	7.2% (95% CI 6.1 – 8.4)	5.7% (95% CI 4.4 – 7.4)
Median no. of adults in HH (IQR)	2 (2 - 3)	2 (2 - 2)
Father HIV positive (%)	5.8% (95% CI 4.9 – 6.9)	7.5% (95% CI 6.0 – 9.3)
Mother HIV positive (%)	9.3% (95% CI 8.1 -10.7)	14.6% (95% CI 12.5 – 17.0)
Community M.tb exposure: av notification rate >30 per 100,000 population per year	39.4% (95% CI 37.3 – 41.6)	46.8% (95% CI 43.7 – 49.9)
Resident in an area where HIV prevalence >8.0%	9.5%	0%
Household TB contact	0.3% (95% CI 0.1 – 0.6)	0.5% (95% CI 0.2 – 1.1)
M.tb incidence rate	1.5 per 100py (95% CI 1.2 – 1.9)	2.2 per 100py (95% CI 1.5 – 3.3)

sd standard deviation; HH household; IQR interquartile range; CI confidence interval SES socioeconomic status; py person-years

References

1. WHO. The 2016 Update, Global Health Workforce Statistics. *(*http://www.who.int/hrh/statistics/hwfstats/ *[Accessed 2017/08/02]*. Geneva: World Health Organisation; 2016.
2. Uplekar M, Weil D, Lonnroth K, et al. WHO's new end TB strategy. *Lancet.* 2015;385(9979):1799-1801.
3. Lienhardt C, Lonnroth K, Menzies D, et al. Translational Research for Tuberculosis Elimination: Priorities, Challenges, and Actions. *PLoS Med.* 2016;13(3):e1001965.
4. Yates TA, Khan PY, Knight GM, et al. The transmission of Mycobacterium tuberculosis in high burden settings. *Lancet Infect Dis.* 2016;16(2):227-238.
5. Dodd PJ, Looker C, Plumb ID, et al. Age- and Sex-Specific Social Contact Patterns and Incidence of Mycobacterium tuberculosis Infection. *Am J Epidemiol.* 2016;183(2):156-166.
6. Shingadia D, Novelli V. Diagnosis and treatment of tuberculosis in children. *Lancet Infect Dis.* 2003;3(10):624-632.
7. Marais BJ, Obihara CC, Warren RM, Schaaf HS, Gie RP, Donald PR. The burden of childhood tuberculosis: a public health perspective. *Int J Tuberc Lung Dis.* 2005;9(12):1305-1313.
8. van Rie A, Beyers N, Gie RP, Kunneke M, Zietsman L, Donald PR. Childhood tuberculosis in an urban population in South Africa: burden and risk factor. *Arch Dis Child.* 1999;80(5):433-437.
9. Rieder H. Annual risk of infection with Mycobacterium tuberculosis. *Eur Respir J.* 2005;25(1):181-185.
10. Seddon JA, Jenkins HE, Liu L, et al. Counting children with tuberculosis: why numbers matter. *Int J Tuberc Lung Dis.* 2015;19 Suppl 1:9-16.
11. Lienhardt C, Fielding K, Sillah J, et al. Risk factors for tuberculosis infection in sub-Saharan Africa: a contact study in The Gambia. *Am J Respir Crit Care Med.* 2003;168(4):448-455.
12. Crampin AC, Dube A, Mboma S, et al. Profile: the Karonga Health and Demographic Surveillance System. *Int J Epidemiol.* 2012;41(3):676-685.
13. Khan PY, Glynn JR, Fielding KL, et al. Risk factors for Mycobacterium tuberculosis infection in 2-4 year olds in a rural HIV-prevalent setting. *Int J Tuberc Lung Dis.* 2016;20(3):342-349.
14. Tebruegge M, Bogyi M, Soriano-Arandes A, Kampmann B, Paediatric Tuberculosis Network European Trials G. Shortage of purified protein derivative for tuberculosis testing. *Lancet.* 2014;384(9959):2026.
15. American Thoracic S. Diagnostic standards and classification of tubreculosis. *Am Rev Respir Dis.* 1990;142:725-735.
16. Boccia D, Hargreaves J, Ayles H, Fielding K, Simwinga M, Godfrey-Faussett P. Tuberculosis infection in Zambia: the association with relative wealth. *Am J Trop Med Hyg.* 2009;80(6):1004-1011.

17. Odone A, Crampin AC, Mwinuka V, et al. Association between socioeconomic position and tuberculosis in a large population-based study in rural Malawi. *PLoS One.* 2013;8(10):e77740.
18. Mboma SM, Houben RM, Glynn JR, et al. Control of (Multi)Drug Resistance and Tuberculosis Incidence over 23 Years in the Context of a Well-Supported Tuberculosis Programme in Rural Malawi. *PLoS One.* 2013;8(3):e58192.
19. Benaglia T, Chauveau D, Hunter DR, Young D. mixtools: An R Package for Analyzing Finite Mixture Models. *Journal of Statistical Software.* 2009;32(6):1-29.
20. Davies GR, Fine PE, Vynnycky E. Mixture analysis of tuberculin survey data from northern Malawi and critique of the method. *Int J Tuberc Lung Dis.* 2006;10(9):1023-1029.
21. Choisy M. Identifying a cutoff value from bimodal data. *Available at:* http://marcchoisy.free.fr/fmm/index.html *[Accessed 5 Aug. 2017].* 2017.
22. Trang NV, Choisy M, Nakagomi T, et al. Determination of cut-off cycle threshold values in routine RT-PCR assays to assist differential diagnosis of norovirus in children hospitalized for acute gastroenteritis. *Epidemiol Infect.* 2015;143(15):3292-3299.
23. Rockhill B, Newman B, Weinberg C. Use and misuse of population attributable fractions. *Am J Public Health.* 1998;88(1):15-19.
24. Andersen S, Geser A. The distribution of tuberculous infection among households in African communities. *Bull World Health Organ.* 1960;22:39-60.
25. Shapiro AE, Variava E, Rakgokong MH, et al. Community-based Targeted Case-finding for Tuberculosis and HIV in Household Contacts of Tuberculosis Patients in South Africa. *Am J Respir Crit Care Med.* 2012.
26. Verver S, Warren RM, Munch Z, et al. Proportion of tuberculosis transmission that takes place in households in a high-incidence area. *Lancet.* 2004;363(9404):212-214.
27. Crampin AC, Floyd S, Ngwira BM, et al. Assessment and evaluation of contact as a risk factor for tuberculosis in rural Africa. *Int J Tuberc Lung Dis.* 2008;12(6):612-618.
28. Buu TN, van Soolingen D, Huyen MN, et al. Tuberculosis acquired outside of households, rural Vietnam. *Emerg Infect Dis.* 2010;16(9):1466-1468.
29. Glynn JR, Guerra-Assuncao JA, Houben RM, et al. Whole Genome Sequencing Shows a Low Proportion of Tuberculosis Disease Is Attributable to Known Close Contacts in Rural Malawi. *PLoS One.* 2015;10(7):e0132840.
30. Horna-Campos OJ, Consiglio E, Sanchez-Perez HJ, Navarro A, Cayla JA, Martin-Mateo M. Pulmonary tuberculosis infection among workers in the informal public transport sector in Lima, Peru. *Occupational and Environmental Medicine.* 2011;68(2):163-165.
31. Feske ML, Teeter LD, Musser JM, Graviss EA. Giving TB wheels: Public transportation as a risk factor for tuberculosis transmission. *Tuberculosis (Edinb).* 2011;91 Suppl 1:S16-23.

32. Andrews JR, Morrow C, Wood R. Modeling the role of public transportation in sustaining tuberculosis transmission in South Africa. *Am J Epidemiol.* 2013;177(6):556-561.
33. Taylor JG, Yates TA, Mthethwa M, Tanser F, Abubakar I, Altamirano H. Measuring ventilation and modelling M. tuberculosis transmission in indoor congregate settings, rural KwaZulu-Natal. *Int J Tuberc Lung Dis.* 2016;20(9):1155-1161.
34. McCreesh N, Looker C, Dodd PJ, et al. Comparison of indoor contact time data in Zambia and Western Cape, South Africa suggests targeting of interventions to reduce Mycobacterium tuberculosis transmission should be informed by local data. *BMC Infect Dis.* 2016;16:71.
35. Mangura BT, Napolitano EC, Passannante MR, McDonald RJ, Reichman LB. Mycobacterium tuberculosis miniepidemic in a church gospel choir. *Chest.* 1998;113(1):234-237.
36. Cook SA, Blair I, Tyers M. Outbreak of tuberculosis associated with a church. *Commun Dis Public Health.* 2000;3(3):181-183.
37. Classen CN, Warren R, Richardson M, et al. Impact of social interactions in the community on the transmission of tuberculosis in a high incidence area. *Thorax.* 1999;54(2):136-140.
38. Lincoln EM. Epidemics of tuberculosis. *Bibliotheca tuberculosea.* 1965;21:157-201.
39. Loudon RG, Roberts RM. Singing and the dissemination of tuberculosis. *Am Rev Respir Dis.* 1968;98(2):297-300.
40. Escombe AR, Oeser CC, Gilman RH, et al. Natural ventilation for the prevention of airborne contagion. *PLoS Med.* 2007;4(2):e68.
41. Cox H, Escombe R, McDermid C, et al. Wind-driven roof turbines: a novel way to improve ventilation for TB infection control in health facilities. *PLoS One.* 2012;7(1):e29589.
42. Cotton MF, Schaaf HS, Lottering G, et al. Tuberculosis exposure in HIV-exposed infants in a high-prevalence setting. *Int J Tuberc Lung Dis.* 2008;12(2):225-227.
43. Cranmer LM, Kanyugo M, Jonnalagadda SR, et al. High prevalence of tuberculosis infection in HIV-1 exposed Kenyan infants. *Pediatr Infect Dis J.* 2014;33(4):401-406.
44. Marquez C, Chamie G, Achan J, et al. Tuberculosis Infection in Early Childhood and the Association with HIV-exposure in HIV-uninfected Children in Rural Uganda. *Pediatr Infect Dis J.* 2016;35(5):524-529.
45. Horton KC, MacPherson P, Houben RM, White RG, Corbett EL. Sex Differences in Tuberculosis Burden and Notifications in Low- and Middle-Income Countries: A Systematic Review and Meta-analysis. *PLoS Med.* 2016;13(9):e1002119.
46. Khan PY, Glynn JR, Mzembe T, et al. Challenges in the estimation of the annual risk of Mycobacterium tuberculosis infection in children aged under 5 years. *Am J Epidemiol.* 2017.
47. Menzies D. Interpretation of repeated tuberculin tests. Boosting, conversion, and reversion. *Am J Respir Crit Care Med.* 1999;159(1):15-21.

48. Menzies D. What does tuberculin reactivity after bacille Calmette-Guerin vaccination tell us? *Clin Infect Dis.* 2000;31 Suppl 3:S71-74.

49. Cauthen GM, Pio A, ten Dam HG. Annual risk of tuberculous infection. 1988. *Bull World Health Organ.* 2002;80(6):503-511; discussion 501-502.

50. Daniels M, Hall IM, Ridehalgh F, Springett VH. *Tuberculosis in Young Adults. Report on the Prophit Tuberculosis Survey 1935-1944.* London: H. K. Lewis & Co.; 1948.

51. Fine PE, Bruce J, Ponnighaus JM, Nkhosa P, Harawa A, Vynnycky E. Tuberculin sensitivity: conversions and reversions in a rural African population. *Int J Tuberc Lung Dis.* 1999;3(11):962-975.

52. Narain R, Nair SS, Chandrasekhar P, Rao GR. Problems connected with estimating the incidence of tuberculosis infection. *Bull World Health Organ.* 1966;34(4):605-622.

53. Cobelens F, van Leth F, van 't Hoog A. Design of pragmatic trials of tuberculosis interventions. *Lancet.* 2014;383(9913):213-214.

54. Pai M, Denkinger CM, Kik SV, et al. Gamma Interferon Release Assays for Detection of Mycobacterium tuberculosis Infection. *Clin Microbiol Rev.* 2014;27(1):3-20.

55. Cobelens F. Measuring TB transmission: comments. Paper presented at: Halting TB transmission in HIV endemic settings, June 2017; Cape Town, South Africa.

56. Esmail H, Barry CE, 3rd, Wilkinson RJ. Understanding latent tuberculosis: the key to improved diagnostic and novel treatment strategies. *Drug discovery today.* 2012;17(9-10):514-521.

57. Mossong J, Hens N, Jit M, et al. Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Med.* 2008;5(3):e74.

58. Middelkoop K, Bekker LG, Morrow C, Lee N, Wood R. Decreasing household contribution to TB transmission with age: a retrospective geographic analysis of young people in a South African township. *BMC Infect Dis.* 2014;14(1):221.

59. Zelner JL, Murray MB, Becerra MC, et al. Age-specific risks of tuberculosis infection from household and community exposures and opportunities for interventions in a high-burden setting. *Am J Epidemiol.* 2014;180(8):853-861.

60. Raffalli J, Sepkowitz KA, Armstrong D. Community-based outbreaks of tuberculosis. *Arch Intern Med.* 1996;156(10):1053-1060.

6 Effect of HIV/ART on the 'infectiousness' of smear-positive pulmonary tuberculosis

6.1 Research paper V

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT
www.lshtm.ac.uk

Registry
T: +44(0)20 7299 4646
F: +44(0)20 7299 4656
E: registry@lshtm.ac.uk

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

RESEARCH PAPER COVER SHEET

PLEASE NOTE THAT A COVER SHEET MUST BE COMPLETED FOR EACH RESEARCH PAPER INCLUDED IN A THESIS.

SECTION A – Student Details

Student	Palwasha Yousafzai Khan
Principal Supervisor	Professor Judith Glynn
Thesis Title	Investigating Mycobacterium tuberculosis transmission in rural Malawi

If the Research Paper has previously been published please complete Section B, if not please move to Section C

SECTION B – Paper already published

Where was the work published?			
When was the work published?			
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion			
Have you retained the copyright for the work?*	Choose an item.	Was the work subject to academic peer review?	Choose an item.

**If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.*

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	International Journal of Tuberculosis and Lung Disease
Please list the paper's authors in the intended authorship order:	Khan PY, Crampin AC, Mzembe T, Koole O, Fielding KL, Kranzer K, Glynn JR
Stage of publication	**In press**

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	I conceived of the idea for the analysis, led on all aspects of the fieldwork, data collection and analysis. I wrote the first and final drafts of the paper

Student Signature: [redacted] **Date:** 28/7/2017

Supervisor Signature: [redacted] **Date:** 28/7/2017

TITLE PAGE

Full title: Does antiretroviral treatment increase the infectiousness of smear-positive pulmonary tuberculosis?

Running head title: ART and infectiousness of smear-positive PTB

Word count: 2130

Authors

1. Palwasha Y. Khan

   Department of Infectious Disease Epidemiology

   London School of Hygiene and Tropical Medicine, London, UK

   Karonga Prevention Study, Chilumba, Malawi
2. Amelia C. Crampin

   Department of Infectious Disease Epidemiology

   London School of Hygiene and Tropical Medicine, London, UK

   Karonga Prevention Study, Chilumba, Malawi
3. Themba Mzembe

   Karonga Prevention Study, Chilumba, Malawi
4. Olivier Koole

   Department of Infectious Disease Epidemiology

   London School of Hygiene and Tropical Medicine, London, UK
5. Katherine L. Fielding

   Department of Infectious Disease Epidemiology

London School of Hygiene and Tropical Medicine, London, UK

6. Katharina Kranzer

   Department of Infectious Disease Epidemiology

   London School of Hygiene and Tropical Medicine, London, UK;

   National and Supranational Mycobacterium Reference Laboratory, Forschungszentrum Borstel, Germany

7. Judith R. Glynn

   Department of Infectious Disease Epidemiology

   London School of Hygiene and Tropical Medicine, London, UK

Corresponding author:

Palwasha Y. Khan

Department of Infectious Disease Epidemiology

London School of Hygiene & Tropical Medicine

Keppel Street, London, WD1E 7HT, United Kingdom

Mobile: +44 (0) 7811 902 455

Email: palwasha.khan@lshtm.ac.uk

Abstract

Background: Understanding of the effects of HIV and anti-retroviral treatment (ART) on *Mycobacterium tuberculosis* (*M.tb*) transmission dynamics remains limited. We undertook a cross-sectional study of household contacts of smear-positive pulmonary tuberculosis cases to assess the effect of established ART on the infectiousness of TB

Method: Prevalence of tuberculin skin test (TST) positivity was compared between contacts aged 2-10 years of index cases who were: HIV-negative; HIV-positive not on ART; on ART <1 year and on ART for ≥1 year. Random effects logistic regression was used to take account of clustering within households.

