Informing development strategies for new tuberculosis vaccines: mathematical modelling and novel epidemiological tools

Rebecca Claire Harris, MBioch MSc

Thesis submitted in accordance with the requirements for the degree of

Doctor of Philosophy

September 2017

Department of Infectious Disease Epidemiology
Faculty of Epidemiology and Population Health
London School of Hygiene & Tropical Medicine
University of London

Funded by the UK Medical Research Council

DECLARATION OF OWN WORK

I, Rebecca Claire Harris, confirm that the work presented in this thesis is my own. Where information has been derived from other sources or others have contributed to the work, I confirm that this has been indicated in the thesis.

Rebecca Claire Harris

September 2017

Abstract

Background: With an estimated 10.4 million incident tuberculosis (TB) cases in 2015, and the current trajectory of 1.5%/year incidence rate decline, new TB vaccines are urgently needed to help meet the WHO goal of tuberculosis elimination by 2050. However, insufficient epidemiological evidence exists to inform TB vaccine development strategies and to assist clinical trial site selection and design. Research to meet these data needs is critical to accelerate TB vaccine development. To maximise the future impact of new TB vaccines, estimates of the population-level impact of vaccine characteristics and implementation strategies are needed to inform design of TB vaccine target product profiles. To accelerate and de-risk clinical trials, appropriate epidemiological data are required to inform trial site selection, sample size calculations and recruitment strategy. However, data availability at trial sites is limited, and prospective studies are resource-intensive, so new methods are needed to collect appropriate data to inform TB vaccine trial design.

To inform data-driven development strategies for new TB vaccines, the aims of this thesis were to 1) estimate the epidemiological impact of new TB vaccine characteristics and implementation in China to inform design of target product profiles; and 2) to develop a novel epidemiological spatial mapping tool capable of informing clinical trial design for new TB vaccines in low-income, high-burden settings.

Methods: A deterministic, age-structured, *Mycobacterium tuberculosis* transmission model was developed and calibrated to age-stratified epidemiological and demographic data from China. This was employed to estimate the population-level epidemiological impact of new TB vaccines over the 2025-2050 time horizon, through an exploration of potential vaccine characteristics and implementation strategies.

A new methodology for empirical data collection to determine spatial distribution of TB notifications was developed. The electronic PArticipant Locator application (ePAL app) combined 3,243 community-identified points of interest with high resolution

satellite maps, within an electronic tablet-based case report form. The app was integrated in to the National Tuberculosis Programme in Blantyre, Malawi, for collection of demographics, health status and coordinates of place of residence for patients ≥18 years initiating TB treatment. Accuracy of ePAL-recorded co-ordinates was evaluated against GPS coordinates obtained at the participant's place of residence.

Results: Mathematical modelling predicted a shift towards a reactivation-driven, ageing TB epidemic in China by 2050. Vaccines protective against disease, effective post-infection and providing at least 5 years protection were essential for achieving higher levels of impact. Vaccination of older adults provided greater population-level impact than vaccinating adolescents for all equivalent vaccines explored, even if much lower coverage were achieved in older adult vaccination. Recommendations for post-infection vaccines were robust to substantial reductions in efficacy and duration of protection in older adults, whereas for pre-infection vaccines in some of these scenarios adolescent vaccination may be equivalent or preferred. Vaccinating older adults with post-infection vaccines provided substantially higher impact than pre-infection vaccines.

1,899 TB patients were registered using ePAL in the 12-month study period, with high patient acceptance (98.7%, 1,899/1,924). ePAL achieved clinic-based collection of patient location of residence accurate to a median of 84 metres (IQR: 35-317 metres) in a high population density urban setting without a municipal address system. Advantages of the ePAL system included real-time availability of high-resolution spatial data, low set up costs, and ease of use by health staff as part of routine TB registration. Data were used to identify areas with high TB burden, potentially suitable for TB vaccine trials.

Conclusions: The research presented in this thesis informs the development of appropriate TB vaccines and target populations to maximise future population-level impact. A prevention of disease vaccine efficacious post-infection and delivered to

older adults would contribute towards maximising population-level impact in China. Adolescent-targeted tuberculosis vaccines are likely to have low impact in ageing, reactivation-driven epidemics like China, which suggests a modification of the current strategic focus on adolescents among certain funders. Clinical trials should assess disease endpoints, include *M.tb*-infected and older adult populations, and extend beyond the usual 2-3 years follow up. To support design of disease endpoint trials, ePAL may provide an accurate, easily implementable, low-cost tool for identification of areas of high TB burden in settings without addresses.

Table of Contents

List of Tables

List of Figures

Abbreviations

App	Application
ARV	Anti-retroviral
ART	Anti-retroviral therapy
BCG	Bacillus-Calmette-Guérin
BMGF	Bill and Melinda Gates Foundation
CDP	Clinical Development Plan
CDR	Case Detection Ratio
CEA	Census Enumeration Area
DOTs	Directly Observed Treatments
eCRF	electronic Case Report Form
EDC	Electronic Data Capture
ePAL	electronic PArticipant Locator
GPS	Global Positioning System
HBC	High Burden Country
HIV	Human Immunodeficiency Virus
HSA	Health Surveillance Assistants (Community Health Workers)
IGRA	Interferon Gamma Release Assay
LSHTM	London School of Hygiene and Tropical Medicine
LTBI	Latent TB infection
MLW	Malawi-Liverpool-Wellcome Trust
NTP	National Tuberculosis Programme
PI	Principal Investigator
PII	Personally Identifiable Information
PLHIV	People living with HIV
POI	Point/Place(s) Of Interest
PPC	Preferred Product Characteristic
P&PI	Pre- and post-infection
PRI	Pre-infection
PSI	Post-infection

SES	Socio-Economic Status
SSA	sub-Saharan Africa
TB	Tuberculosis
TST	Tuberculin Skin Test
TPP	Target Product Profile
UR	Uncertainty Range
VE	Vaccine Efficacy
VE-POD	Vaccine Efficacy for Prevention of Disease
VE-POI	Vaccine Efficacy for Prevention of Infection
VE-POI&D	Vaccine Efficacy for Prevention of Infection and Disease
WHO	World Health Organization

Acknowledgements

This PhD would not have been possible without the help and support of many colleagues, collaborators, friends and family.