Results: *M.tb* infection prevalence in contacts of HIV-negative, HIV-positive on ART ≥1 year and HIV-positive not on ART/ on ART <1 year index cases was 44%, 21% and 22% respectively. Compared to contacts of HIV-positive index cases not on ART or recently started on ART, the odds of TST positivity was similar in contacts of HIV-positive index cases on ART ≥1 year (aOR 1.0; 95% CI 0.3 – 3.7). The odds were 2.9 times higher in child contacts of HIV-negative index cases (aOR 2.9; 95% CI 1.0 – 8.2).

Conclusions: We found no evidence that established ART increased the infectiousness of smear positive HIV-positive index cases.

Word count: 196

Key Words: *M.tb* infection; infectiousness; tuberculosis; antiretroviral treatment; HIV

Introduction

The human immunodeficiency virus (HIV) pandemic continues to challenge global tuberculosis (TB) control,[1,2] yet our understanding of the effects of HIV and anti-retroviral treatment (ART) on *Mycobacterium tuberculosis* (*M.tb*) transmission dynamics remains limited.[3-7]

In HIV-positive individuals, ART reduces TB incidence across all CD4 cell counts,[8] yet despite long-term ART, TB incidence remains higher than in HIV-negative people in both high and low TB-burden settings.[9,10] As life expectancy is greatly extended by ART, the cumulative lifetime risk of TB for HIV-positive people remains very high.[9] While HIV-positive TB patients with advanced immunosuppression are less likely to transmit to household contacts than their HIV-negative counterparts,[11-17] partly mediated through a lower bacillary load in sputum,[11,16,17] ART may increase the infectiousness of TB by shifting the clinical manifestation to be more similar to HIV-negative patients.[18,19]

Concerns have been raised that the increased life-expectancy and possible increased infectiousness due to ART might increase TB incidence at a population-level.[20] This might be negated by reduced HIV transmission,[21,22] but a rebound in TB incidence, exceeding levels present before ART roll-out, is possible if good adherence to ART is not sustained.[23] Programmatic data from South Africa, Malawi and Zimbabwe have shown a reduction in TB incidence, as inferred from trends in TB case notification, in association with ART scale-up.[24,25] However, short-term reductions in TB incidence may be due to protection from progression to disease rather than a reduction in *M.tb* transmission *per se*.

We examine the effect of HIV and ART on intra-household *M.tb* transmission by measuring the prevalence of *M.tb* infection among child contacts of adult smear-positive tuberculosis cases.

Methods

Study setting

Karonga District, northern Malawi, is predominantly rural, with an adult HIV prevalence of 9% and new smear-positive TB incidence of 87 per 100,000 adults per year.[26] 60% of TB cases are HIV-positive.[26] The first ART clinic opened in 2005 and by 2012, 16 clinics in the district were certified to initiate and provide ART.[27]

Study design

A cross-sectional household study of all diagnosed smear-positive TB cases in the district was conducted from January 2013 to April 2015. Households were eligible if a smear-positive case had lived there for at least 2 weeks after the onset of symptoms and prior to initiation of treatment. Bacteriological, demographic and clinical (including HIV and ART status) data from all patients starting TB treatment in the district have been collected in a TB case cohort study since 1988 to date, which has been previously described.[26,28]

Households were visited approximately 6 weeks after TB diagnosis of index case (date of first smear-positive sputum). All children aged 2 to 10 years-old resident within the household were included. A tuberculin skin test (TST) was administered and read according to standard international guidelines[29] using 2

international units of RT23 (Statens Serum Institute, Copenhagen, Denmark) and induration was measured 48-72 hours later. A positive TST was defined as induration ≥ 10mm. Children under the age of 2 years were excluded to minimize misclassification of infection status due to false-positive TST as a result of recent BCG vaccination.[30]

A questionnaire was completed which included data on demographics, BCG vaccination status, exposure to index case (whether index case was mother, duration of sleeping in same room and of living in the same household whilst index case was symptomatic), and household characteristics (number of residents, socio-economic indicators including quality of dwelling place). A composite score for household socioeconomic status was created using head of household employment, number of assets, food security and availability of soap, and a composite score for quality of dwelling place was based on building materials, type of roof, number of rooms, water source, presence of glass windows, electricity and latrine type.

Any child with symptoms suggestive of TB (fever, weight loss, failure to thrive, night sweats or cough) was reviewed by a clinician and referred to the district hospital where appropriate. All children aged <5 years without evidence of active disease were commenced on 6-month isoniazid preventive treatment (5mg/kg once daily) irrespective of TST induration size (as per Malawi National TB programme guidelines).[31]

Ethics approval

The study was approved by the Malawi National Health Sciences Research Committee (#1049) and the London School of Hygiene & Tropical Medicine Ethics committee (#6285). At the time of study recruitment, smear positive pulmonary TB patients were asked for written consent to visit their household(s) to screen household members for infection and disease. Written informed consent was then obtained from a parent or guardian of each participating child at the time of household visit.

Statistical analysis

Prevalence of tuberculin skin test (TST) positivity, was compared between household contacts by HIV and ART status of the index case, distinguishing those not on ART or on ART for <1 year and on ART for ≥1 year at tuberculosis diagnosis. We performed univariate analyses for covariates known to be risk factors for *M.tb* infection. They were assessed as confounders of the association between HIV/ART status and TST positivity, first individually and then in a multivariable model, using random effects logistic regression to account for clustering within household.

Sputum smear grade and duration of symptoms were not included in the initial multivariable model as they were considered to be on the causal pathway between HIV/ART status and prevalence of TST positivity in the child contact. However, they were examined as mediators of the association in the subsequent models.

Sensitivity analyses: The analysis was repeated (i) grouping all patients on ART irrespective of time on ART, (ii) separating patients on ART for ≥2 years, and (iii) using a TST cut-off ≥ 15mm.

Results

388 child contacts of 187 index cases were eligible for inclusion (Figure 1). 309 child contacts had a TST placed and read within 48-72 hours (80%; 153 index cases). HIV/ART status was missing for 3 index cases (7 contacts). Therefore 302 child contacts of 150 index cases were included in the final analysis. One hundred and seventy-seven children (58.4%) had no induration; the frequency distribution of those children with non-zero TST induration (n=125) is shown in Figure 2.

Index case characteristics

Of the 150 index cases, 63% were male. Median age was 33.6 years (interquartile range [IQR] 28.2 – 39.3) for female index cases and 37.6 years (IQR 30.6 – 44.9) for male. 22% of index cases were on ART at TB diagnosis, with a median duration on ART of 2.8 years (IQR 1.2 – 4.5; n=33). Table 1 shows index case characteristics by HIV/ART status. The median duration of symptoms prior to TB diagnosis (by self-report) was shortest in HIV-positive on ART ≥1 year (8.3 weeks IQR: 7.0 – 16.3).

TST positivity in child contacts

The prevalence of TST positivity among all child contacts was 34.4%. TST positivity in the child contacts of HIV-positive index cases not on ART was 23.1%

(15/65); as there were only 11 child contacts of 8 HIV-positive index cases on ART for <1 year (TST positivity 18.2%) this category was combined with the HIV-positive not on ART index cases. TST positivity was 44.1% (75/170), 21.4% (12/56) and 22.4 % (17/76) in contacts of index cases who were HIV-negative, HIV-positive on ART ≥1 year, and HIV-positive not on ART/on ART for <1 year, respectively (Table 2).

Factors associated with TST positivity in child contacts included: HIV/ART status of index case; sex of index case; whether index case was the mother; index case sputum smear grade; and degree of exposure of contact (Table 2).

Compared to contacts of HIV-positive index cases not on ART or on ART for <1 year, the odds of a positive TST were higher in the contacts of HIV-negative index cases (crude OR 4.3: 95% CI 1.4-13.1, reducing to OR 2.9: 95%CI 1.0–8.2 after adjustment for socio-demographic factors) but not in contacts of HIV-positive index cases who had been on ART for ≥ 1 year (crude OR 1.1: 95%CI 0.3–4.6, adjusted OR 1.0; 95% CI 0.3–3.7, Table 2). Further adjustment assessed the effect of factors that may be on the causal pathway. Adjustment for duration of symptoms (model 2a, Table 3) made little difference to the odds of TST positivity in contacts of HIV-negative index cases but adjustment for smear grade (model 2b) reduced the association (aOR 2.2: 95% CI 0.8–6.3). Among contacts of HIV-positive index cases there was no association of TST positivity with ART duration in any model (Table 3).

The results of the sensitivity analyses are also shown in Table 3. Using a cut-off TST ≥15mm, regrouping those on ART ignoring duration, or separating those on ART >2 years made little difference to the results.

Discussion

We found no evidence that child contacts of HIV-positive TB patients on ART were more likely to have a positive TST compared to child contacts of HIV-positive TB patients not on ART. However, child contacts of HIV-negative individuals had nearly three times the odds of having a positive TST compared to child contacts of HIV-positive TB patients not on ART; this was partly explained by differences in the degree of smear positivity.

Some TB household contact studies have found no difference in infectiousness between HIV-positive TB patients compared to HIV-negative TB patients if HIV-positive index cases were less immunosuppressed (CD4>250)[16] or were smear-positive and/or had cavitatory disease.[17] Heterogeneity observed in the estimates of infectiousness of HIV-positive TB patients compared to HIV-negative TB patients has been well-described,[16,17,32] although no studies to date have examined the effect of ART status of the HIV-positive index case. Possible reasons for heterogeneity seen include differences in patient eligibility (smear-positive only *vs* all TB patients), study settings (high *vs* low HIV and TB background prevalence), household contacts screened for *M.tb* infection (adults and children *vs* children only) and study-related biases, such as exposure assessment bias and recall bias.[33] Study calendar period may also influence the

estimates of infectiousness as the degree of immunosuppression of HIV-positive TB cases on a population-level will be a function of the maturity of the HIV epidemic and the time since roll-out of ART.

The lack of evidence for increased infectiousness in TB patients established on ART, compared to those not on ART in our study may be due to earlier diagnosis resulting in a shorter duration of infectiousness. However, adjusting for duration of symptoms in model 2a did not alter the odds of a positive TST in contacts of index cases who had been on ART ≥ 1 years. This might be due to the fact that duration of symptoms at time of TB diagnosis is notoriously difficult to recall accurately. Interestingly, a contemporaneous study of attendance at the HIV/ART clinic undertaken at the Karonga District Hospital found that HIV-positive individuals on ART attended clinic much more regularly than those not on ART; median number of visits per year were 5 (IQR 4 – 6) for ART patients and 1 (IQR 1-2) for HIV patients not on ART (unpublished data). This gives much greater opportunity for early diagnosis of TB. Another reason for the absence of evidence for increased infectiousness in TB patients established on ART may be that in our population in northern Malawi, the effect of ART is masked by the degree of immunodeficiency of the index case because ART is started in individuals with more advanced HIV infection.

CD4 cell count testing is not done routinely in Karonga, northern Malawi resulting in the absence of a biological marker for the level of immunosuppression of HIV-positive TB patients and degree of immune reconstitution in those established on ART in this study. Limitations of our study

also include the absence of radiological data to assess extent of lung cavitation by HIV/ART status. These data would have helped to interpret the apparent lack of effect of ART on infectiousness. Knowing the HIV status of the child contacts would have also strengthened this study, as an HIV-positive child with advanced immunosuppression may be more likely to have a false-negative TST than an HIV-negative child, and an HIV-positive child is more likely to be resident in the household of an HIV-positive index case. This is one potential explanation for the lower prevalence of *M.tb* infection in the child contacts of index cases with HIV infection irrespective of ART status found in this study. However, ART for prevention of mother to child transmission is widely used and the proportion of infected children will be very low.

Conclusion

We found a higher prevalence of *M.tb* infection among child contacts of HIV-negative tuberculosis patients compared to contacts of HIV-positive index cases, irrespective of ART status. We found no evidence that HIV-positive index cases on ART for ≥ 1 year at tuberculosis diagnosis were more likely to transmit than other HIV-positive index cases. Frequent health service contacts of HIV-positive individuals on ART leading to prompt diagnosis of TB may mitigate the effects of any increase in infectiousness. Further studies are required to definitively establish whether ART has a biological effect on the infectiousness of HIV-positive TB patients.

References

1. WHO. Global tuberculosis control: WHO report 2015. http://apps.who.int/iris/bitstream/10665/191102/1/9789241565059_eng.pdf?ua=1. Geneva: World Health Organisation; 2015.
2. Chaisson RE, Churchyard GJ. Recurrent tuberculosis: relapse, reinfection, and HIV. *J Infect Dis.* 2010;201(5):653-655.
3. Yates TA, Khan PY, Knight GM, et al. The transmission of Mycobacterium tuberculosis in high burden settings. *Lancet Infect Dis.* 2016;16(2):227-238.
4. Odhiambo JA, Borgdorff MW, Kiambih FM, et al. Tuberculosis and the HIV epidemic: increasing annual risk of tuberculous infection in Kenya, 1986-1996. *Am J Public Health.* 1999;89(7):1078-1082.
5. Egwaga SM, Cobelens FG, Muwinge H, Verhage C, Kalisvaart N, Borgdorff MW. The impact of the HIV epidemic on tuberculosis transmission in Tanzania. *AIDS.* 2006;20(6):915-921.
6. Middelkoop K, Bekker LG, Myer L, Dawson R, Wood R. Rates of tuberculosis transmission to children and adolescents in a community with a high prevalence of HIV infection among adults. *Clinical Infectious Diseases.* 2008;47(3):349-355.
7. Rieder HL. Editorial commentary: on the risk of being and becoming infected with Mycobacterium tuberculosis. *Clin Infect Dis.* 2008;47(3):356-357.
8. Suthar AB, Lawn SD, del Amo J, et al. Antiretroviral therapy for prevention of tuberculosis in adults with HIV: a systematic review and meta-analysis. *PLoS Med.* 2012;9(7):e1001270.
9. Lawn SD, Harries AD, Williams BG, et al. Antiretroviral therapy and the control of HIV-associated tuberculosis. Will ART do it? *Int J Tuberc Lung Dis.* 2011;15(5):571-581.
10. Gupta RK, Rice B, Brown AE, et al. Does antiretroviral therapy reduce HIV-associated tuberculosis incidence to background rates? A national observational cohort study from England, Wales, and Northern Ireland. *Lancet HIV.* 2015;2(6):e243-251.
11. Elliott AM, Hayes RJ, Halwiindi B, et al. The impact of HIV on infectiousness of pulmonary tuberculosis: a community study in Zambia. *AIDS.* 1993;7(7):981-987.
12. Cauthen GM, Dooley SW, Onorato IM, et al. Transmission of Mycobacterium tuberculosis from tuberculosis patients with HIV infection or AIDS. *Am J Epidemiol.* 1996;144(1):69-77.
13. Espinal MA, Perez EN, Baez J, et al. Infectiousness of Mycobacterium tuberculosis in HIV-1-infected patients with tuberculosis: a prospective study. *Lancet.* 2000;355(9200):275-280.
14. Carvalho AC, DeRiemer K, Nunes ZB, et al. Transmission of Mycobacterium tuberculosis to contacts of HIV-infected tuberculosis patients. *Am J Respir Crit Care Med.* 2001;164(12):2166-2171.

15. Kenyon TA, Creek T, Laserson K, et al. Risk factors for transmission of Mycobacterium tuberculosis from HIV-infected tuberculosis patients, Botswana. *Int J Tuberc Lung Dis.* 2002;6(10):843-850.
16. Huang CC, Tchetgen ET, Becerra MC, et al. The effect of HIV-related immunosuppression on the risk of tuberculosis transmission to household contacts. *Clin Infect Dis.* 2014;58(6):765-774.
17. Martinez L, Sekandi JN, Castellanos ME, Zalwango S, Whalen CC. Infectiousness of HIV Seropositive Tuberculosis Patients in a High-burden African Setting. *Am J Respir Crit Care Med.* 2016.
18. Munthali L, Khan PY, Mwaungulu NJ, et al. The effect of HIV and antiretroviral therapy on characteristics of pulmonary tuberculosis in northern Malawi: a cross-sectional study. *BMC Infect Dis.* 2014;14:107.
19. van Halsema CL, Fielding KL, Chihota VN, et al. Brief Report: The Effect of Antiretroviral Therapy and CD4 Count on Markers of Infectiousness in HIV-Associated Tuberculosis. *Journal of acquired immune deficiency syndromes (1999).* 2015;70(1):104-108.
20. Godfrey-Faussett P, Ayles H. Can we control tuberculosis in high HIV prevalence settings? *Tuberculosis (Edinb).* 2003;83(1-3):68-76.
21. Williams BG, Granich R, De Cock KM, Glaziou P, Sharma A, Dye C. Antiretroviral therapy for tuberculosis control in nine African countries. *Proc Natl Acad Sci U S A.* 2010;107(45):19485-19489.
22. Pretorius C, Menzies NA, Chindelevitch L, et al. The potential effects of changing HIV treatment policy on tuberculosis outcomes in South Africa: results from three tuberculosis-HIV transmission models. *AIDS.* 2014;28 Suppl 1:S25-34.
23. Dodd PJ, Knight GM, Lawn SD, Corbett EL, White RG. Predicting the long-term impact of antiretroviral therapy scale-up on population incidence of tuberculosis. *PLoS One.* 2013;8(9):e75466.
24. Zachariah R, Bemelmans M, Akesson A, et al. Reduced tuberculosis case notification associated with scaling up antiretroviral treatment in rural Malawi. *Int J Tuberc Lung Dis.* 2011;15(7):933-937.
25. Middelkoop K, Wood R, Bekker LG. The impact of antiretroviral treatment programs on tuberculosis notification rates. *Int J Tuberc Lung Dis.* 2011;15(12):1714; author reply 1714-1715.
26. Mboma SM, Houben RM, Glynn JR, et al. Control of (Multi)Drug Resistance and Tuberculosis Incidence over 23 Years in the Context of a Well-Supported Tuberculosis Programme in Rural Malawi. *PLoS One.* 2013;8(3):e58192.
27. Koole O, Houben RM, Mzembe T, et al. Improved retention of patients starting antiretroviral treatment in Karonga District, northern Malawi, 2005-2012. *Journal of acquired immune deficiency syndromes (1999).* 2014;67(1):e27-33.
28. Crampin AC, Glynn JR, Fine PE. What has Karonga taught us? Tuberculosis studied over three decades. *Int J Tuberc Lung Dis.* 2009;13(2):153-164.

29. Arnadottir T, Rieder HL, Trebucq A, Waaler HT. Guidelines for conducting tuberculin skin test surveys in high prevalence countries. *Tuber Lung Dis.* 1996;77 Suppl 1:1-19.
30. Khan PY, Glynn JR, Fielding KL, et al. Risk factors for Mycobacterium tuberculosis infection in 2-4 year olds in a rural HIV-prevalent setting. *Int J Tuberc Lung Dis.* 2016;20(3):342-349.
31. Ministry of Health Malawi. National Tuberculosis Control Programme Manual. 2007.
32. Cruciani M, Malena M, Bosco O, Gatti G, Serpelloni G. The impact of human immunodeficiency virus type 1 on infectiousness of tuberculosis: a meta-analysis. *Clin Infect Dis.* 2001;33(11):1922-1930.
33. Pai M, McCulloch M, Colford JM, Jr. Meta-analysis of the impact of HIV on the infectiousness of tuberculosis: methodological concerns. *Clin Infect Dis.* 2002;34(9):1285-1287.

Tables and figures

Table 1. Index case characteristics by HIV/ART status of index case (n=150)

Case characteristics		HIV/ART status of index case							
		HIV-		HIV+ ART ≥1 yr		HIV+ ART < 1yr		HIV+ no ART	
		(N=82)		(N=25)		(N=8)		(N=35)	
Male (%)		50	(61.7)	14	(56.0)	5	(62.5)	25	(71.4)
Age in years (mean (sd))		36.1	(14.7)	39.9	(6.1)	36.1	(4.8)	36.7	(8.7)
Index case mother (%)		15	(18.3)	7	(28.0)	2	(25.0)	5	(14.3)
Smear grade	Scanty	2	(2.5)	3	(12.0)	1	(12.5)	4	(11.4)
	+1	11	(13.4)	4	(16.0)	2	(25.0)	9	(25.7)
	+2	27	(32.9)	9	(36.0)	3	(37.5)	4	(11.4)
	+3	42	(51.2)	9	(36.0)	2	(25.0)	18	(51.5)
Duration of symptoms in weeks (median (IQR))		13.2	(7.6 – 20.6)	8.3	(7.0 – 16.3)	18.6	(8.2 – 28.6)	12.4	(8.0 – 19.7)

sd standard deviation; IQR interquartile range

Table 2. Demographic and clinical characteristics of index case and contact: risk factors for TST positivity (n=302)

Characteristics			TST≥10mm (N=302) n/N (%)		Crude OR (95% CI)	P value
Index Case	HIV/ART status	HIV+ no ART/ ART<1y	17/76	(22.4)	1	0.009
		HIV+ on ART ≥ 1y	12/56	(21.4)	1.1 (0.3 – 4.6)	
		HIV-	75/170	(44.1)	4.3 (1.4 – 13.1)	
	Sex	Female	52/116	(44.8)	1	0.01
		Male	52/186	(28.0)	0.3 (0.1 – 0.8)	
	Age (years)	<30	35/86	(41.9)	1	0.3
		30 - 44	53/156	(34.0)	0.6 (0.2 – 1.7)	
		≥45	15/60	(25.0)	0.3 (0.1 – 1.3)	
	Index case relationship	Not mother	70/243	(28.8)	1	0.002
		Mother	34/59	(57.6)	4.7 (1.7 – 13.0)	
	Smoker	No	72/207	(34.8)	1	0.8
		Yes	32/95	(33.7)	0.9 (0.3 – 2.5)	

Continued…

Table 2 (Continued). Demographic and clinical characteristics of index case and contact: risk factors for TST positivity (n=302)

Characteristics			TST≥10mm (N=302) n/N (%)		Crude OR (95% CI)	P value
	Sputum smear grade	Scanty	1/19	(5.3)	0.02 (0.001 – 0.5)	0.005
		+1	12/57	(21.1)	0.2 (0.06 – 0.8)	
		+2	35/92	(38.0)	0.7 (0.2 – 1.9)	
		+3	56/134	(41.8)	1	
	Duration of symptoms (weeks)	< 8	34/118	(28.8)	1	0.1
		≥ 8	70/184	(38.0)	2.1 (0.8 – 5.6)	
Contact	Sex	Female	52/148	(35.1)	1	0.6
		Male	52/154	(33.8)	0.8 (0.4 – 1.7)	
	BCG scar*	No	17/41	(41.5)	1	0.3
		Yes	82/248	(33.1)	0.6 (0.2 – 1.7)	
	Age (years)	2 - 3	30/75	(40.0)	1	0.7
		4 - 5	24/62	(38.7)	1.2 (0.4 – 3.3)	
		6 - 7	22/77	(28.6)	0.7 (0.3 – 1.7)	
		≥ 8	28/88	(31.8)	0.8 (0.3 – 1.9)	
	Degree of exposure to index case whilst symptomatic	Resident	44/175	(25.1)	1	0.008
		Sleep same room ≤ 30days	16/46	(34.8)	1.5 (0.5 – 4.3)	
		Sleep same room > 30days	42/81	(51.9)	4.0 (1.7 – 10.4)	
Household	No. of adults	1 - 2	42/101	(41.6)	1	0.3
		3 - 4	47/139	(33.8)	0.6 (0.2 – 1.7)	
		≥ 5	15/62	(24.2)	0.3 (0.1 – 1.3)	
	Socio-economic	Lowest score	54/148	(36.5)	1	0.08
		Middle	33/80	(41.3)	1.3 (0.4 – 3.9	
		Highest score	17/74	(23.0)	0.3 (0.09 – 1.1)	
	Quality of dwelling score	lowest	28/64	(43.8)	1	0.1
		middle	45/118	(38.1)	0.6 (0.2 – 2.1)	
		highest	31/117	(26.5)	0.3 (0.08 – 1.0)	
	Crowding (persons/room)	1 - 2	45/131	(34.4)	1	0.6
		3 - 4	45/138	(32.6)	1.0 (0.4 – 2.7)	
		≥ 5	14/33	(42.2)	2.3 (0.5 – 11.7)	

OR odds ratio; CI confidence interval; *missing BCG scar status on 13 child contacts

Table 3. Multivariable analysis of association of HIV/ART status of index case with TST positivity in child contacts

		TST positive (n=302)		Crude OR		Multivariable model 1[a] (n=302)		Multivariable model 2a[b] (n=302)		Multivariable model 2b[c] (n=302)	
		n/N	(%)	OR (95% CI)	P value	OR (95% CI)	P value	OR (95% CI)	P value	OR (95% CI)	P value
Index case HIV/ART status	HIV+ no ART/ ART<1y	17/76	(22.4)	1	0.009	1	0.05	1	0.06	1	0.2
	HIV+ on ART≥1y	12/56	(21.4)	1.1 (0.3 – 4.6)		1.0 (0.3 – 3.7)		1.1 (0.3 – 4.4)		0.9 (0.2 – 3.4)	
	HIV-	75/170	(44.1)	4.3 (1.4 – 13.1)		2.9 (1.0 – 8.2)		3.0 (1.0 – 8.7)		2.2 (0.8 – 6.3)	
Sensitivity analyses											
	HIV+ no ART	15/65	(23.1)	1	0.009	1	0.05	1	0.06	1	0.2
	HIV+ on ART	14/67	(20.9)	0.9 (0.2 – 3.5)		0.9 (0.3 – 3.3)		1.0 (0.3 – 3.7)		0.9 (0.2 – 3.3)	
	HIV-	75/170	(44.1)	3.8 (1.2 – 12.2)		2.8 (1.0 – 8.1)		2.9 (1.0 – 8.5)		2.2 (0.7 – 6.4)	
	HIV+ no ART	15/65	(23.1)	1	0.03	1	0.1	1	0.1	1	0.3
	HIV+ on ART<2y	8/41	(19.5)	0.9 (0.1 – 5.5)		0.8 (0.1 – 4.2)		0.8 (0.1 – 4.3)		0.6 (0.1 –3.5)	
	HIV+ on ART ≥2y	6/26	(23.1)	0.9 (0.2 – 4.3)		1.0 (0.2 – 4.8)		1.2 (0.3 – 5.9)		1.2 (0.2 – 5.6)	
	HIV-	75/170	(44.1)	3.8 (1.2 – 12.2)		2.8 (1.0 – 8.1)		2.9 (1.0 – 8.5)		2.2 (0.7 – 6.5)	
TST≥15mm											
	HIV+ no ART/ ART<1y	12/76	(15.8)	1	0.008	1	0.08	1	0.1	1	0.2
	HIV+ on ART ≥ 1y	7/56	(12.5)	0.6 (0.1 – 3.5)		0.6 (0.1 – 2.8)		0.7 (0.1 – 3.4)		0.5 (0.1 – 2.4)	
	HIV-	58/170	(34.1)	4.2 (1.2 – 15.0)		2.3 (0.7 – 7.5)		2.4 (0.7 – 8.1)		1.7 (0.5 – 5.5)	

[a] Adjusted for household clustering; age and sex of index case; age and sex of contact; degree of exposure of contact to case; household SES; quality of dwelling structure; whether index case was the mother; and number of adults in the household

[b] Adjusted for household clustering and for all risk factors in the model 1 plus duration of symptoms

[c] Adjusted for household clustering and for all risk factors in the model 1 plus duration of symptoms and smear grade

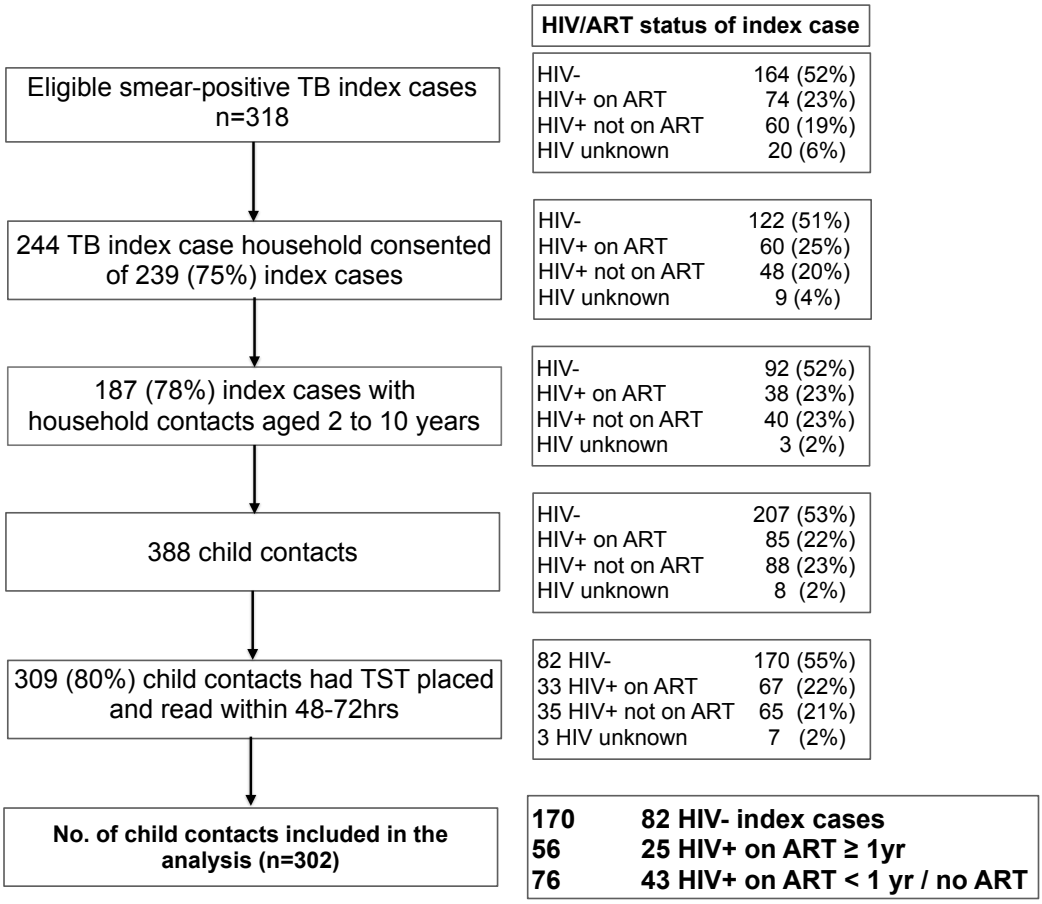


Figure 1. Study flowchart: TB case through to child contact TST data included in analysis

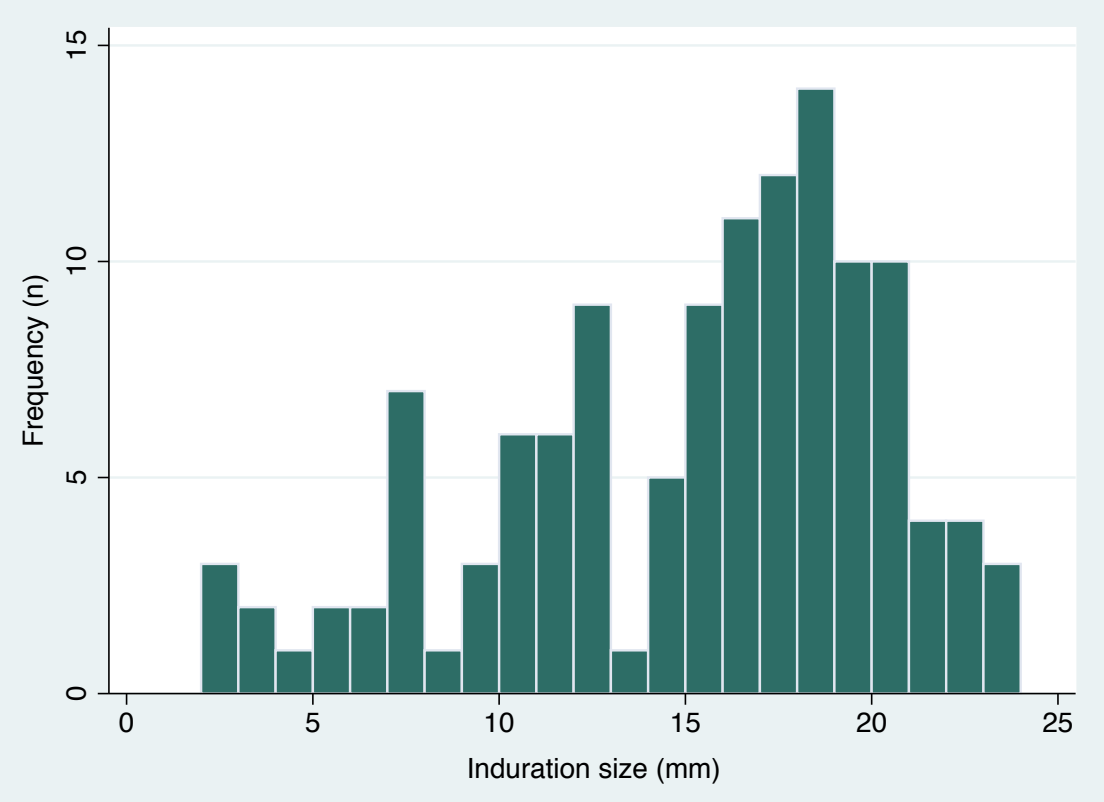


Figure 2. Histogram illustrating the frequency distribution of non-zero induration in child contacts (n=125)

Competing interests

There are no conflicts of interest (including financial and other relationships) for the authors.