Firstly, my supervisors Prof. Richard White and Dr Tom Sumner, to whom I owe an eternal debt of gratitude for their guidance, patience and belief in me. I have learnt so much from you academically, professionally and personally. Your passion for the work you do is an inspiration, and shines through in the day-to-day of your work, making this PhD a pleasure to have undertaken. I am truly thankful for the opportunities and guidance you have given, and look forward to continuing to work together.

I would like also to thank the members of my advisory committee (Liz Corbett, Emily Webb, Rein Houben, Tom Evans, Vicky Cardenas and Bernard Fritzell) for technical advice, valuable insight, and willingness to share their wealth of knowledge to help guide and improve my research. Bernard, Vicky, Rein and Tom each brought their own brand of infectious enthusiasm to discussions about the work throughout the PhD, and have always helped remind me of why the work that we do is important. A special thanks to Emily for your kindness and patience with all of my questions. And most of all Liz, for welcoming me in to her research group (and on several occasions her home) in Blantyre. It was truly a pleasure to work with you and your team. I would also like to thank our collaborators at Aeras, WHO, Gates Foundation and TBVI for their willingness and enthusiasm to engage with this work to ensure it met the needs of the TB vaccine community, and many of whom took roles on my advisory committee.

Many thanks to the LSHTM TB modelling group - throughout the PhD it has made a big difference to feel part of a supportive team. Similarly, to the TB and Vaccine Centres for making me feel part of the broader school environment, and providing opportunities to work with so many excellent academics from across the school.

Eternal thanks go to John Edmunds, Suzanne Filteau and the MRC, without whom I would not have had a funded place here at the school.

On a more personal note, my friends at the School, PPM, around London, champ. international, and around the world have been the best of cheerleaders helping me through the PhD, many thanks for always believing in me.

And finally, but most importantly, to my dearest Mob, Caroline, Mum and Dad. I can't find the words to thank you enough for supporting my decision to take on a PhD, and providing the love, patience, proofreading and home cooked meals that I needed to make it through, I would never have made it across the finish line without all of you there to support me. My parents have always instilled in me the importance of following my passions, and have provided the love and support throughout my life to enable me to do just that. Without you I wouldn't be here writing this acknowledgement, so this PhD is dedicated to you.

CHAPTER 1: Introduction

Summary of Chapter 1

This chapter defines the context and rationale for the research presented in this thesis. This includes a discussion of tuberculosis as a public health problem, and the potential role of new vaccines in tackling the burden of disease within the context of WHO 2035 and 2050 goals. Two key epidemiological research needs are identified, and the rationale for the development of new mathematical models and an epidemiological tool to help inform and accelerate development of TB vaccines is described. The two overall aims of the thesis are defined, and translated into research objectives.

An outline of the thesis to is provided help guide the reader, including a brief summary of the chapters and research papers.

CHAPTER 1 Introduction

1.1 Background

The global Millennium Development Goal to halt and reverse tuberculosis (TB) incidence trends by 2015 was achieved, yet with an estimated 10.4 million incident TB cases and 1.8 million deaths in 2015, a substantial burden of disease still exists.[1] The current trajectory of incidence decline (1.5%/year) will not achieve the WHO 2050 goal of eliminating TB as a public health problem, defined as less than 1 case per million population per year. Historical trends and mathematical modelling indicate that with socioeconomic development and aggressive scale-up of existing interventions, up to 10% annual reduction could potentially be achieved,[2-4] but falls short of the 20% per year required for elimination (Figure 1.1).[5] Development of new tools, such as new TB vaccines, will be essential for achieving these goals.[5]

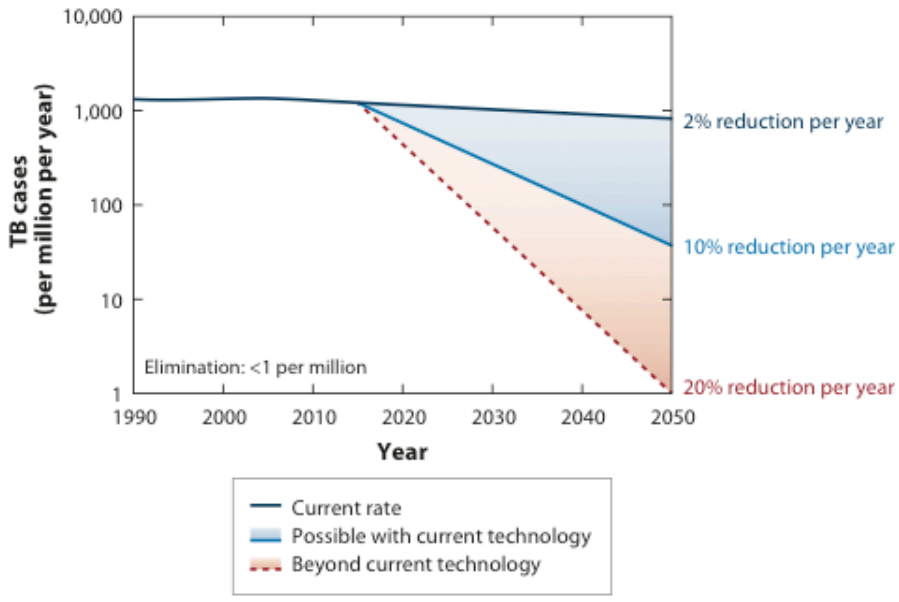


Figure 1.1: Reduction in incidence rate required to achieve WHO 2050 goal of elimination of TB as a public health problem. Reproduced with permission from Dye et al. 2013.[5]

The existing TB vaccine, Bacillus Calmette-Guérin (BCG), is delivered to more than 88% of BCG-eligible newborns, totalling more than 100 million newborns annually.[6-8] Meta-analyses suggest that in children vaccinated at birth, BCG efficacy is 59% (95% confidence interval (CI), 42-71%) against pulmonary TB and 90% (95% CI, 23-99%) against extrapulmonary disease.[9] However, highly variable protection has been observed in adults, ranging by setting from 0% to 80% against pulmonary disease.[9] Given that 90% of global TB incidence occurs in the adult (≥15 years) population, there is a clear and urgent need for a new prophylactic TB vaccine effective in adults.[1]

The new TB vaccine pipeline is the strongest to date, with 13 candidates currently in clinical trials.[10] Yet progression of vaccines through the pipeline remains slow, due to limited funding and the technical challenges in conducting clinical trials. Particularly limiting is the lack of an immune correlate of protection and relatively low incidence of disease, necessitating large efficacy trials with long periods of follow up, without prior knowledge of whether the elicited response is likely to be protective.