Acknowledgements

All authors have seen and approved the manuscript, and all have contributed significantly to the work. PYK conceived the idea of undertaking the analysis with support from KK. KLF directed the final analytic strategy. The original studies which provided the data for this analysis were designed by ACC and JRG. PYK directed the implementation of the original studies with OK and TM. OK supervised field activities with the assistance of TM and supervised data collection and data entry. PYK wrote the first draft of manuscript. ACC, KLF, KK, OK and JRG edited the subsequent drafts.

Funding: This work was supported by a Wellcome Trust Strategic Award for the Karonga Prevention Study [grant number 098610/Z/12/Z] and a Wellcome Trust clinical research training fellowship [PYK: grant number 100137/Z/12/Z].

Data presented in part at the Conference on Retroviruses and Opportunistic Infections (**CROI**) 2016 in Boston, 22-25 February 2016 (Young Investigator Scholarship)

7 Discussion

7.1 Introduction

Tuberculosis (TB) is estimated to have caused more deaths than any other infectious disease in history.[1] Since 1993, when TB was declared a global health emergency, we have seen the successful roll out of the DOTS strategy which, despite being one of the most cost-effective public health interventions to date,[2,3] was unable to mitigate the impact of the human immunodeficiency virus (HIV) epidemic on TB control.[4] The ensuing Stop TB strategy, in conjunction with the concurrent roll-out of anti-retroviral treatment (ART) across sub-Saharan Africa did, however, result in achieving the Millennium Development Goal target "to halt and begin to reverse the incidence of tuberculosis by 2015".[5] Despite a concerted global effort which has undoubtedly saved millions of lives,[6] 80 years after Wade Hampton Frost predicted the eventual eradication of TB, we are far off target for elimination[ii] by 2050 and nowhere near eradication[iii]. The aim of this research was to address some of the knowledge gaps and gain a better understanding of *M.tb* transmission dynamics, especially in high HIV burden settings, to inform how best to interrupt transmission thereby accelerating the decline in TB incidence required to achieve the goal of TB elimination and a "world free of TB".[5]

7.2 Key research findings

Key research findings are summarised in Table 1 below.

[ii] Elimination is defined as achieving an incidence of less than 1 case of all forms of TB per million population

[iii] Eradication defined as the permanent reduction to zero of the worldwide incidence of the infection and interventions measures are no longer required[7]

Table 1. Summary of key research findings

Area of research	Key Findings	Comments
Measuring M.tb transmission at a population-level	*M.tb infection prevalence estimates:* Rust and Thomas was the only method which adjusted for BCG-attributable induration in the very young, most recently vaccinated children	For the Rust and Thomas method, data need to be collected at scale and in household TB contacts within the same population as non-contacts (Chapter 3) Logistically more complicated to execute than a simple tuberculin survey
	Estimates of risk of M.tb infection: Marked differences in estimates of the risk of infection when using cross-sectional data versus longitudinal data (0.3% vs 1.6% per annum respectively)	Underestimates of *M.tb* infection risk using cross-sectional data are discussed in Chapter 5. A cross-sectional analysis of the 2nd round of TST data using TST≥15mm cut-off method gives an *M.tb* prevalence estimate of 3.5% compared to 1.1% found at the baseline survey [See Appendix III].
	Age: Risk of *M.tb* infection is dependent on age, even in the under-5s; approximately 1.5-fold increase in the risk of incident *M.tb* infection for each one year increase in age	This finding is remarkably consistent when using increasingly more specific cut-offs to define TST conversion. [See Appendix IV: Table A2]
Drivers and locations of community M.tb transmission	*Household smear-positive TB contact:* - Associated with a 5-fold increase in the odds of prevalent *M.tb* infection; PAF 2% - Associated with a 7-fold increase in the risk of incident *M.tb* infection; PAF 2%	PAF of household TB contact for prevalent *M.tb* infection: 15% if based on data from TB household study [See Appendix V]. This estimate is similar to the PAF estimates from other studies conducted in Karonga and much higher than the 2% estimated from the baseline TST survey and the TST cohort study.
	Neighbourhood smear-positive TB contact: (defined as resident within 200m of an infectious TB case excluding household TB contact)*:* - Associated with a 2-fold increase in odds of prevalent *M.tb* infection; PAF 17%	No evidence of an association of neighbourhood TB contact (excluding household TB contact) with risk of incident *M.tb* infection

Continued…

Table 1 (Continued). Summary of key research findings

Area of research	Key Findings	Comments
Drivers and locations of community M.tb transmission	*Community M.tb exposure:* (defined as the av. smear-positive TB notification rate of >30 per 100,000 pop per year): - Associated with 2-fold increase in the odds of prevalent *M.tb* infection; PAF 39% - Associated with a 2-fold increase in risk of incident *M.tb* infection; PAF 29%	PAFs for community *M.tb* exposure are relatively similar for prevalent and incident *M.tb* infection. In the multivariable model for both prevalent *M.tb* infection and incident *M.tb* infection, household and neighbourhood infectious TB contact were not included in the model due to collinearity. However, including household and neighbourhood TB contact in the model did not alter the effect estimate of community *M.tb* exposure. [See main text below (Section 7.3)]
	HIV infection in the family: - HIV exposure *in utero* associated with nearly 4-fold increase in odds of prevalent *M.tb* infection; PAF 6% - HIV-positive father associated with a 2.5-fold increase in risk of incident *M.tb* infection; PAF 7%	[See main text below (Section 7.3)]
	Church attendance: - Associated with a nearly 4-fold increase in risk of incident *M.tb* infection; PAF 68%	Highlight that poorly-ventilated crowded indoor spaces pose substantial risk of onward *M.tb* transmission, even in settings with a low-to-moderate TB burden. Although inference is limited due to the wide confidence interval around the PAF for church attendance which includes 0.
	Travel on mini-buses: - Associated with a nearly 2-fold increase in risk of incident infection; PAF 27%	
Effect of HIV and ART on M.tb transmission dynamics	*Difference in infectiousness:* HIV-positive smear-positive TB cases are less infectious than HIV-negative smear-positive TB cases; no evidence that HIV-positive on long-term ART are more likely to transmit *M.tb* infection than HIV-positive TB cases not on ART	Likely to be dependent on biological factors, such as degree of immunosuppression of HIV+ index TB cases at a population-level (which is in turn dependent on the maturity of HIV epidemic), contextual factors, such as the stage at which HIV is being diagnosed (early vs late) and quality and amount of TB screening of HIV+ individuals at a health systems-level and behavioural, such as health-seeking behaviour and how good ART adherence is) Need further studies to definitively establish whether ART has a biological effect on the infectiousness of HIV+ TB patients

BCG bacille Calmette-Guerin; TST tuberculin skin test; ARTI average annual risk of *M.tb* infection; CI confidence interval; PAF population attributable fraction; ART antiretroviral treatment

7.3 Interpretation of key findings

Measures of population-level M.tb transmission using tuberculin data

The challenges of using tuberculin data to measure population-level *M.tb* transmission have been discussed in Chapter 3 with regards to estimating the average annual risk of *M.tb* infection (ARTI) derived from *M.tb* prevalence estimates, and in Chapter 5 with regards to estimating the incidence of *M.tb* infection using serial TSTs.

An additional analysis (Appendix III) was undertaken as an internal validity check of the data to estimate the prevalence of *M.tb* infection at the 2nd round of skin testing. The eligibility criteria for this analysis was as per the baseline TST survey presented in Chapter 3 (i.e. restricted to children aged between 2.0 to 4.9 years at the time of the 2nd TST) but these children had undergone two TSTs as opposed to one. ARTI derived from the 2nd round was significantly higher (ARTI 1.0%: 95% CI 0.8 – 1.2%) than the estimate derived from the baseline round (Chapter 4: 0.3%; 95% CI 0.2 – 0.4%), and closer to the estimate of the annual incidence risk estimated from the longitudinal analysis (1.6%; 95% CI 1.3 – 2.0: Chapter 5). Thus leading one to postulate whether 'boosting' from the 1st TST 'unmasked' the waned immunity from remote *M.tb* infection that was not apparent at the 1st round in some children, which then contributed to the ARTI estimate derived from the 2nd round. However, it is not possible to make any such conclusions based on the data available.

Even if tuberculin surveys at a single point in time do underestimate the ARTI within a population, repeat TST surveys in the same age group of children over time, will probably still allow estimates of the trend in force of infection to be gauged.[8,9] That is as long as the method used to estimate the prevalence of *M.tb* infection is able to appropriately account for cross-reactions due to BCG and/or NTM infections, even as the prevalence of *M.tb* infection decreases over time.[10] With regards to longitudinal studies, which involve serial skin testing to estimate the incidence of *M.tb* infection, the difficulty remains in disentangling what is a true conversion (new incident *M.tb* infection), boosting of remote *M.tb* infection (which occurred prior to the start of the study and therefore should not contribute to the estimate of incidence) and boosting of previous BCG vaccination and/or sensitisation to NTM.

Community vs household transmission in an HIV prevalent setting

The proportion of prevalent and incident *M.tb* infections attributable to household infectious TB contact in this study population was similar, at approximately 2% using data from the DSS area alone. However, the PAF of household TB contact for prevalent *M.tb* infection, which was based on the ratio of the prevalence of *M.tb* infection in household contacts of smear-positive TB cases in the TB case-contact household study (Chapter 3 Table 2: 'Higher risk' group aged ≥ 2 years) to the prevalence of *M.tb* infection in those not known to have had any household contact with an infectious TB case from the baseline survey (Chapter 4 Table 1) was more than 7-fold higher than the PAF presented in Chapter 4 (15% vs 2% respectively: see Appendix V). A PAF of 15% for

household/close contact is nearer to the estimates found in other studies conducted in Karonga district.

A molecular study of epidemiologically linked cases found that only 9-13% of tuberculosis cases were estimated to be attributable to recent transmission from identifiable close contacts,[11] and a more recently published molecular study using whole genome sequencing (WGS) found that the proportion of TB attributable to known contacts was estimated to be 9.4% overall.[12] Similar results were demonstrated in another study using a case-control study design which estimated that identifiable recent contact with known smear-positive cases accounted for 12.5% of the burden of disease in the district.[13] A cross-sectional household case-contact study conducted in the district also found that an estimated 20% of *M.tb* infection in children under 10 years was attributable to smear-positive household contact.[14]

The higher PAF observed is primarily driven by the higher absolute risk in the household contacts of an infectious TB case in the TB case-contact household study, compared to the absolute risk derived from the baseline TST survey. One reason for this difference in risk may be because the TB household study, and the other studies mentioned above, were all conducted district-wide and not confined to the DSS. TB screening of all adults is undertaken as part of the annual re-census in the DSS, which consists of a simple symptom screen and if symptomatic, individuals are asked to submit a sputum sample for smear microscopy. More generally, research has been conducted in the DSS for the last 12 years, which may have impacted on the behaviour of the population with

regards to health-seeking behaviour- a form of 'Hawthorne effect'.[15] This may be leading to earlier case detection of smear-positive TB cases in the DSS and a shorter duration of infectiousness, resulting in a lower risk of intra-household transmission to children compared to smear-positive TB cases in the district outside of the DSS, where the majority of TB cases are diagnosed through passive case-finding.

Although household/neighbourhood contacts of a known infectious TB case are at high risk for *M.tb* infection, the majority of *M.tb* transmission to children appears to be occurring through casual contact in the community or from contact with undiagnosed TB cases within the family (including extended family members and close non-familial contacts), even in this low-to-moderate burden setting. The PAF of neighbourhood contact (defined as being resident within 200m of a diagnosed smear-positive TB case excluding known household contact) was only 17% for prevalent *M.tb* infection, and there was no evidence for an association of neighbourhood contact with incident *M.tb* infection. This may have been due to the fact that all diagnosed smear-positive TB cases during the lifetime of the child were included for the analysis of prevalent *M.tb* infection whereas only those individuals diagnosed with smear-positive TB from 2013 to 2015 were included in the *M.tb* infection incidence analysis, to assess the association of *M.tb* infection incidence with concurrent (rather than historical) TB in neighbourhood, thereby reducing the power to identify an association with neighbourhood contact.

Approximately 40% of prevalent *M.tb* infections and 30% of incident *M.tb* infections in young children were attributable to living in an area with a higher community *M.tb* exposure (defined as the average smear-positive TB notification rate per 100,000 population per year per residential area). Areas with a higher TB notification rate in the DSS may be a marker of communities where there is a higher burden of undiagnosed TB. If the higher average notification rates of smear-positive TB were due to better case detection in these residential areas then one would expect that duration of infectiousness of TB cases would be shorter and therefore there would be a decreased, rather than an increased risk of *M.tb* infection to children. Together these findings *suggest* that the majority of *M.tb* transmission to young children is occurring from close contacts with undiagnosed infectious TB and/or unidentifiable casual contacts with (known or unknown infectious TB) in the community. The fact that age is so strongly associated with an increasing risk of incident *M.tb* (as children become more mobile), and that 68% of incident *M.tb* infections *may* be attributable to church attendance (bearing in mind the large confidence interval around the PAF) and approximately 28% are attributable to travel on mini-buses, all point to the interpretation that unidentifiable casual contacts with (known or unknown infectious TB) is playing a role in driving community *M.tb* transmission to pre-school children.

HIV infection in the family

HIV infection in the family, although not accounting for the majority of *M.tb* infections in children in this setting, was associated with a significantly increased risk of *M.tb* infection through a number of mechanisms. As discussed in Chapter

5, maternal HIV infection was not associated with incident *M.tb* infection, although it was strongly associated with prevalent *M.tb* infection. This was most likely because children who had a TST≥10mm at baseline (prevalent *M.tb* infection) were not eligible for a repeat skin test and were thus excluded from the cohort analysis.

An exploratory analysis of population-level *M.tb* infection prevalence was undertaken, stratified by *in utero* HIV-exposure status in the DSS [see Appendix VI]. Using cut-off methods and mixture analysis, the prevalence was similar in the HIV-exposed and un-exposed children, whereas using the Rust and Thomas method, the prevalence of *M.tb* infection was over 4-fold higher in HIV-exposed children (2.5%) compared to HIV-unexposed children (0.6%). However, this observed difference in prevalence estimates is primarily driven by very small numbers in the *nth* category (TST≥20mm): 1/264 in the HIV-exposed versus 6/4461 in the unexposed children. Hence this exploratory analysis was not included in the paper on estimating the *M.tb* prevalence using the Rust and Thomas method which was published in *American Journal of Epidemiology* (Chapter 3). Notwithstanding this limitation, it does triangulate with the finding that HIV-exposed children in the DSS had a 4-fold increase in the odds of prevalent *M.tb* infection compared to HIV-unexposed children in the risk factor analysis of baseline survey (Chapter 4).

An increased risk of *M.tb* infection among HIV-exposed children has been shown in a number of other studies across sub-Saharan Africa.[16-21] Increased *M.tb* exposure in the home may underlie this association, as HIV-exposed children

even if not infected with HIV may have a higher risk of acquiring *M.tb* infection from undiagnosed TB in their mothers (or fathers) than children whose parents are HIV-negative. They may be at a greater risk of nosocomial *M.tb* exposure from accompanying HIV-positive mothers to clinic and they may also be more likely to acquire *M.tb* infection following *M.tb* exposure due to alterations in innate and adaptive immunity that affect susceptibility to infection.[16,22]

HIV infection in the father was associated with a 2.5-fold increase in risk of *M.tb* infection in the longitudinal analysis (Chapter 5). This increased risk of incident infection observed in young children may be due to undiagnosed TB in their fathers. This potentially sheds some light as to why the more holistic TB-HIV intervention at the household level of newly diagnosed TB patients, compared to enhanced case-finding in the community in the ZAMSTAR study (a cluster-randomised trial of interventions to reduce *M.tb* transmission in high HIV prevalence settings) had an effect on *M.tb* transmission at community level.[23] This was a somewhat surprising finding at the time, as contact investigation of known TB cases has always been deemed low priority in low-income settings due to it being resource-intensive to conduct, and had been thought to have limited impact as a control strategy to reduce community *Mtb* transmission,[24] as most secondary cases arise outside identified contacts in high-incidence areas. However, it is possible that through the household intervention in ZAMSTAR, more men with undiagnosed TB (with and without HIV infection) were screened for TB, than would have been accessed through the enhanced case finding which required men to actively seek a community sputum collection point, thereby

reducing the duration of infectiousness in the communities randomised to the household intervention arm.

Differences in infectiousness by HIV/ART status

Although index TB cases with HIV infection (irrespective of ART status) have a lower risk of transmission within the household, as inferred from TST positivity in child contacts in this setting (Chapter 6), on a population-level, HIV positive TB cases may still play a substantial role in onward *M.tb* transmission.

A large-scale WGS study of 72% of all culture-positive TB episodes in Karonga district from 1995 to 2010 found no association in transmissibility with HIV infection.[25] One reason cited was that this may have been due to the social clustering of HIV-positive individuals, which increases the opportunities for transmission to susceptible individuals who manifest disease, balancing out any decreased transmissibility.[25] The relative risk (RR) of having transmissions confirmed using WGS from HIV-positive compared to HIV-negative smear-positive named contacts was 0.78 (95% CI 0.54 – 1.1).[26] Approximately 60% of smear-positive TB patients are HIV-positive, so if transmission were reduced by 22%, HIV-associated TB would still account for 47% of all transmission resulting in disease in Karonga.

Much quoted epidemiological studies undertaken in the mid-2000s in South Africa and Zimbabwe, found that despite a huge difference in the TB disease incidence between HIV-negative and HIV-positive individuals, the point

prevalence of active TB disease (the main driver of onward *M.tb* transmission) differed little by HIV status.[27,28] Thus it was surmised that the rapid progression and lower per-case infectivity of HIV-associated TB, resulted in maintaining a static *M.tb* transmission rate, despite the large increase in TB incidence observed.[27-29] However, some of these findings may have been specific to the context in which they were conducted and not generalisable, as contrary results have been documented in other studies. For example, stable age-specific incidence rates were observed in HIV-negative gold miners in a single workforce in the Free State Province, South Africa,[29] whereas a retrospective cohort among men working in four mines in the Gauteng Province, South Africa conducted over the same time period found that incidence of new pulmonary TB in HIV-negative men doubled, reflecting increased *M.tb* transmission, which had occurred despite regular TB screening and readily available treatment administered under direct supervision.[30] Other conflicting results include estimates around duration of infectiousness stratified by HIV status. A TB prevalence survey conducted by Wood *et al.* in Cape Town in 2005, estimated that the duration of infectiousness was similar in HIV-positive and HIV-negative individuals, and that 87% of the total person-years of undiagnosed smear-positive TB in the community were among HIV-positive individuals,[31] which differs from the conclusions drawn from the studies referred to above,[27,28] and from a TB prevalence survey conducted in Harare, Zimbabwe in 2005-2006, which estimated that duration of infectiousness was much shorter in HIV-positive individuals.[27,28,32]

Conflicting results from studies assessing the impact of HIV on *M.tb* transmission at a population-level done nearly 10-20 years ago have also not been resolved.[33-

[37] Trends in the observed ARTI, as measured by tuberculin surveys in school children, differed by country. In Kenya, a downward trend in ARTI in the pre-HIV era was reversed in areas where there was a higher HIV prevalence,[33] whereas in Tanzania the ARTI continued to fall as the HIV epidemic took off.[34,35] However, this sustained downward trend in ARTI observed in Tanzania coincided with significant strengthening of the TB control programme in the country.[35] In Cape Town, South Africa, the ARTI, although extremely high at ~4%, was found to be consistent across age groups between the ages of 5 to 17 years, leading to the conclusion that HIV-associated TB was not having a major impact on the ARTI.[36] The validity of this conclusion was questioned, because ARTI as an indicator of the risk of infection from a single survey, ignores the effects of cohort and age. Thus, despite what may be a true increase in the risk of infection over time, the increase in the risk of infection in the younger children may be masked due to the levelling effect of the lower risk in older children from times gone by.[37]

Interestingly, the results of an exploratory ecological spatial analysis in the Karonga DSS, using a distance-based mapping approach in R,[38] undertaken by the candidate identified a 'hot spot' of incident *M.tb* infection in children in the same localised geographical area as a larger 'hot spot' of higher than expected adult HIV prevalence for the population (unpublished data: Appendix VII). However, it is not clear whether clustering of *M.tb* infection, which is manifest as spatial heterogeneity as seen with TB disease incidence in other settings,[11,39-41] is occurring as result of localised 'hot spots' of *M.tb* transmission due to undiagnosed infectious cases or reflects a concentration of people with shared risk factors for susceptibility to infection and/or progression to disease.[9]

7.4 Limitations

The limitations of each of the studies have been discussed in the appropriate discussion sections in each paper. However, one limitation which has not been addressed is that this research has not examined the role of re-infection in *M.tb* transmission dynamics, which is likely to have important implications for the development of TB control strategies depending on the epidemiological context.[42,43] Also there is minimal multidrug-resistant TB in this area (<1%, even in retreatment cases),[44] most likely as a result of the well-implemented TB control programme, and therefore findings from this research may not be generalisable to settings where there are poorly functioning TB control programmes and a higher burden of drug-resistant TB.

7.5 Recommendations

Improve surveillance to enable context-specific targeting of control strategies

The availability of socio-demographic and geo-positional data of diagnosed TB cases in this setting enabled elucidation of risk factors driving *M.tb* transmission at a community-level. Collecting more detailed programmatic data would allow the identification of factors driving the epidemic at a more local level,[45] although this will require building up surveillance infrastructure which is currently lacking in most high burden settings. 'Know your epidemic', which was originally coined by the Joint United Nations Programme on HIV/AIDS (UNAIDS) as its mantra for HIV prevention,[46] needs to be urgently adopted by the TB control community. Targeted TB control strategies specific to the local epidemiological context are much more likely to be effective than a 'one size fits all' approach.

Active case-finding in the household as a component of integrated HIV/TB care in high HIV prevalence settings, may be more effective in interrupting the *M.tb* transmission cycle,[47] whereas in other settings, for example, where the prevalence of HIV is not so high such as Peru or Brazil, targeting certain 'hotspots' of disease incidence may be more successful.[38,48]

It may be that in Karonga district, a targeted approach focussing on the households of newly diagnosed TB patients, irrespective of HIV status, and HIV-positive patients starting ART may be most effective. A proposed strategy would involve a similar intervention to that used in ZAMSTAR, which would include HIV testing and TB screening for all members of the household. Isoniazid preventive therapy (IPT) would not be restricted to young children but would include older children and adolescents irrespective of HIV status,[49] and all HIV-positive individuals would be started on ART and IPT.[50,51]

Children with M.tb infection as targets of control strategies

Children are the seedbeds of future epidemics, and as *M.tb* infection can be lifelong,[52] cases will continue to emerge over time. Although in our setting only 4% of children were born to an HIV-positive mother; each year worldwide over a million infants are born to HIV-positive mothers[20], and HIV-exposed uninfected children now account for as many as 30% of all births in parts of southern Africa.[21] This growing population, although not infected with HIV, may be more vulnerable to infection following *M.tb* exposure,[19,22] and TB compared to HIV-unexposed children.[53] This is especially worrying in some high HIV prevalence settings, such as South Africa where there is ongoing transmission of drug-

resistant *M.tb,* as there may be a rapidly growing reservoir of drug-resistant *M.tb* infection waiting to reactivate in the coming years.

Better infection control practices in congregate settings

There is no doubt that better infection control practices are required in congregate settings, such as churches and public transport. Simple infection control practices, such as opening windows or even holding congregations in outdoor spaces may go some way to mitigating the risk of *M.tb* transmission. These have been highlighted in Chapter 2 and Chapter 5.

7.6 Future work

As stated by Rieder in the Epidemiological Basis of Tuberculosis Control:[52]

> "…tuberculous infection takes center stage for providing an understanding of the dynamics of the tuberculosis epidemic in the community."

M.tb infection incidence and prevalence are the primary determinants of current and future TB disease incidence,[54] so accurately quantifying these metrics is essential for future planning and to assess the effectiveness of control interventions to reduce *M.tb* transmission. But, as highlighted by this research, accurate measurements of these parameters at a population-level remain problematic.

Conducting robust TB epidemiological studies is resource-intensive and expensive and collecting epidemiological data at scale will be even less possible

in budget-constrained times. Thus, we are somewhat dependent on models to predict the future burden and effectiveness of control interventions in different geographical areas, over time and at scale. These models are entirely dependent on a number of priority parameters; *M.tb* infection risk being one of the most integral to improving the predictive ability of virtually all TB models.[55] Further research is required to attain a robust proxy measure for *M.tb* transmission at a population-level.

Despite the improved specificity of interferon gamma release assays (IGRA) over TST,[56] skin testing remains the only practical method of measuring infection in populations, especially in low-income settings with limited laboratory infrastructure.[57] Through a combination of newer statistical techniques such as latent variable modelling, systematic calibration studies of TST, IGRAs,[58-60] and the newer skin tests, such as C-Tb skin test (a novel skin test containing *M.tb*-specific antigens, ESAT-6 and CFP-10),[61] in a variety of populations (infant, child, adolescent, adult) and geographical settings with high and low HIV/TB burden, it may be possible to derive a method to quantify a robust proxy measure of *M.tb* transmission.

We also need a better understanding of what is actually being measured by TST/IGRAs. A recent study by Andrews *et al.* found that interferon-gamma values showed a bimodal distribution in infants, which was not seen in adolescents in the same area in Western Province.[62] Such findings raise the questions of whether we are actually measuring *M.tb* exposure, or susceptibility to infection following *M.tb* exposure, and how to distinguish the two? Improved

diagnostics of *M.tb* infection status would go a long way in helping to understand the natural history of *M.tb* exposure through to established *M.tb* infection.

Studies which were undertaken in the pre-ART era need to be repeated in the post ART roll-out era as it likely that the epidemiology of HIV-associated TB has changed. In HIV-positive individuals, ART reduces TB incidence across all CD4 cell counts,[63] yet despite long-term ART, TB incidence remains higher than in HIV-negative people in both high and low TB-burden settings.[50,64] As life expectancy is greatly extended by ART, the cumulative lifetime risk of TB for HIV-positive people remains very high,[64] and ART may even increase the infectiousness of TB by shifting the clinical manifestations to be more similar to that in HIV-negative patients.[65,66] We need to re-assess the population-level impact of HIV on *M.tb* transmission, and gain a better understanding of the effect of ART on *M.tb* transmission dynamics. More data are required to address the issue of conflicting evidence on TB disease duration ratio between HIV-positive and HIV-negative individuals,[67] and this is likely to have changed in the ART era. There has been also accumulating evidence from prevalence surveys that in some settings there may be a large hidden burden of asymptomatic subclinical TB cases,[68,69] even in people with HIV infection,[28,31,70,71] thus raising the question as to whether passive diagnosis is indeed enough in some settings to prevent onward *M.tb* transmission.[72] However, the evidence in support of community-wide active-case finding as a means to reduce *M.tb* transmission at a population-level is scarce,[73,74] therefore reiterating the point that greater insights into the *M.tb* transmission dynamics in populations with moderate-to-high HIV

prevalence are required before advocating the widespread implementation of resource-expensive control strategies.

7.7 Concluding remarks

A hallmark of TB epidemiology is its variability among populations.[75] *M.tb* and its host population undergo continuous change over time which mutually affect each other.[54] Factors driving *M.tb* transmission in one setting may not be relevant in another; and what was seen historically may no longer be relevant. Understanding the TB epidemic at a local level through continual surveillance allows the implementation of the most effective context-specific control interventions. Only then will we accelerate the decline in TB incidence towards the goal of TB elimination.

7.8 References

1. Paulson T. Epidemiology: A mortal foe. *Nature.* 2013;502(7470):S2-3.
2. Murray CJ, DeJonghe E, Chum HJ, Nyangulu DS, Salomao A, Styblo K. Cost effectiveness of chemotherapy for pulmonary tuberculosis in three sub-Saharan African countries. *Lancet.* 1991;338(8778):1305-1308.
3. Baltussen R, Floyd K, Dye C. Cost effectiveness analysis of strategies for tuberculosis control in developing countries. *BMJ.* 2005;331(7529):1364.
4. De Cock KM, Chaisson RE. Will DOTS do it? A reappraisal of tuberculosis control in countries with high rates of HIV infection. *Int J Tuberc Lung Dis.* 1999;3(6):457-465.
5. WHO. End TB Strategy. Draft global strategy and targets for tuberculosis prevention, care and control after 2015. *Documentation for World Health Assembly 67.* http://apps.who.int/gb/ebwha/pdf_files/WHA67/A67_11-en.pdf. Geneva: World Health Organisation; 2014.
6. Lienhardt C, Glaziou P, Uplekar M, Lonnroth K, Getahun H, Raviglione M. Global tuberculosis control: lessons learnt and future prospects. *Nature reviews Microbiology.* 2012;10(6):407-416.
7. Dowdle WR. The principles of disease elimination and eradication. *Bull World Health Organ.* 1998;76 Suppl 2:22-25.
8. Borgdorff M. Annual risk of infection- time for an update? *Bull World Health Organ.* 2002;80(6):501-502.
9. Yates TA, Khan PY, Knight GM, et al. The transmission of Mycobacterium tuberculosis in high burden settings. *Lancet Infect Dis.* 2016;16(2):227-238.
10. Khan PY, Glynn JR, Mzembe T, et al. Challenges in the estimation of the annual risk of Mycobacterium tuberculosis infection in children aged under 5 years. *Am J Epidemiol.* 2017.
11. Crampin AC, Glynn JR, Traore H, et al. Tuberculosis transmission attributable to close contacts and HIV status, Malawi. *Emerg Infect Dis.* 2006;12(5):729-735.
12. Glynn JR, Guerra-Assuncao JA, Houben RM, et al. Whole Genome Sequencing Shows a Low Proportion of Tuberculosis Disease Is Attributable to Known Close Contacts in Rural Malawi. *PLoS One.* 2015;10(7):e0132840.
13. Crampin AC, Floyd S, Ngwira BM, et al. Assessment and evaluation of contact as a risk factor for tuberculosis in rural Africa. *Int J Tuberc Lung Dis.* 2008;12(6):612-618.
14. Ho T. *The risk of household transmission of Mycobacterium tuberculosis infection to children under 11 years in rural Malawi*. London: Department of Infectious Disease Epidemiology, London School of Hygiene and Tropical Medicine; 2012.

15. Sedgwick P, Greenwood N. Understanding the Hawthorne effect. *BMJ.* 2015;351:h4672.

16. Cotton MF, Schaaf HS, Lottering G, et al. Tuberculosis exposure in HIV-exposed infants in a high-prevalence setting. *Int J Tuberc Lung Dis.* 2008;12(2):225-227.

17. Cotton MF, Slogrove A, Rabie H. Infections in HIV-exposed uninfected children with focus on sub-Saharan Africa. *Pediatr Infect Dis J.* 2014;33(10):1085-1086.

18. Cranmer LM, Kanyugo M, Jonnalagadda SR, et al. High prevalence of tuberculosis infection in HIV-1 exposed Kenyan infants. *Pediatr Infect Dis J.* 2014;33(4):401-406.