Key activities to help overcome some of the challenges in TB vaccine development have been identified in the TB Vaccine Blueprint.[11] Included in the Blueprint activities was the need to "determine TB prevalence and incidence, select trial sites and choose target populations for TB vaccines that result in the greatest reduction of disease".[11] This constitutes two epidemiological research needs, summarised in Box 1 below.

Box 1: Epidemiological research needs, based on the TB Vaccine Blueprint[11]

1) Data-driven identification of vaccine characteristics and target populations for vaccination that could achieve the greatest reduction in disease burden.
2) A better understanding of local epidemiology to inform the selection of trial sites, recruitment populations and design clinical trials.

For the first research need, mathematical modelling provides a data-driven framework for estimating the potential impact of vaccine characteristics and vaccination strategies, aiding identification of those likely to achieve the greatest

reduction in burden ('ideal' candidates). Where 'ideal' candidates are not feasible, modelling can help identify characteristics that would provide a minimum desirable impact. This research is needed to inform development of target product profiles (TPPs) for TB vaccines. TPPs are strategic documents outlining the desired and minimum vaccine label. Vaccine characteristics and implementation strategies considered in the TPP include, but are not limited to, the vaccine efficacy for prevention of infection or disease, whether the vaccine is effective in populations uninfected or infected with *Mycobacterium tuberculosis* (*M.tb)*, duration of protection, and target populations that could maximise impact (e.g. age groups).

Existing modelling literature explores the epidemiological impact of a variety of these vaccine characteristics in global, regional and country-level settings, but in any given study typically only a subset of characteristics was explored, limiting use for informing vaccine development. A comprehensive exploration of possible vaccine characteristics and implementation strategies using mathematical modelling is required.

Results from a phase III prophylactic TB vaccine trial in China are anticipated this year.[12] As more information becomes available about the characteristics of this vaccine, mathematical modelling can be used to predict future impact. China is the third largest contributor to global TB incidence,[1] so opportunities to maximize the vaccine-avoidable disease burden in this setting could also have a sizeable impact on the global burden. Available TB vaccine modelling in China is limited, exploring relatively unrealistic profiles (for example assuming 100% efficacy and lifelong protection),[5,13] and overlooking the dynamic co-evolution of demographics and TB epidemiology in this ageing population. Although global-level modelling has explored age targeting of vaccination to infants versus adolescents and adults,[14,15] given the ageing populations in several of the WHO high TB burden countries (e.g. China, Russia, Brazil and Thailand),[1,16] it is surprising that vaccinating older adults or the elderly has not been explored. A significant research gap clearly exists in fully

exploring the impact of vaccine characteristics and age targeting of vaccination in this setting.

Target product profiles, ideally informed by mathematical modelling as described above, are translated in to clinical development plans (CDPs) designed to develop appropriate candidates ready for delivery to identified target populations. As highlighted in the TB vaccine Blueprint's second epidemiological research need, designing such trials, selecting trial sites, and identifying recruitment populations all require appropriate epidemiological data.

There is a substantial global burden of TB disease, but relatively low incidence rate compared to other diseases such as malaria, therefore clinical trials for prevention of TB disease can be long, large, expensive, and challenging to design. As has been discussed in the literature, there is a dearth of trial-ready research sites with known sufficient TB burden for vaccine trials.[17] Country and regional level programmatic tuberculosis data, and prevalence surveys where available, can be used to identify regions experiencing a trial-suitable burden of disease. However, data of greater granularity and specific to communities where recruitment is planned is needed. Such data inform sample size calculations that balance the risk of failure from insufficient endpoints against minimising costs and the number of participants exposed to the investigational product. Such data can also help direct recruitment through identification of geographical hotspots in the intended recruitment populations (e.g. HIV-negative populations or specific age groups).

However, data with such granularity are rarely available. Prospective prevalence or incidence studies are considered the gold standard, but are expensive, inflexible once initiated and can take months to years from inception to delivery of results. Readily available TB notification data can be collected from national tuberculosis programme (NTP) clinic records, but clinic catchment areas may be large, descriptions of residence unclear, and patients may not always seek care in their local community due to stigma associated with the disease. In many low- and middle-income countries

where formal address systems are not in place, verbal descriptions of place of residence are collected in clinic records (e.g. 'Near Maketa market, by the bus station look for the maize mill and ask for Mr X'). However, a study in Malawi has highlighted the difficulty in accurately identifying place of residence from such descriptions, with two research assistants only able to identify the cluster (mean cluster size: 1,342 adults) in which the participant was resident for 40.3% and 68.2% of 129 ART-initiators.[18] Therefore, existing data may be of limited use for improving granularity of available data for study design.

In the same research site in Malawi, a low cost "Map book" was developed for identification of place of residence, and was found to improve the accuracy of cluster residency identification to 97% (95% CI, 90-100%). Already this rapid and low cost approach could greatly improve the granularity of TB notification data in low resource, high burden settings. However, this concept could be further developed, with the aim of providing granularity closer to the household level, combined with information regarding patient characteristics and real-time reporting. Such a validated, low cost, rapidly implementable tool could meet the need for improved data to inform TB vaccine clinical trial design and recruitment of TPP-identified populations.

1.2 Rationale for the present study

Development of new TB vaccines to reduce the burden of adult TB disease is a clear and urgent public health need. However, to inform TPPs for new TB vaccines, research to fully explore the vaccine characteristics and target populations that could provide the greatest epidemiological impact is required. In terms of target populations, given the ageing demographics in several high burden countries (e.g. China), comparing vaccinating the currently prioritised adolescent population to vaccinating older adult populations is an important research gap. To help carry these TPPs through to implementable clinical trials, appropriate data are needed to inform trial design and identify recruitment population hotspots. Given the expense and

time requirement of prospective prevalence and incidence studies, development of a low-cost, rapidly implementable tool to be able to identify hotspots in these target populations for recruitment could improve TB vaccine trial design and reduce costs.