19. Marquez C, Chamie G, Achan J, et al. Tuberculosis Infection in Early Childhood and the Association with HIV-exposure in HIV-uninfected Children in Rural Uganda. *Pediatr Infect Dis J.* 2016;35(5):524-529.

20. Sugandhi N, Rodrigues J, Kim M, et al. HIV-exposed infants: rethinking care for a lifelong condition. *AIDS.* 2013;27 Suppl 2:S187-195.

21. Shapiro RL, Lockman S. Mortality among HIV-exposed infants: the first and final frontier. *Clin Infect Dis.* 2010;50(3):445-447.

22. Afran L, Garcia Knight M, Nduati E, Urban BC, Heyderman RS, Rowland-Jones SL. HIV-exposed uninfected children: a growing population with a vulnerable immune system? *Clin Exp Immunol.* 2014;176(1):11-22.

23. Ayles H, Muyoyeta M, Du Toit E, et al. Effect of household and community interventions on the burden of tuberculosis in southern Africa: the ZAMSTAR community-randomised trial. *Lancet.* 2013;382(9899):1183-1194.

24. Morrison J, Pai M, Hopewell PC. Tuberculosis and latent tuberculosis infection in close contacts of people with pulmonary tuberculosis in low-income and middle-income countries: a systematic review and meta-analysis. *Lancet Infect Dis.* 2008;8(6):359-368.

25. Guerra-Assuncao J, Crampin A, Houben R, et al. Large-scale whole genome sequencing of provides insights into transmission in a high prevalence area. *eLife.* 2015;4.

26. Glynn J. Does HIV infection reduce the probability of transmission of pulmonary tuberculosis? Paper presented at: CROI 2015; Seattle, USA.

27. Corbett EL, Charalambous S, Moloi VM, et al. Human immunodeficiency virus and the prevalence of undiagnosed tuberculosis in African gold miners. *Am J Respir Crit Care Med.* 2004;170(6):673-679.

28. Corbett EL, Bandason T, Cheung YB, et al. Epidemiology of tuberculosis in a high HIV prevalence population provided with enhanced diagnosis of symptomatic disease. *PLoS Med.* 2007;4(1):e22.

29. Corbett EL, Charalambous S, Fielding K, et al. Stable incidence rates of tuberculosis (TB) among human immunodeficiency virus (HIV)-negative

South African gold miners during a decade of epidemic HIV-associated TB. *J Infect Dis.* 2003;188(8):1156-1163.

30. Sonnenberg P, Glynn JR, Fielding K, Murray J, Godfrey-Faussett P, Shearer S. HIV and pulmonary tuberculosis: the impact goes beyond those infected with HIV. *Aids.* 2004;18(4):657-662.

31. Wood R, Middelkoop K, Myer L, et al. Undiagnosed tuberculosis in a community with high HIV prevalence: implications for tuberculosis control. *Am J Respir Crit Care Med.* 2007;175(1):87-93.

32. Corbett EL, Bandason T, Cheung YB, et al. Prevalent infectious tuberculosis in Harare, Zimbabwe: burden, risk factors and implications for control. *Int J Tuberc Lung Dis.* 2009;13(10):1231-1237.

33. Odhiambo JA, Borgdorff MW, Kiambih FM, et al. Tuberculosis and the HIV epidemic: increasing annual risk of tuberculous infection in Kenya, 1986-1996. *Am J Public Health.* 1999;89(7):1078-1082.

34. Tanzania Tuberculin Survey C. Tuberculosis control in the era of the HIV epidemic: risk of tuberculosis infection in Tanzania, 1983-1998. *International Journal of Tuberculosis & Lung Disease.* 2001;5(2):103-112.

35. Egwaga SM, Cobelens FG, Muwinge H, Verhage C, Kalisvaart N, Borgdorff MW. The impact of the HIV epidemic on tuberculosis transmission in Tanzania. *AIDS.* 2006;20(6):915-921.

36. Middelkoop K, Bekker LG, Myer L, Dawson R, Wood R. Rates of tuberculosis transmission to children and adolescents in a community with a high prevalence of HIV infection among adults. *Clinical Infectious Diseases.* 2008;47(3):349-355.

37. Rieder HL. Editorial commentary: on the risk of being and becoming infected with Mycobacterium tuberculosis. *Clin Infect Dis.* 2008;47(3):356-357.

38. Zelner JL, Murray MB, Becerra MC, et al. Identifying Hotspots of Multidrug-Resistant Tuberculosis Transmission Using Spatial and Molecular Genetic Data. *J Infect Dis.* 2016;213(2):287-294.

39. Verver S, Warren RM, Munch Z, et al. Proportion of tuberculosis transmission that takes place in households in a high-incidence area. *Lancet.* 2004;363(9404):212-214.

40. Buu TN, van Soolingen D, Huyen MN, et al. Tuberculosis acquired outside of households, rural Vietnam. *Emerg Infect Dis.* 2010;16(9):1466-1468.

41. Brooks-Pollock E, Becerra MC, Goldstein E, Cohen T, Murray MB. Epidemiologic inference from the distribution of tuberculosis cases in households in Lima, Peru. *J Infect Dis.* 2011;203(11):1582-1589.

42. Vynnycky E, Fine PE. The natural history of tuberculosis: the implications of age-dependent risks of disease and the role of reinfection. *Epidemiol Infect.* 1997;119(2):183-201.

43. Cohen T, Colijn C, Finklea B, Murray M. Exogenous re-infection and the dynamics of tuberculosis epidemics: local effects in a network model of transmission. *J R Soc Interface.* 2007;4(14):523-531.

44. Mboma SM, Houben RM, Glynn JR, et al. Control of (Multi)Drug Resistance and Tuberculosis Incidence over 23 Years in the Context of a Well-Supported Tuberculosis Programme in Rural Malawi. *PLoS One.* 2013;8(3):e58192.

45. Theron G, Jenkins HE, Cobelens F, et al. Data for action: collection and use of local data to end tuberculosis. *Lancet.* 2015;386(10010):2324-2333.

46. UNAIDS. Practical Guidelines for Intensifying HIV Prevention: Towards Universal Access. http://data.unaids.org/pub/manual/2007/20070306_prevention_guidelines_towards_universal_access_en.pdf *[Accessed 10 August 2017]*. Geneva2007.

47. Shapiro AE, Variava E, Rakgokong MH, et al. Community-based Targeted Case-finding for Tuberculosis and HIV in Household Contacts of Tuberculosis Patients in South Africa. *Am J Respir Crit Care Med.* 2012.

48. Dowdy DW, Golub JE, Chaisson RE, Saraceni V. Heterogeneity in tuberculosis transmission and the role of geographic hotspots in propagating epidemics. *Proc Natl Acad Sci U S A.* 2012;109(24):9557-9562.

49. Zelner JL, Murray MB, Becerra MC, et al. Bacillus Calmette-Guerin and isoniazid preventive therapy protect contacts of patients with tuberculosis. *Am J Respir Crit Care Med.* 2014;189(7):853-859.

50. Gupta RK, Rice B, Brown AE, et al. Does antiretroviral therapy reduce HIV-associated tuberculosis incidence to background rates? A national observational cohort study from England, Wales, and Northern Ireland. *Lancet HIV.* 2015;2(6):e243-251.

51. Lawn SD, Wilkinson RJ. ART and prevention of HIV-associated tuberculosis. *Lancet HIV.* 2015;2(6):e221-222.

52. Rieder HL. *Epidemiological basis of tuberculosis control.* Paris: International Union Against Tuberculosis and Lung Disease; 1999.

53. Madhi SA, Nachman S, Violari A, et al. Primary isoniazid prophylaxis against tuberculosis in HIV-exposed children. *N Engl J Med.* 2011;365(1):21-31.

54. Rieder HL. The dynamics of tuberculosis epidemiology. *The Indian journal of tuberculosis.* 2014;61(1):19-29.

55. Dowdy DW, Dye C, Cohen T. Data needs for evidence-based decisions: a tuberculosis modeler's 'wish list'. *Int J Tuberc Lung Dis.* 2013;17(7):866-877.

56. Pai M, Denkinger CM, Kik SV, et al. Gamma Interferon Release Assays for Detection of Mycobacterium tuberculosis Infection. *Clin Microbiol Rev.* 2014;27(1):3-20.

57. Dye C, Bassili A, Bierrenbach AL, et al. Measuring tuberculosis burden, trends, and the impact of control programmes. *Lancet Infect Dis.* 2008;8(4):233-243.

58. Dodd PJ, Millington KA, Ghani AC, et al. Interpreting tuberculin skin tests in a population with a high prevalence of HIV, tuberculosis, and nonspecific tuberculin sensitivity. *Am J Epidemiol.* 2010;171(9):1037-1045.

59. Bakir M, Dosanjh DP, Deeks JJ, et al. Use of T cell-based diagnosis of tuberculosis infection to optimize interpretation of tuberculin skin testing for child tuberculosis contacts. *Clin Infect Dis.* 2009;48(3):302-312.

60. Pai M, Dendukuri N, Wang L, Joshi R, Kalantri S, Rieder HL. Improving the estimation of tuberculosis infection prevalence using T-cell-based assay and mixture models. *International Journal of Tuberculosis and Lung Disease.* 2008;12(8):895-902.

61. Hoff ST, Peter JG, Theron G, et al. Sensitivity of C-Tb: a novel RD-1-specific skin test for the diagnosis of tuberculosis infection. *Eur Respir J.* 2016;47(3):919-928.

62. Andrews JR, Nemes E, Tameris M, et al. Serial QuantiFERON testing and tuberculosis disease risk among young children: an observational cohort study. *Lancet Respir Med.* 2017;5(4):282-290.

63. Suthar AB, Lawn SD, del Amo J, et al. Antiretroviral therapy for prevention of tuberculosis in adults with HIV: a systematic review and meta-analysis. *PLoS Med.* 2012;9(7):e1001270.

64. Lawn SD, Harries AD, Williams BG, et al. Antiretroviral therapy and the control of HIV-associated tuberculosis. Will ART do it? *Int J Tuberc Lung Dis.* 2011;15(5):571-581.

65. Munthali L, Khan PY, Mwaungulu NJ, et al. The effect of HIV and antiretroviral therapy on characteristics of pulmonary tuberculosis in northern Malawi: a cross-sectional study. *BMC Infect Dis.* 2014;14:107.

66. van Halsema CL, Fielding KL, Chihota VN, et al. Brief Report: The Effect of Antiretroviral Therapy and CD4 Count on Markers of Infectiousness in HIV-Associated Tuberculosis. *Journal of acquired immune deficiency syndromes (1999).* 2015;70(1):104-108.

67. Williams B, Maher D. Tuberculosis fueled by HIV: putting out the flames. *Am J Respir Crit Care Med.* 2007;175(1):6-8.

68. Hoa NB, Sy DN, Nhung NV, Tiemersma EW, Borgdorff MW, Cobelens FG. National survey of tuberculosis prevalence in Viet Nam. *Bull World Health Organ.* 2010;88(4):273-280.

69. Qadeer E, Fatima R, Yaqoob A, et al. Population Based National Tuberculosis Prevalence Survey among Adults (>15 Years) in Pakistan, 2010-2011. *PLoS One.* 2016;11(2):e0148293.

70. Ayles H, Schaap A, Nota A, et al. Prevalence of tuberculosis, HIV and respiratory symptoms in two Zambian communities: implications for tuberculosis control in the era of HIV. *PLoS One.* 2009;4(5):e5602.

71. van't Hoog AH, Laserson KF, Githui WA, et al. High prevalence of pulmonary tuberculosis and inadequate case finding in rural western Kenya. *Am J Respir Crit Care Med.* 2011;183(9):1245-1253.

72. Dowdy DW, Basu S, Andrews JR. Is passive diagnosis enough? The impact of subclinical disease on diagnostic strategies for tuberculosis. *Am J Respir Crit Care Med.* 2013;187(5):543-551.

73. Kranzer K, Afnan-Holmes H, Tomlin K, et al. The benefits to communities and individuals of screening for active tuberculosis disease: a systematic review. *Int J Tuberc Lung Dis.* 2013;17(4):432-446.

74. Kranzer K, Khan P, Godfrey-Fausset P, Ayles H, Lonnroth K. Tuberculosis control. *Lancet.* 2016;387(10024):1159-1160.

75. Fine PE, Small PM. Exogenous reinfection in tuberculosis. *N Engl J Med.* 1999;341(16):1226-1227.

Appendix I

Copyright agreements

Research paper I: Lancet Infectious Disease

ELSEVIER LICENSE
TERMS AND CONDITIONS

May 14, 2017

This Agreement between Palwasha Y Khan ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

License Number	4107591065590
License date	May 14, 2017
Licensed Content Publisher	Elsevier
Licensed Content Publication	The Lancet Infectious Diseases
Licensed Content Title	The transmission of Mycobacterium tuberculosis in high burden settings
Licensed Content Author	Tom A Yates,Palwasha Y Khan,Gwenan M Knight,Jonathon G Taylor,Timothy D McHugh,Marc Lipman,Richard G White,Ted Cohen,Frank G Cobelens,Robin Wood,David A J Moore,Ibrahim Abubakar
Licensed Content Date	February 2016
Licensed Content Volume	16
Licensed Content Issue	2
Licensed Content Pages	12
Start Page	227
End Page	238
Type of Use	reuse in a thesis/dissertation
Portion	full article
Format	both print and electronic
Are you the author of this Elsevier article?	Yes
Will you be translating?	No
Order reference number	
Title of your thesis/dissertation	Investigating Mycobacterium tuberculosis transmission in rural Malawi
Expected completion date	Sep 2017
Estimated size (number of pages)	200
Elsevier VAT number	GB 494 6272 12
Requestor Location	Palwasha Y Khan 23 Windlass Court Cardiff, CF10 4NG United Kingdom Attn: Palwasha Y Khan
Publisher Tax ID	GB 494 6272 12
Total	0.00 USD

Research paper II: International Journal of Tuberculosis and Lung Disease

From: **Clare Pierard** cpierard@theunion.org
Subject: RE: [BULK] TheUnion.org Contact Form: Request to use copyright material in thesis
Date: 15 May 2017 at 11:33
To: palwasha.khan@lshtm.ac.uk
Cc: journal@theunion.org

Dear Palwasha

Many thanks for your request. Permission is granted for the purpose indicated.

Please use the following wording: "Reprinted with permission of the International Union Against Tuberculosis and Lung Disease. Copyright © The Union."

And cite: Author(s).Title. Int J Tuberc Lung Dis year; volume: pages.

Kind regards

Clare Pierard

Clare Pierard, Managing Editor
International Journal of Tuberculosis and Lung Disease
CALL: (+33) 1 44 32 03 60 SKYPE: clare.pierard
E-MAIL: cpierard@theunion.org | journal@theunion.org
http://www.ingentaconnect.com/journals/browse/iuatld/ijtld (/pha)

The Union
68, boulevard Saint-Michel
F-75006 Paris
France

theunion.org
worldlunghealth.org

De : Palwasha Khan [mailto:palwasha.khan@lshtm.ac.uk]
Envoyé : dimanche 14 mai 2017 12:54
À : documents@theunion.org
Objet : [BULK] TheUnion.org Contact Form: Request to use copyright material in thesis
Importance : Faible

There has been a submission made through your form, The Union: Contact Us:

ID	772
Recipient	documents@theunion.org
Inquiry Type	Ordering a scientific publication
Member	yes
Member Number	CU-
Last Name	Khan
First Name	Palwasha
Email	palwasha.khan@lshtm.ac.uk
Organization	London School of Hygiene & Tropical Medicine

Appendix II

Mycobacterium tuberculosis transmission meeting agenda

M.tb transmission 2014

PROGRAMME AT-A-GLANCE

0830-0900 Registration

0900-0905 Brief introduction and housekeeping - Dave Moore, LSHTM

~ Transmission - what is known ~

UPDATE Chair, Ali Zumla, UCL
0905-0925 A potted history of early research on *M.tb* transmission - Paul Fine, LSHTM
0925-0945 What have molecular techniques taught us about *M.tb* transmission? - Judith Glynn, LSHTM

~ Active case finding ~

DEBATE Chair, Katharina Kranzer, LSHTM
0945-1030 'FURTHER IMPROVEMENTS IN ACTIVE CASE FINDING CAN ONLY HAVE A LIMITED IMPACT ON TB CONTROL IN THE COMING DECADES'
- David Dowdy, John Hopkins (proposing motion)
- Helen Ayles, ZAMBART/LSHTM (opposing motion)

1030-1100 Coffee

~ Where does *M.tb* transmission occur? ~

NEW SCIENCE Chair, Jimmy Whitworth, Wellcome Trust
1100-1115 Carbon dioxide exposure as a proxy for *M.tb* exposure - Robin Wood, Desmond Tutu HIV Centre
1115-1130 Inferring chains of transmission using whole genome sequencing - Josie Bryant, UCL/Sanger Institute
1130-1145 Directly detecting *M.tb* in the environment - Allan Bennett, Porton Down
1145-1200 Measuring and modelling social contact carrying a transmission risk - Pete Dodd, CREATE Consortium

NEXT STEPS Chair, Alison Grant, LSHTM
1200-1210 Fantasy study- Daren Caruana, UCL
1210-1255 **DISCUSSION** - 'HOW CAN WE MOST EFFECTIVELY LOCATE *M.TB TRANSMISSION*?'