This thesis describes mathematical modelling to estimate the potential impact of new TB vaccines, and identifies vaccine profiles and target populations that could result in the greatest reduction in TB disease burden in China; and develops a new methodology for empirical data collection for determining spatial distribution of TB notifications in potential trial populations.

Overall, this research will provide an important contribution towards fulfilling the epidemiological research needs to ensure appropriate TB vaccines are developed for and delivered to populations that could maximise population-level impact of new TB vaccines.

1.3 Thesis aims

In order to facilitate data-driven development strategies for new TB vaccines, the overall aims of this research were to estimate the potential epidemiological impact of new TB vaccine characteristics and target populations in China, to identify those predicted to provide the greatest reduction in TB disease; and to generate a novel epidemiological spatial mapping tool capable of informing clinical trial design for new TB vaccines in low-income, high-burden settings with limited availability of relevant data.

This translated into two streams of work, each with a defined aim:

1. Generate mathematical models of age-stratified demography and TB epidemiology to explore the population-level epidemiological impact of vaccine characteristics and implementation strategies for potential new TB vaccines, using China as a case study.

2. Development and evaluation of a low-cost, rapidly implementable TB spatial mapping application ('app') in Blantyre, Malawi, capable of mapping the place of residence of populations experiencing disease outcomes relevant to TB vaccine clinical trials.

1.4 Thesis objectives

Aim 1 (modelling to inform vaccine characteristics and implementation) was achieved through the following objectives:

1. Conduct a systematic review of the TB vaccine mathematical modelling literature to a) summarise existing literature exploring the effect of vaccine characteristics and implementation on the potential epidemiological impact of new TB vaccines, and b) identify research needs with respect to mathematical modelling of possible vaccine characteristics and implementation strategies for new TB vaccines (Chapter 2).

2. a) Develop a mathematical model calibrated to epidemiological and demographic temporal and age distribution trends in China, accounting for uncertainty in natural history parameters, to predict the evolution of the TB epidemic over the 2050 time horizon (Chapter 3).

   b) Using the calibrated China TB model, simulate the introduction of new TB vaccines to investigate the population-level epidemiological impact of varying TB vaccine characteristics identified in objective 1 (Chapter 3).

c) Identify the combination of vaccine characteristics that would be most likely to deliver pre-specified minimum incidence rate reductions compared to the no new vaccine baseline in 2050 (Chapter 3).

3. Using the calibrated model of the TB epidemic in China and main vaccine characteristics (objective 2), investigate the impact of age-targeted vaccination programmes to adolescents versus older adults on population-level TB epidemiology (Chapter 4).

Aim 2 (development of epidemiological tool for informing future trial design) was achieved through the following objectives:

4. a) Develop and implement a novel, low resource, rapidly implementable and easy to use app for spatial mapping of the place of residence of TB cases registering for treatment in Blantyre, Malawi (Chapter 5).

   b) Evaluate the accuracy of patient coordinates measured via the app compared to the gold standard of measurements taken using GPS at the patient's place of residence (Chapter 5)

   c) Based upon place of residence collected using the app, generate spatial maps of disease numbers and rates in populations potentially relevant for TB vaccine trial recruitment. (Chapter 5).

1.5 Thesis outline

This is a research paper style thesis, submitted in accordance with London School of Hygiene and Tropical Medicine regulations. As such, submitted and published articles are included without adaptation; therefore, although minimised where possible, some repetition exists between chapters. Chapters containing research papers are structured around the manuscript, starting with an introduction to provide context for the paper within the thesis, followed by the manuscript, and finally any

supplementary materials associated with the paper and/or unpublished additional analyses and discussion.

The thesis consists of four sections, and contains six chapters, which include one published, one submitted, and one unpublished manuscript. The first section (Chapters 1 and 2) comprises the background materials, aims and objectives of the work. Section two (Chapters 3 and 4) reports on the mathematical modelling research conducted to inform characteristics and implementation of new TB vaccines in China. Section three (Chapter 5) reports the development, validation and results of the novel spatial mapping tool for informing TB vaccine trial recruitment. Finally, section four (Chapter 6) provides discussion, conclusions, policy implications and future work. A more detailed outline of each chapter follows.

Chapter 1: This first chapter provides a brief overview of the epidemiological and TB vaccine development context within which this research resides, and the research needs identified. These provide the rationale for the aims of this PhD and the objectives identified to achieve those aims.

Chapter 2: The background to the thesis includes an overview of tuberculosis disease burden, the need for new TB vaccines, and for new tools and data to provide the evidence base for development and implementation planning for new TB vaccines. This chapter includes ***research paper 1***, a systematic review published in Human Vaccines and Immunotherapeutics, summarising the modelling literature of the epidemiological impact of new TB vaccines and implementation strategies (*objective 1*).[19] This is followed by a summary of the current options for epidemiological data collection tools to inform trial recruitment of populations. Research gaps in both of these areas are identified.

Citation (paper 1): Harris RC, Sumner T, Knight GM, and White RG. Systematic review of mathematical models exploring the epidemiological impact of future TB vaccines. *Hum Vaccin Immunother* 2016; **12**(11): 2813-32.

Chapter 3: This chapter describes the development of the mathematical model of age-stratified TB demographics and epidemiology in China. Modelled epidemiological projections to 2050 without introduction of new TB vaccines are described (*objective 2a*). Using this calibrated epidemiological model, the population-level epidemiological impact of a wide range of vaccine characteristics was explored in depth, and incidence and mortality results in terms of absolute and rate reductions in incidence and mortality are reported (*objective 2b and c*).

Chapter 4: This chapter reports a further adapted version of the China vaccine model, exploring the impact of vaccination of adolescents (15-19 years) versus older adults (60-64 years) with new TB vaccines (*objective 3*). A subset of vaccines, selected based upon the results of the modelling in chapter 3, are explored in this chapter. This work is reported as ***research paper 2***, currently under review at the Lancet Global Health.

Citation (paper 2): Harris RC, Sumner T, Knight GM, Evans T, Cardenas V, Chen C, and White RG. New tuberculosis vaccines – the impact of age targeted vaccination in China and implications for vaccine development. *Under review*

As a body of work, these models give insight in to the characteristics and implementation strategies for new TB vaccines that need to be developed to maximise population level impact. The results are indicative of populations and endpoints important for clinical trials, which are the focus of the next section of the thesis. This next section describes the tool developed for mapping key populations experiencing outcomes relevant to TB vaccine clinical trials.

Chapter 5: This chapter reports the development and evaluation results (*objective 4*) of ePAL (electronic PArticipant Locator), a novel, low cost, rapidly implementable and easy to use app for spatial mapping of the place of residence of TB cases registering for treatment in Blantyre, Malawi. The development and evaluation results are reported as ***research paper 3***. The results of 12 months data collection by NTP TB officers using ePAL in Blantyre, Malawi are summarised in this chapter. This includes

spatial heatmaps (*objective 4c*) of all TB, and demographic and clinical subgroups relevant to TB trial recruitment populations or outcomes.

Citation (paper 3): Harris RC, Kaswaswa K, Choko AT, Molineux A, MacPherson P, Webb E, Shonga W, White RG*, and Corbett EL*. Development and evaluation of novel software to identify place of residence for clinic-based TB patients in Blantyre, Malawi. *(proposed authors, author order and title. In review with co-authors. *Joint senior authors)*

Chapter 6: This chapter summarises the research undertaken in the thesis and the implications of the results for the TB vaccine development community. Key findings, strengths, limitations and contributions to the field are discussed. I provide recommendations for development of TB vaccines from the perspective of vaccine developers, clinical triallists and country-level decision makers. Opportunities for future research are identified and discussed.

1.6 Contribution of the author

The need for mathematical modelling research to support TB vaccine development was identified, and initial research questions developed, by myself and Professor Richard White. These research questions evolved during the PhD based upon research needs identified from the systematic literature review, and feedback by stakeholders and collaborators (including Aeras, BMGF, WHO and TBVI). From previous work, I had identified the need for new solutions for generating epidemiological data to support TB vaccine clinical trial design, which was fortuitously married with the idea for an electronic mapping tool proposed by Professor Liz Corbett and Dr Peter MacPherson.

Contributors to the research presented in the three research papers are detailed in their respective cover pages. For all three papers, I wrote the first draft and incorporated changes based upon co-author review.

1.7 Funding

This PhD was funded by the UK Medical Research Council (MRC) under the London School of Hygiene and Tropical Medicine MRC vaccines scholarship program. The work presented in Chapter 3 was partially funded by the Bill and Melinda Gates Foundation, and funding for the work in Chapter 4 was received from Aeras, a non-profit TB vaccine product development partnership. Tool development and field work presented in chapter 5 was co-funded by the author's MRC scholarship and Professor Liz Corbett's Wellcome Trust grant.

1.8 Chapter 1 References

1. World Health Organization. Global Tuberculosis Report 2016. 2016. http://apps.who.int/iris/bitstream/10665/250441/1/9789241565394-eng.pdf?ua=1 (accessed 10th May 2017).
2. Marais BJ, Raviglione MC, Donald PR, et al. Scale-up of services and research priorities for diagnosis, management, and control of tuberculosis: a call to action. *Lancet (London, England)* 2010; **375**(9732): 2179-91.
3. Dye C, Garnett GP, Sleeman K, Williams BG. Prospects for worldwide tuberculosis control under the WHO DOTS strategy. *Lancet (London, England)* 1998; **352**(9144): 1886-91.
4. Fletcher HA, Schrager L. TB vaccine development and the End TB Strategy: importance and current status. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 2016; **110**(4): 212-8.
5. Dye C, Glaziou P, Floyd K, Raviglione M. Prospects for tuberculosis elimination. *Annu Rev Public Health* 2013; **34**: 271-86.
6. World Health Organization. Reported estimates of BCG coverage. 2015. http://apps.who.int/immunization_monitoring/globalsummary/timeseries/tscoveragebcg.html (accessed 23rd December 2015).
7. Bao QS, Du YH, Lu CY. Treatment outcome of new pulmonary tuberculosis in Guangzhou, China 1993-2002: a register-based cohort study. *BMC Public Health* 2007; **7**: 344.
8. World Health Organization. WHO/UNICEF Estimates of National Immunization Coverage (WUENIC). 20 Oct 2016 2016. http://www.who.int/immunization/monitoring_surveillance/data/en/ (accessed 03 July 2017.
9. Abubakar I, Pimpin L, Ariti C, et al. Systematic review and meta-analysis of the current evidence on the duration of protection by bacillus Calmette-Guerin vaccination against tuberculosis. *Health technology assessment (Winchester, England)* 2013; **17**(37): 1-372, v-vi.
10. Aeras. The Aeras Annual Report 2015. 2015. http://www.aeras.org/annualreport2015 (accessed 20th January 2017).
11. Brennan MJ, Thole J. Tuberculosis vaccines: a strategic blueprint for the next decade. *Tuberculosis (Edinburgh, Scotland)* 2012; **92 Suppl 1**: S6-13.
12. Anhui Zhifei Longcom Biologic Pharmacy Co. Phase III Clinical Study of Efficacy and Safety of Vaccae™ to Prevent Tuberculosis. 27th December 2016. https://clinicaltrials.gov/show/NCT01979900 (accessed 3rd January 2017).
13. Liu S, Li Y, Bi Y, Huang Q. Mixed vaccination strategy for the control of tuberculosis: A case study in China. *Math Biosci Eng* 2017; **14**(3): 695-708.
14. Knight GM, Griffiths UK, Sumner T, et al. Impact and cost-effectiveness of new tuberculosis vaccines in low- and middle-income countries. *Proc Natl Acad Sci U S A* 2014; **111**(43): 15520-5.
15. Arregui S, Sanz J, Marinova D, et al. A data-driven model for the assessment of age-dependent patterns of Tuberculosis burden and impact evaluation of novel vaccines. *bioRxiv* 2017: Online first.