1255-1355 Lunch

M.TB TRANSMISSION MEETING - 11TH NOVEMBER 2014
VENUE: JOHN SNOW LECTURE THEATRE, LSHTM

M.tb transmission 2014

~ How can we interrupt transmission in high burden settings? ~

NEW SCIENCE Chair, Charles Mgone, EDCTP

1355-1410 Should we focus on transmission hotspots? - Ted Cohen, Yale SPH

1410-1425 What guinea pigs have taught me - Rod Escombe, Imperial College

1425-1440 Retrofitting public spaces to reduce *M.tb* transmission - Jonathon Taylor, UCL

1440-1455 The dynamics of drug-resistant *M.tb* transmission - Gwen Knight, LSHTM

NEXT STEPS Chair, Katherine Fielding, LSHTM

1455-1505 Fantasy study- Andrew Hayward, UCL

1505-1550 **DISCUSSION** - 'INTERRUPTING TRANSMISSION - HOW SHOULD WE DESIGN THE NEXT INTERVENTION STUDIES?'

~ Moving forwards ~

1550-1610

1610-1615 **Closing remarks** - Ibrahim Abubakar, UCL

1615-1645 Coffee

1645-1730 SMALL GROUP DISCUSSIONS ON FUTURE COLLABORATIONS

Observational studies Chair, Molebogeng Rangaka, UCL

Modelling studies Chairs, Emilia Vynnycky, PHE/LSHTM & Michael Kimerling, BMGF

Intervention studies Chair, Marc Lipman, UCL

1730-1900 Wine Reception

M.TB TRANSMISSION MEETING - 11TH NOVEMBER 2014
VENUE: JOHN SNOW LECTURE THEATRE, LSHTM

List of attendees

	Surname	First Name
1	Abdul Rahman	Norana
2	Abubakar	Ibrahim
3	Adamu	Aishatu
4	Ahmedov	Sevim
5	Al-Bakri	Isbah
6	Altamirano	Hector
7	Altass	Lynn
8	Appleby	Yasmin
9	arnold	amber
10	arnold	amber
11	Astley	Philip
12	Aung	Htin
13	Ayles	Helen
14	Bager	Line
15	Bailey	Sarah Lou
16	Barer	Mike
17	Barnham	Kate
18	Bennett	Allan
19	Bidonde	Julia
20	Blake	Sara
21	Bolt	Hikaru
22	Booker	Sharon
23	Booth	Helen
24	Brown	James
25	Browne	Catherine
26	Burkitt	Andy
27	Byng-Maddick	Rachel
28	Cadogan	Paula
29	carroll	kevin
30	Caruana	Daren
31	Casali	Nicki
32	Cavany	Sean
33	Chiang	Nicole
34	Cliff	Jackie
35	Cobelens	Frank
36	Cohen	Theodore
37	Colijn	Caroline
38	Connell	David
39	connor	matthias
40	Crampin	Amelia
41	Dardamissis	Dr Evdokia
42	Davidson	Jennifer
43	de Jager	Bernadette
44	DE VRIES	GERARD
45	Dedicoat	Martin
46	dewan	mary
47	Dockrell	Hazel
48	Dodd	Pete
49	Dowdy	David
50	Dyson	Hugh
51	Edoo	Mohammud
52	El Haj	Salima
53	ESTIBALIZ	MONDEJAR
54	Fielding	Katherine
55	Fine	Paul
56	Fletcher	Helen
57	Floyd	Sian
58	Foley	Phil
59	Gibbons	Chris
60	Gibertoni Cruz	A
61	Glynn	Judith
62	Gomes	Dulce
63	Gorak-Stolinska	Patricia
64	Grgic	Ljuban
65	Guerra-Assuncao	Jose Afonso
66	Gupta	Rishi
67	Gupta-Wright	Ankur
68	Guzzetta	Giorgio
69	HANIFA	YASMEEN
70	Head	Michael
71	Heinsbroek	Ellen
72	Herman	Jo
73	Hickson	Vicky
74	Highfield	Sam
75	Hippner	Piotr
76	Hoppe	Anne
77	Horton	Katherine
78	Hotton	Majella
79	Houben	Rein
80	Isosomppi	Sanna
81	Jackson	Charlotte
82	Jacobs	Ashley
83	Jones	Eben
84	KAJIWARA	MARI
85	Khan	Palwasha
86	kiff	jeremy
87	Knight	Gwen
88	Koeser	Claudio
89	kong	karen
90	Kranzer	Katharina
91	kuma	george
93	Kumar	Kartik
94	Lalor	Maeve
95	Lamb	Georgia

	Surname	First Name
96	Lamden	Kenneth
97	Laurence	Yoko
98	Law	Irwin
99	Lee	Jaehee
100	Lee	Nathaniel
101	Lewis	James
102	Li	Matthew
103	Lipman	Marc
104	Loutet	Miranda
105	MacPherson	Peter
106	maguire	helen
107	Marsh	Philip
108	Martin	Jenni
109	Martineau	Adrian
110	Marziano	Valentina
111	Masu	Adelaide
112	McCreesh	Nicky
113	McHugh	Tim
114	Mears	Jessica
115	Menezes	Dee
116	MENSAH	CAESAR
117	Mgone	Charles
118	Mikhail	Amy
119	Millard	James
120	Miller	Alex
121	Moodley	Riya
122	Moore	Dave
123	Moores	Alexandra
124	Muliaditan	Morris
125	MULOIWA	RUDZANI
126	Munang	Melinda
127	Muzyamba	Morris
128	Nguipdop-Djomo	Patrick
129	Nice	Jako
130	Nimmo	Camus
131	Nunes	Carla
132	Nunn	Hazel
133	Nyirenda	Moffat
134	O'Callaghan	Chris
135	Oguti	Blanche
136	ortiz canseco	julio
137	Orton	Jennie
138	Pareek	Manish
139	Parwati	Cicilia Gita
140	Penna	Maria Lucia
141	Phelan	Jody
142	Phillips	Patrick
143	Pillay	Timesh
144	Plazzotta	Giacomo
145	Pollara	Gabriele
146	Potton	Elspeth
147	Prabowo	Satria Arief
148	Pritchard	Jennifer
149	purcell	bernadette
150	Quay	Doris
151	Rahman	Yasmine Grace
152	Rangaka	Molebogeng
153	Ranzani	Otavio
154	Ratner	Jennifer
155	Rawkins	Ann
156	Roe	Jennifer
157	Ross	David
158	Roy	Anjana
159	Rudgard	William
160	Saavedra-Campos	Maria
161	Saito	Makoto
162	SALIMEE	Sultan
163	Schultze	Anna
164	Scuffell	Jamie
165	seymore	rose
166	Sharma	Sunny
167	Sole	Genna
168	Southern	Jo
169	Stevenson	Daniel
170	Story	Alistair
171	Stuart	Arabella
172	Sumner	Tom
173	surey	julian
174	Swann	Catherine
175	Takaya	Saho
176	Taubman	Diana
177	Taylor	Jonathon
178	Telisinghe	Lily
179	Thomas	Lucy
180	Thomas	Sherine
181	Thompson	Mary
182	Thompson	Courtney
183	van Hest	Rob
184	VAN HUNEN	RIANNE
185	Ward	Charlotte
186	White	Richard
187	Whitehead	Nuala

	Surname	First Name
188	Whitworth	Jimmy
189	Wills	Genevieve
190	Winter	Joanne
191	Witt	Karolina
192	Wood	Robin
193	Wright	Cameron
194	Wu	Ping
195	Wurie	Fatima
196	Yassine	Bilal
197	Yates	Tom
198	Zelmer	Andrea
199	Grant	Alison

Appendix III

Cross-sectional analysis of 2nd round of TST data

The incidence of *M.tb* infection estimated from the TST cohort study (Chapter 5) was much higher than the average annual risk of *M.tb* infection (ARTI) estimated from the cross-sectional survey (Chapter 4). As an internal validity check of the data, the second round of TST data was analysed as cross-sectional data, using the TST≥15mm cut-off to estimate *M.tb* prevalence, with the premise that ARTI should not have changed significantly between rounds.

Only those children aged between 2.0 to 4.9 years of age at the time of the 2nd TST were eligible to be included in this analysis of the 2nd round of TST, as per eligibility of the 1st round (baseline). 14 of the prevalent *M.tb* infections identified in the first round were included in the numerator and denominator, as these children would have been eligible to have been tested (i.e. aged between 2.0 – 4.9 years when skin testing was taking place in their residential area as part of the 2nd round). Of the 60 children who had a TST≥15mm at the 2nd round of skin testing, 36/60 (60%) had zero induration at baseline. Table A1 shows the difference in ARTI between the 1st and 2nd round.

Table A1. illustrating the ARTI estimated from *M.tb* prevalence at the 2 rounds

TST cohort study	***M.tb* prevalence (TST≥15mm)**			**Mean age**	**ARTI**	
	n/N	%	95 % CI	(years)	%	95 % CI
1st Round	35/3170	1.1	0.8 – 1.5	3.5	0.3	0.2 – 0.4
2nd Round	74/2105	3.5	2.8 – 4.4	3.7	1.0	0.8 – 1.2

Interpretation

The ARTI estimated from the 2nd round of skin testing is significantly higher than the ARTI estimated from the 1st round suggesting that this increase in *M.tb* prevalence observed in the 2nd round is a result of 'boosting' of the host response from the TST given in the 1st round.

Appendix IV

Sensitivity analysis of the association of incident M.tb infection and age

Table A2. Sensitivity analysis using different criteria to define incident *M.tb* infection and the association with age

		Increment between 1st and 2nd TST (1st TST <10mm)									
		10mm		13mm		15mm		17mm		19mm	
		Incidence per 100py	IRR (95% CI)	Incidence per 100py	IRR (95% CI)	Incidence per 100py	IRR (95% CI)	Incidence per 100py	IRR (95% CI)	Incidence per 100py	IRR (95% CI)
Overall		2.5 (2.1 – 2.9)	-	1.6 (1.3 – 2.0)	-	1.1 (0.8 – 1.4)	-	0.6 (0.4 – 0.8)	-	0.2 (0.1 – 0.4)	-
Age at 1st TST (years)	**<1.0**	1.4 (0.8 – 2.3)	1	0.8 (0.4 – 1.7)	1	0.3 (0.1 – 1.0)	1	0.1 (0.01 – 0.7)	1	0.1 (0.02 – 0.7)	1
	1.0 - 1.9	1.9 (1.3 – 2.7)	1.4 (0.7 – 2.7)	1.0 (0.6 – 1.6)	1.2 (0.5 – 2.8)	0.8 (0.5 – 1.3)	2.5 (0.7 – 8.8)	0.5 (0.2 – 1.0)	4.6 (0.6 – 36.7)	0.06 (0.008 – 0.4)	0.6 (0.04 – 9.1)
	2.0 - 2.9	2.6 (1.9 – 3.5)	1.9 (1.0 – 3.6)	1.9 (1.3 – 2.7)	2.3 (1.1 – 5.0)	1.4 (0.9 – 2.1)	4.4 (1.3 – 14.6)	0.6 (0.3 – 1.1)	5.9 (0.8 – 46.1)	0.3 (0.09 – 0.7)	2.4 (0.3 – 21.1)
	3.0 - 3.9	3.5 (2.5 – 4.9)	2.5 (1.3 – 4.7)	2.4 (1.6 – 3.6)	2.8 (1.2 – 6.3)	1.6 (1.0 – 2.6)	5.0 (1.4 – 17.3)	0.8 (0.4 – 1.7)	8.1 (1.0 – 64.5)	0.4 (0.2 – 1.1)	4.0 (0.5 – 36.0)
	4.0 - 4.9	4.6 (2.9 – 7.4)	3.3 (1.6 – 6.9)	3.3 (1.9 – 5.7)	3.8 (1.6 – 9.5)	1.9 (0.9 – 4.0)	6.2 (1.6 – 24.4)	1.1 (0.4 – 2.9)	10.5 (1.2 – 94.2)	0.8 (0.3 – 2.5)	7.8 (0.8 – 74.7)
Age as a linear variable			1.3 (1.2 – 1.6)		1.4 (1.2 – 1.7)		1.5 (1.2 – 1.8)		1.5 (1.1 – 2.0)		2.0 (1.2 – 3.2)

TST tuberculin skin test; py person-years; IRR incidence rate ratio; CI confidence interval

Appendix V

Population attributable fraction of known household contact

It was possible to estimate the population attributable fraction (PAF) of known household contact with an infectious TB case for prevalent *M.tb* infection using two different methods in this thesis. The two methods used were as follows:

1. $\text{PAF} = \text{p'}\left(\frac{\text{RR}-1}{\text{RR}}\right)$

where p' = *proportion of cases exposed* to household contact with infectious TB case; RR = relative risk of the association of prevalent *M.tb* infection in children and household contact with infectious TB case

2. $\text{PAF} = \frac{\text{p}(\text{RR}-1)}{\text{p}(\text{RR}-1)+1}$

where p = *proportion of the population exposed* to household contact with infectious TB case; RR = as above

To ensure consistency of PAF estimates using the 2 methods, the definition of prevalent *M.tb* infection is based on the TST≥15mm cut-off method and restricted to the children aged between 2.0 to 4.9 years.

Estimates of PAF

Method 1 estimates the PAF of household TB contact to be 2.1% as calculated in Chapter 4, where p'=1/35 and RR=3.6 (see Chapter 4 page 347 of published paper with erratum noted at end of chapter). Even when the unadjusted RR is used (RR=4.6), the PAF is 2.2%.

Method 2 uses data from the TB case-contact household study where prevalence of *M.tb* infection in children resident with a smear-positive TB index case is 0.319 (Chapter 3 Table 2: estimate of *M.tb* prevalence in 'Higher Risk' group aged ≥2 years). The comparator (prevalence of *M.tb* infection in children from DSS with no known household contact with an infectious TB case) is the same as Method 1, (0.011; Chapter 3 Table 2, which is the same as Chapter 4 Table 1 excluding household contact: 34/3150); yields a relative risk of 29 (0.319/0.011). The proportion of the population exposed to an infectious TB case in the household is 0.0063 (20/3170), thus giving a PAF of 15%. Repeating the above PAF calculation using data from the baseline TST survey alone gives a PAF of 2.3%, which is similar to the PAF estimated from method 1.

The key difference driving this discrepancy between the PAF estimates is the much higher absolute risk in the child household contacts of an infectious TB case derived from the TB case-contact household study.

Appendix VI

M.tb infection prevalence stratified by HIV exposure status

Background

Several studies have found a high prevalence of *M.tb* infection among HIV-exposed uninfected infants in sub-Saharan Africa,[1-6] which may be one of the contributing factors for the increased risk of morbidity and mortality observed in this population despite the successful roll-out Prevention of Mother-to-Child Transmission (PMTCT) programmes.[7-9] As data on HIV-exposure status was available for children in this study, it thought that an exploratory analysis stratifying by HIV-exposure status would be of interest as these data are unlikely to be available on a population-level in other settings. However, this analysis was not included in the published paper due to the very limited sample size.