16. The World Bank. Population ages 65 and above as a percentage of the total population. 2017. http://data.worldbank.org/indicator/SP.POP.65UP.TO.ZS (accessed 31st July 2017).
17. McShane H. Need for more TB vaccine field sites. *Indian journal of experimental biology* 2009; **47**(6): 445-6.
18. MacPherson P, Choko AT, Webb EL, et al. Development and validation of a global positioning system-based "map book" system for categorizing cluster residency status of community members living in high-density urban slums in Blantyre, Malawi. *American journal of epidemiology* 2013; **177**(10): 1143-7.
19. Harris RC, Sumner T, Knight GM, White RG. Systematic review of mathematical models exploring the epidemiological impact of future TB vaccines. *Hum Vaccin Immunother* 2016; **12**(11): 2813-32.

CHAPTER 2: Background

Summary of Chapter 2

In this chapter, I provide a detailed exploration of the research context, beginning with an outline of tuberculosis as a global public health problem, the need for new TB vaccines, and for new tools and data to provide the evidence base for development and implementation planning for new TB vaccines. This chapter includes ***research paper 1***, a systematic review published in Human Vaccines and Immunotherapeutics, summarising the modelling literature of the epidemiological impact of new TB vaccines and implementation strategies (*objective 1*).[1] This is followed by a summary of the current options for epidemiological data collection tools to inform trial recruitment of populations. Research gaps in both of these areas are identified (*objective 1*).

Citation (research paper 1): Harris RC, Sumner T, Knight GM, and White RG. Systematic review of mathematical models exploring the epidemiological impact of future TB vaccines. *Hum Vaccin Immunother* 2016; **12**(11): 2813-32.

This chapter is derived in part from an article I published in *Human Vaccines and Therapeutics* on 22nd July 2016, available online: http://wwww.tandfonline.com 10.1080/21645515.2016.1205769.

CHAPTER 2 Background

2.1 Tuberculosis natural history

The vast majority of tuberculosis (TB) disease in humans is caused by the bacterial pathogen *Mycobacterium tuberculosis* (*M.tb*).[2] There are seven other related *Mycobacterium* species, some of which are also capable of causing disease to varying degrees.[2] TB disease symptoms depend upon factors such as the site of infection, but common symptoms include night sweats, weight loss, chronic cough and haemoptysis.[3] Confirmation of TB disease is most commonly via one or a combination of chest x-ray, sputum smear testing, culture, and Genexpert nucleic acid amplification test.

M.tb transmission occurs via the airborne route, when infected droplet nuclei expulsed (e.g. coughing, shouting) from the lungs of an individual with pulmonary or laryngeal TB disease are inhaled into the lungs of another individual. The likelihood of a transmission event is dependent on the infectiousness of the individual with TB disease,[4] susceptibility and extent of exposure of the exposed individual, and environmental factors.[5] Infectiousness can be modified by the location of disease foci (i.e. pulmonary vs. extrapulmonary), bacterial load, comorbidities (e.g. HIV), and clinical symptoms such as cough; whereas susceptibility is mostly associated with co-morbidities of the exposed individual (e.g. HIV, malnutrition) and other exposures that may weaken the lungs (e.g indoor air pollution, smoking).[6-8] Exposure to infected droplets is related to proximity and duration of exposure to infected individuals.[5] The persistence of infected droplet nuclei in the air may be influenced by abiotic factors such as ventilation, humidity and exposure to ultraviolet light.[9] In combination, these factors determine the likelihood of a transmission event.

Unlike many infectious diseases that have relatively short incubation periods (e.g. cholera, influenza), *M.tb* infection can remain latent for extended periods of time.[10] In fact, the majority of individuals infected with *M.tb* never develop disease. The age-weighted lifetime risk of progression to pulmonary disease has been estimated at

12% in low transmission settings.[10] This risk is elevated in certain populations, for example the risk of progression in HIV co-infected populations is estimated at approximately 10% per year.[11] The risk of developing disease also varies by age at infection and time since infection, with incubation periods becoming shorter in older age groups.[10,12] Risk of progression to disease is highest in the first year after infection, falling year-on-year beyond this.[11-14]

Rapid primary ("fast") progression to disease occurs when bacterial replication produces symptomatic disease relatively rapidly after entry into the lungs; whereas reactivation ("slow") progression is characterised by an extended period in the latent state. Latency involves the establishment of granuloma, formed when *M.tb* bacteria replicate intracellularly in alveolar macrophages and initiate cytokine recruitment of both innate and adaptive immune cells.[15] Eventually a state of equilibrium is attained between dormant *M.tb* bacteria contained within the granuloma core and the host immune system. Disruption of this equilibrium leads to bacterial replication and development of "reactivation" active disease.[11]

Pulmonary disease is the most frequent presentation. Given the airborne route of *M.tb* transmission, pulmonary cases are the most infectious, particularly those testing positive by sputum smear microscopy. In 2015, approximately 15% of notified TB cases occurred at extrapulmonary sites, such as miliary TB and TB meningitis.[16] Though the proportion of extrapulmonary disease is higher at the age extremes, due to immature immunity in children and immunosenescence in the elderly.[16,17]

Without treatment, smear negative 10-year case fatality is estimated at 20%, increasing up to 70% in smear positive disease.[18] Including those receiving treatment and those remaining untreated, the global case fatality ratio in 2015 was estimated at 17%, ranging from 5% to >20% by country.[16]

2.2 Tuberculosis as a global public health concern

Tuberculosis disease (TB) has become the largest single-pathogen cause of mortality globally,[19] responsible for 1.8 million deaths in 2015. During the same period, there were an estimated 10.4 million incident TB cases, including 480,000 multi-drug resistant (MDR) cases. Although the current burden of disease remains the most immediate public health concern, a substantial reservoir of latent TB infection (LTBI) exists, estimated at approximately 1.7 billion of the global population.[20] This population remains at risk of developing reactivation disease, so should be considered when planning for TB prevention in the long term.

Regional heterogeneity exists in the distribution of TB burden. South-East Asia and sub-Saharan Africa contributed 46% and 26% of the global incident cases in 2015, respectively.[19] More specifically, six countries account for 60% of incident TB cases: India, Indonesia, China, Nigeria, Pakistan and South Africa. Populations in sub-Saharan Africa have experienced some of the highest TB incidence rates, reaching 834/100,000 population per year in South Africa.[19]

Heterogeneity also exists in the distribution of disease within populations, due to environmental, biological and social risk factors. Environmental risk factors include exposure to smoking or indoor air pollution, and collectively are thought to be a contributory risk factor for an estimated 38.1% of TB cases across the pre-2015 WHO 22 high burden countries (HBCs).[21] Key biological factors that weaken immunity and therefore increase the risk of TB infection and progression to disease include HIV, malnutrition and age-related factors. HIV exists as a co-morbidity in 11% of TB cases globally,[19] and is estimated to be associated with the highest reported relative risk (RR) of disease (RR: 26.7).[21] For some countries in Africa, the population attributable fraction (PAF) for HIV is estimated at greater than 50% of cases. Although the relative risk of TB associated with malnutrition is much lower (RR: 3.2) than HIV, due to the prevalence of malnutrition across the 22 HBCs, it is associated with the highest PAF (26.9%). Immunological changes with age also influence the risk of disease, with immune immaturity in infants and immune senescence in older adults and the elderly

increasing the risk of development of primary or reactivation disease.[17] However, given the population demographic structure at the global level, and the burden of HIV co-infection in younger adults, 90% of TB cases globally are estimated to occur in adults.[19]

Social factors also play an important role in the risk of infection, as contact patterns with groups with higher disease rates, such as particular age or social groups, can influence the risk of exposure and therefore infection. In addition, time spent in particular settings with crowding or poor ventilation can increase the risk of infection.[5] This is particularly pronounced in prison populations or crowded places of residence, and may also be associated with crowded social locations such as workplaces or churches.[22]

2.3 Global TB targets

Since the WHO 1993 declaration identifying tuberculosis as a "public health emergency of international concern", significant investments have been made in the global fight against tuberculosis,[23] with funding almost doubling in the last decade, reaching USD 6.5 billion in 2016.[19]

The WHO's commitment to tackle the burden of TB was enshrined in Millennium Development Goal (MDG) 6c, aiming to "halt and begin to reverse the incidence of TB by 2015". This principal global goal was accomplished ahead of time in 2012.[24] Concurrently with the MDGs, additional targets were set by the Stop TB Partnership, aiming for a 50% reduction by 2015 in both TB prevalence and mortality rates compared to 1990 levels, and the longer-term target of elimination of TB as a public health problem by 2050 (<1 case per million population per year).[19] Globally, both targets were narrowly missed, with TB prevalence falling by 42% and TB mortality falling by 47%.[19]

Although significant progress has been made at the global level, clear discrepancies exist in regional and country-level attainment of the MDG 6c and 2015 Stop TB

targets. TB incidence trends were falling at the regional level in all six WHO regions in 2015, but several countries had not achieved the reversal of trend, including six of the WHO 22 HBCs: Democratic Republic of Congo, Mozambique, Nigeria, Afghanistan, Pakistan and Bangladesh.[25] The Stop TB prevalence targets were met in the region of the Americas (AMR), South-East Asia (SEAR) and Western Pacific (WPR), but were missed in the remaining three regions and in 59% of the 22 HBCs. Mortality targets were attained in the same regions plus the Eastern Mediterranean region (EMR), but were missed in 50% of HBCs.[25] Despite meeting the 2015 goals,[25] India and China still contribute 27% and 9% of global burden, respectively. Consideration of these geographical differences will be important in tackling region- and country-level variation.

Post-2015, although the Stop TB target of TB elimination by 2050 still stands,[25] new milestones and targets have been developed in tandem with the UN sustainable Development Goals (SDGs) as checkpoints on the road to elimination (Table 2.1).[26] SDG 3.3 aims to "end the epidemic of… tuberculosis" by 2030, with the indicator for progress defined as 90% reduction in number of TB deaths and 80% reduction in incidence rate compared to 2015.[26]

Table 2.1: WHO End TB and Stop TB goals for reduction in epidemiological burden of TB from 2020 to 2050[19,25]

	End TB strategy milestones and targets[19,26]				Stop TB strategy[25]
Target	2020	2025	2030	2035	2050
Reduction in absolute number of TB deaths compared to 2015	35%	75%	90%	95%	-
Reduction in TB incidence rate compared to 2015	20%	50%	80%	90%	-
Incidence rate target	<85 per 100,000 population per year	<55 per 100,000 population per year	<20 per 100,000 population per year	<10 per 100,000 population per year	<1 per million population per year

Progress in reducing the global burden of TB has been slower than anticipated, with the annual rate of incidence decline estimated at 1.5% in 2015.[19] To achieve the 2020 milestone and 2025 target, an acceleration to 4-5% annual decline is required by 2020, increasing to 10%/year by 2025 (Figure 2.1).[26] Historical trends are indicative of the feasibility of declines of approximately 10% per year through socioeconomic development and implementation of current tools.[27,28] In modelling studies, annual reductions of 6% have been projected as feasible by implementation of the Global Plan,[29] though up to 10% per year could be considered feasible with aggressive scale up of existing interventions and improved social protection.[27,29,30]

Beyond 2025, assuming the previous goals are achieved, reductions of 17% per year will be required to meet 2035 targets.[30] Looking further ahead to the 2050 goals, a 20% decline per year from 2015 has been estimated as necessary to achieve elimination by 2050.[31] This is beyond the acceleration that has been achieved historically or has been estimated to be possible with existing measures in modelling studies, therefore novel tools will be essential to reach the WHO 2050 goal of elimination of TB as a public health problem.[31]