The idea for stratifying by HIV-exposure status had come from the original paper by Rust and Thomas which had also included a stratified analysis based on calendar period and region. The findings had demonstrated substantial regional variation in *M.tb* infection prevalence consistent with TB epidemiology at that time in the US, which were not as apparent when using the traditional cut-off method of TST≥10mm. The authors also found that infection prevalence had declined by 8.5% per annum between 1961 to 1968 in the US Navy recruits using their method, rather than a decline of 4.5% per annum using the traditional cut-off method, postulating that reactions to atypical mycobacteria had obscured the extent of decline in *M.tb* infection prevalence over time.[10]

Study hypothesis: M.tb infection prevalence is higher in HIV-exposed children than in HIV-unexposed children at a population-level in the DSS.

Definition of HIV-exposure status: Children whose mothers were known to be HIV-positive at the time of delivery were defined as HIV-exposed, while children whose mothers were HIV-negative after or up to 1 year before their birth were defined as HIV-unexposed.

Result of exploratory data analysis

Data on HIV-exposure status of the children was available on 4840/5119 (94.5%). Web figure 4 in the paper (chapter 3) illustrates the percentage distribution of induration category stratified by HIV-exposure and contact status of the child and Table A3 presents the frequency distribution of induration. Table A4 (below) shows the different estimates of *M.tb* infection prevalence stratified by risk group and HIV-exposure status. The population of interest for this analysis was the 'lower risk group' which are the children resident in the DSS.

All the methods of deriving infection prevalence demonstrate similar estimates for HIV-exposed and HIV-unexposed children apart from the Rust and Thomas model. The point estimate for *M.tb* infection prevalence is 4.3 times higher in the HIV-exposed population than HIV-unexposed children (2.5% vs 0.6%), although the 95% confidence interval for the HIV-exposed group includes zero due to the very small sample size of HIV-exposed children in the DSS. The ARTI for HIV-exposed children in the DSS was 0.9% (95% CI: 0 – 7.2%) and 0.2% (95% CI: 0.04 – 0.6%) for HIV-unexposed children.

Table A3. Frequency Distribution of Tuberculin Data in Children Aged Under 5 years Stratified by Risk Group and HIV-exposure status

	FREQUENCY DISTRIBUTION OF TUBERCULIN DATA							
INDURATION SIZE (MM)	LOWER RISK GROUP (N=4725)				HIGHER RISK GROUP (N=115)			
	HIV-EXPOSED		HIV-UNEXPOSED		HIV-EXPOSED		HIV-UNEXPOSED	
	n	%	n	%	n	%	n	%
0	225	85.2	3770	84.5	91	52.9	54	55.1
2 – 3	2	0.8	24	0.5	1	5.9	1	1.0
4 – 5	2	0.8	67	1.5	1	5.9	1	1.0
6 – 7	7	2.7	150	3.4	0	0	3	3.1
8 – 9	14	5.3	171	3.8	0	0	1	1.0
10 – 11	3	1.1	90	2.0	0	0	2	2.0
12 – 13	3	1.1	79	1.8	1	5.9	5	5.1
14 – 15	5	1.9	58	1.3	0	0	6	6.1
16 – 17	1	0.4	29	0.7	2	11.8	8	8.2
18 – 19	1	0.4	17	0.4	2	11.8	9	9.2
20 – 21	1	0.4	2	0.1	1	5.9	6	6.1
22+	0	0	4	0.1	0	0	2	2.0

Table A4. *M.tb* infection prevalence estimates and 95% confidence intervals stratified by HIV exposure and contact status using different methods

Method to estimate *M.tb* infection prevalence		LOWER RISK GROUP (DSS) (N=4725)				HIGHER RISK GROUP (TB HOUSEHOLD) (N=115)			
		HIV-UNEXPOSED N=4461		HIV-EXPOSED N=264		HIV-UNEXPOSED N=98		HIV-EXPOSED N=17	
		%	(95 % CI)	%	(95 % CI)	%	(95 % CI)	%	(95 % CI)
TST cut-off	10 mm	6.3	(5.6 – 7.0)	5.2	(2.8 – 8.5)	38.8	(29.1 – 49.2)	35.3	(14.2 – 61.7)
	15mm	1.9	(1.5 – 2.3)	1.8	(0.6 – 4.2)	29.6	(20.8 – 39.7)	29.4	(10.3 – 56.0)
Fixed mirror		1.3	(1.0 – 1.6)	1.5	(0.4 – 3.7)	40.8	(31.0 – 51.2)	41.2	(18.4 – 67.1)
Mixture model		7.0	(3.7 – 10.6)	7.5	(0.07 – 15.2)	32.7	(11.8 – 55.5)	37.2	(19.7 – 52.3)
Rust & Thomas model		**0.6**	**(0.1 – 1.5)**	**2.5**	**(0.0 – 19.1)**	**35.2**	**(23.8 – 46.4)**	**39.7**	**(0.0 – 79.3)**

Interpretation of findings

Using the Rust and Thomas method, prevalence of *M.tb* infection was higher in children born to HIV-positive mothers compared to children born to HIV-negative mothers in the Karonga DSS as hypothesised. But findings of this exploratory analysis should be interpreted with caution and not over-stated due to the limited sample size.

Interestingly, the magnitude of the effect estimate of HIV-exposure status on *M.tb* prevalence using the Rust and Thomas method to analyse tuberculin data in this study is similar to that seen in a cross-sectional study in Uganda where children born to HIV-infected mothers had a 4-fold higher prevalence of Quantiferon Gold-in-Tube assay (QFT-GIT) positivity compared to those children born to HIV-negative mothers (6.4% vs. 1.5%; prevalence ratio 4.3).[4] The corresponding ARTI for HIV-exposed children and HIV-unexposed children in the Ugandan study was approximately 2.5% and 0.5% respectively. This is much higher than the ARTI seen in both HIV-exposed and HIV-unexposed children in the Karonga DSS and is in keeping with local epidemiology. The annual TB incidence (all cases) is 400 per 100,000 population in the rural district in which the study in Uganda was conducted, whereas in Karonga it is approximately 120 per 100,000 population.

However, the studies are not entirely comparable because the prevalence of *M.tb* infection in the HIV-exposed group in our setting, reflects the prevalence of *M.tb* infection in HIV-exposed uninfected and HIV-positive children as the child's HIV status was unknown, whereas the Ugandan study only included HIV-exposed

uninfected children. It maybe that *M.tb* infection prevalence in HIV-exposed uninfected children in our study may be even higher if HIV-positive children with advanced immunosuppression have been included, as these children are less likely to mount a response to tuberculin even if infected with *M.tb*,[11] leading to an underestimate of *M.tb* infection prevalence. Alternatively, HIV-positive children may be more susceptible to *M.tb* infection following exposure than HIV-exposed uninfected children resulting in an over-estimate of the prevalence of *M.tb* infection in HIV-exposed uninfected children in our population. Either way, ART for prevention of mother to child transmission is widely used in the DSS,[12] and the proportion of infected children is likely to be very low.

References

1. Cotton MF, Schaaf HS, Lottering G, et al. Tuberculosis exposure in HIV-exposed infants in a high-prevalence setting. *Int J Tuberc Lung Dis.* 2008;12(2):225-227.
2. Cotton MF, Slogrove A, Rabie H. Infections in HIV-exposed uninfected children with focus on sub-Saharan Africa. *Pediatr Infect Dis J.* 2014;33(10):1085-1086.
3. Cranmer LM, Kanyugo M, Jonnalagadda SR, et al. High prevalence of tuberculosis infection in HIV-1 exposed Kenyan infants. *Pediatr Infect Dis J.* 2014;33(4):401-406.
4. Marquez C, Chamie G, Achan J, et al. Tuberculosis Infection in Early Childhood and the Association with HIV-exposure in HIV-uninfected Children in Rural Uganda. *Pediatr Infect Dis J.* 2016;35(5):524-529.
5. Sugandhi N, Rodrigues J, Kim M, et al. HIV-exposed infants: rethinking care for a lifelong condition. *AIDS.* 2013;27 Suppl 2:S187-195.
6. Shapiro RL, Lockman S. Mortality among HIV-exposed infants: the first and final frontier. *Clin Infect Dis.* 2010;50(3):445-447.
7. Gupta A, Nayak U, Ram M, et al. Postpartum tuberculosis incidence and mortality among HIV-infected women and their infants in Pune, India, 2002-2005. *Clin Infect Dis.* 2007;45(2):241-249.
8. Landes M, van Lettow M, Chan AK, Mayuni I, Schouten EJ, Bedell RA. Mortality and health outcomes of HIV-exposed and unexposed children in a PMTCT cohort in Malawi. *PLoS One.* 2012;7(10):e47337.

9. Afran L, Garcia Knight M, Nduati E, Urban BC, Heyderman RS, Rowland-Jones SL. HIV-exposed uninfected children: a growing population with a vulnerable immune system? *Clin Exp Immunol.* 2014;176(1):11-22.
10. Rust P, Thomas J. A method for estimating the prevalence of tuberculosis infection. *Am J Epidemiol.* 1975;101(4):311-322.
11. Mandalakas AM, Kirchner HL, Walzl G, et al. Optimizing the detection of recent tuberculosis infection in children in a high tuberculosis-HIV burden setting. *Am J Respir Crit Care Med.* 2015;191(7):820-830.
12. Price AJ, Kayange M, Zaba B, et al. Uptake of prevention of mother-to-child-transmission using Option B+ in northern rural Malawi: a retrospective cohort study. *Sex Transm Infect.* 2014;90(4):309-314.

Appendix VI

Exploratory geo-spatial analyses in the Karonga DSS

Geo-positional data were available for: (i) all children in the baseline TST survey and cohort study; (ii) diagnosed smear-positive TB patients from 2007 to 2012 (during the lifetime of the children included in the baseline survey); and all adults aged ≥15 years in the DSS with HIV and antiretroviral (ART) use status.

Exploratory analyses were undertaken to examine if there was any evidence for 'hotspots' in the DSS of:

i. smear-positive TB cases
ii. adults with diagnosed HIV infection
iii. adult ART use
iv. children with prevalent *M.tb* infection from the baseline TST survey
v. children with incident *M.tb* infection from TST cohort

Brief statistical methods

Maps were generated which highlighted areas in continuous space with greater-than-expected risk of each outcome. A non-parametric distance-based mapping approach was used, as this approach is less sensitive to asymmetric spatial patterns of cases and controls than Kernel-density estimation.[1]

These 'hot spot' maps, present a visual representation of the relative risk of being a 'case' or a 'control' at each spatial location. Cases were defined as follows: for analysis (i) all adults aged ≥15 years resident in the DSS from 2007 to 2012 with diagnosed smear-positive TB were designated as cases, and all adults aged ≥15 years without smear-positive TB were designated as controls; for analysis (ii) all

adults aged ≥15 years resident in the DSS in 2014 diagnosed with HIV infection were defined as cases and all those with known HIV-negative status as controls. The same principle was repeated for (iii) adult ART use, and (iv) prevalent and (v) incident *M.tb* infection in children. A colour scheme to visually highlight 'hot spots' was generated by creating 100 simulated datasets with randomly permutated case labels, with each simulated dataset representing a scenario in which the overall ratio of cases to controls remains constant (and identical to observed data) but in which the case/control status is independent of household location. The distance-based algorithm is applied to each of the randomised simulated datasets and minimum and maximum score values are obtained using the same 100 x 100 grid. These values were then used to create a colour scale for the map.

Different definitions of prevalent and incident *M.tb* infections were examined, including TST≥10mm cut-off and TST≥15mm cut-off for prevalent *M.tb* infection and TST conversion≥10mm and TST conversion≥13mm for incident *M.tb* infection. Datasets were created using Stata version 13.1 for Mac (Stata Corporation, College Station, TX) and ArcGIS 10® software (Environmental Systems Research Institute, Redlands, CA, USA). All the geospatial analyses were undertaken in R (R Foundation for Statistical Computing, Vienna) using the *hotspotr* package.[3]

Results

Figure A1 summarises the results of the geospatial analyses. Due to confidentiality issues, hot spot maps were displayed as *x y* co-ordinates on a grid

rather than on geographical maps as areas of higher HIV prevalence/ART use were easily identifiable from the map of the DSS. The grids generated for adult HIV prevalence and adult ART use were identical so the only the grid depicting ART use in the DSS is displayed in the figure. There was no evidence of a hot spot of incident *M.tb* infection in the DSS using the definition of TST conversion ≥ 13mm. Map (d) shown below is generated using TST conversion ≥10mm to define incident *M.tb* infection and in map (c) definition of prevalent infection was TST≥15mm in those aged 2-4 years. There was no evidence of a hot spot of prevalent *M.tb* infection in children using the more sensitive cut-off of TST≥10mm.

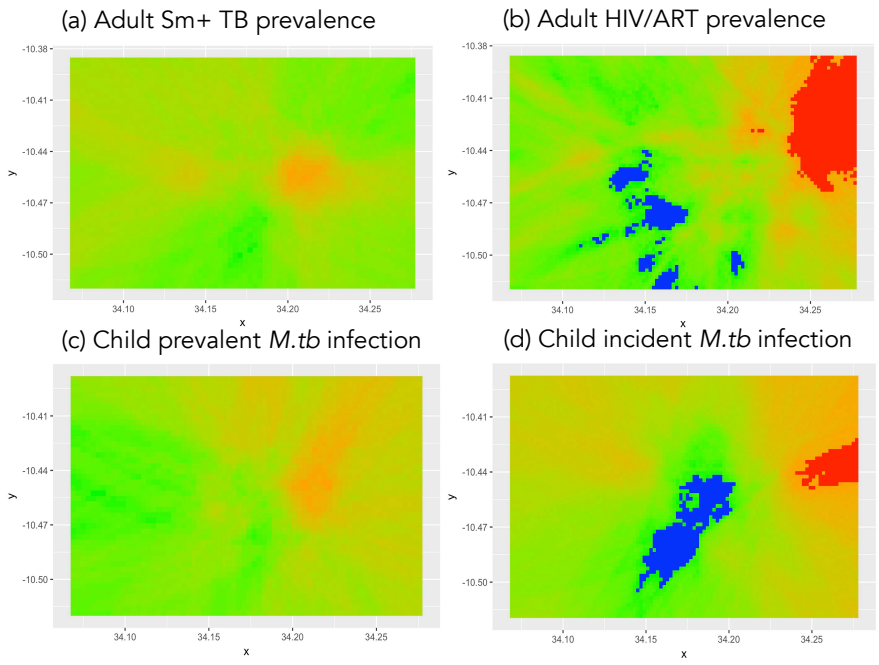


Figure A1. 'Hotspot' maps in the Karonga demographic surveillance area.

KEY: Red areas highlight score values greater than the 95th percentile of the maximum value from the randomly permuted maps; blues area highlight score values less than 5th percentile of the minimum value from the randomly permuted maps.

Interpretation of findings

There was no evidence of a hot spot of either prevalent smear-positive TB cases in the DSS between 2007-2012 and no evidence of a hot spot of prevalent *M.tb* infection in children aged 2-4 years in the DSS. There was also no evidence of hotspot of incident *M.tb* infection using the definition used in the risk factor analysis in Chapter 5 (TST conversion ≥13mm). However, when using the less stringent cut-off of TST conversion ≥ 10mm to define incident *M.tb* infection, there was evidence of a hotspot of incident *M.tb* infection in the same area of the DSS as the hotspot of adult HIV prevalence/ART use in the DSS.

This is an exploratory ecological analysis and findings should not to be over-interpreted. Especially in view of the fact that there was no evidence of a hot spot using the TST conversion ≥13mm as the definition of incident *M.tb* infection. This may have been because of reduced power due to a smaller sample of 'cases', compared to using TST conversion≥10mm as the definition of *M.tb* infection. However, the finding of the hot spot of incident *M.tb* infection albeit using the less stringent definition in the same location as the hotspot of adult HIV/ART use does lend itself to generating a study hypothesis that should be further investigated. Namely, is HIV/ART use driving community *M.tb* transmission, as inferred from TST conversion in pre-school children in some areas in the DSS?

References

1. Zelner JL, Murray MB, Becerra MC, et al. Identifying Hotspots of Multidrug-Resistant Tuberculosis Transmission Using Spatial and Molecular Genetic Data. *J Infect Dis.* 2016;213(2):287-294.

2. Manjourides J, Lin HH, Shin S, et al. Identifying multidrug resistant tuberculosis transmission hotspots using routinely collected data. *Tuberculosis (Edinb).* 2012;92(3):273-279.
3. *Hotspotr* [computer program]. https://jzelner.gitlab.io/hotspotr/2016.