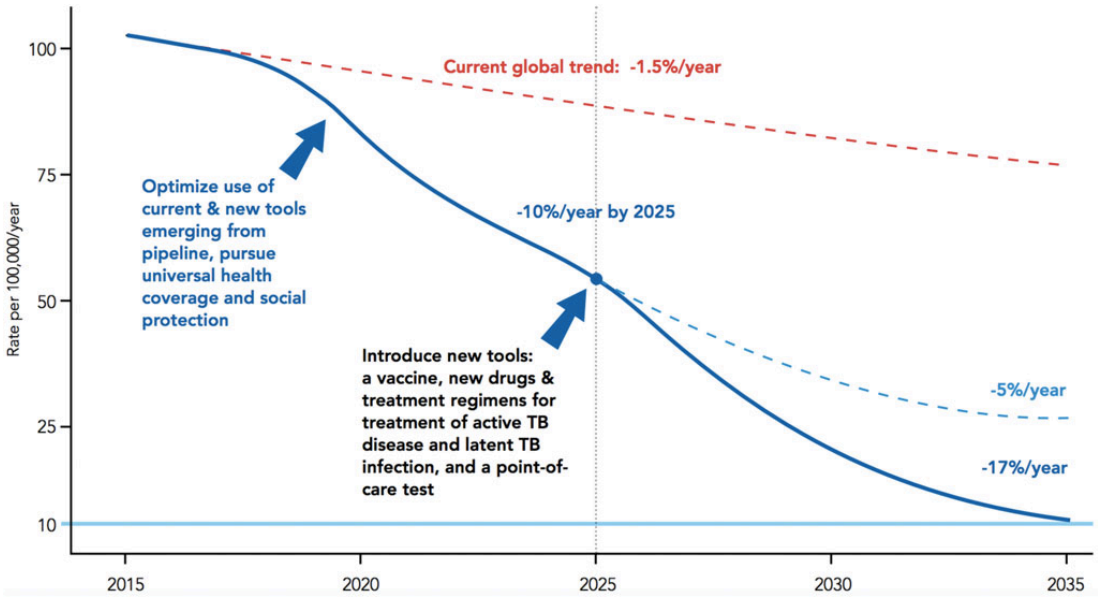


Figure 2.1: WHO global trends required to achieve 2025 and 2035 goals. Reprinted with permission, from the WHO End TB Strategy, WHO/HTM/TB/2015.19, http://www.who.int/tb/End_TB_brochure.pdf?ua=1, Accessed 3rd January 2017. [26]

New tools with the potential to contribute to TB care and prevention include new vaccines, new drugs and regimens, improved diagnostics, and measures to prevent reactivation disease in those latently infected.

With the current global burden of latent infection alone, TB incidence would be anticipated to decline only as far as 16.5 cases per 100,000 population per year by 2035 and 8.3 cases per 100,000 population per year by 2050;[20] therefore measures to reduce reactivation of latent infection will be important for achieving elimination. In settings with a significant pool of latent infection, the Abu-Raddad modelling study suggests that only mass isoniazid preventative treatment (IPT) or mass vaccination of this population with a new effective vaccine could come close to achieving the 2050 targets.[32] IPT is effective, but adherence issues and duration of protection may limit its potential impact, and widespread use as monotherapy could lead to development of resistance.[33] With the anticipation of longer-term protection and reduced adherence issues, vaccines could provide impact in a more sustainable manner. In another modelling study, results indicated that a combination of measures to tackle latent TB and active disease may be needed to achieve the WHO 2050 goals at the

global level.[31] The WHO suggests that "moving forward to the 2035 targets requires… effective pre- and post-exposure vaccines", in addition to better diagnostics, treatment regimens and LTBI treatment.[26]

2.4 BCG: an effective vaccine for prevention of TB?

Unlike many other high-burden infectious diseases, an effective vaccine exists for tuberculosis: Bacillus Calmette-Guérin (BCG). BCG is a live, attenuated *Mycobacterium bovis* vaccine, used for neonatal vaccination as part of the Expanded Program on Immunization (EPI) schedule. In 2015, global coverage was estimated at 88%, with the vaccine provided to over 100 million neonates annually to protect against childhood tuberculosis.[34] Recent meta analyses suggesting 59% (95% CI, 42-71%) vaccine efficacy against pulmonary TB,[35] 90% (95% CI, 23-99%) efficacy against miliary and meningeal TB,[35,36] and 66% (95% CI, 8-88%) efficacy against TB mortality.[36] BCG is estimated to prevent one case of miliary TB and one of TB meningitis in the first five years of life per 9314 and 3453 vaccinated children, respectively.[37] However, BCG has exhibited highly variable efficacy against adult pulmonary forms of tuberculosis, ranging from an "absence of clinically important benefit"[36] in studies in Malawi and India,[36,38,39] to almost 80% protection in the UK MRC trial and in a North American Indian population.[36,39,40] Many factors may influence these age and regional differences in efficacy, but prior infection or exposure to environmental bacteria has been observed as an important factor associated with lower efficacy against pulmonary TB.[35]

The protective effect of BCG has been demonstrated to last at least 10 years, with declining efficacy over time.[36] Five studies have demonstrated a measurable effect after 15 years, but loss to follow up and insufficient endpoints after 10 years limit the conclusions that can be drawn based upon available data beyond 10 years after vaccination.

BCG is contraindicated in HIV-infected populations due to the increased risk of disseminated BCG disease.[41] Therefore, HIV-infected children and adults at high risk of TB remain unprotected.

A combination of variable geographical efficacy against adult disease, insufficient duration of efficacy to protect from birth to adulthood, and safety concerns in epidemiologically important populations, mean that BCG is not an appropriate tool for tackling the global burden of adult TB disease. Given the need highlighted by existing models for a new vaccine to help achieve the WHO 2050 targets and these limitations in the use of BCG, the development of an efficacious novel TB vaccine that can protect in adulthood is a clear research priority.

2.5 The TB vaccine pipeline and development

Although a very clear and urgent need exists for new TB vaccines, particularly those able to reduce the burden of TB in adults, no new prophylactic vaccine has been licensed since BCG. Only one new candidate, the SRL-172 whole cell vaccine, has demonstrated safety and efficacy in phase III studies in HIV-positive populations,[42] but the agar-based production methodology was not scalable.[43] This candidate has since been adapted for broth-based production, re-entered clinical trials as candidate DAR-901, and is currently undergoing a phase IIB prevention of infection study.[44]

However, DAR-901 is not the only vaccine in development, with 13 candidates currently in clinical trials, the TB vaccine pipeline is the strongest to-date. One candidate is being tested in a phase III efficacy trial, two candidates in addition to DAR-901 in phase IIB proof of concept trials, four are undergoing phase IIA safety and immunogenicity studies, and five are in phase I first-in-human safety trials (Figure 2.2).[45]

The phase III *Vaccae* candidate is currently under evaluation for efficacy against disease in 10,000 participants in China, with results anticipated in 2017.[46] In addition,

the M72-AS01E and VPM1002 candidate phase IIB studies, and H4:IC31 phase IIA study, are expected to publish results within the next year.[47-49]

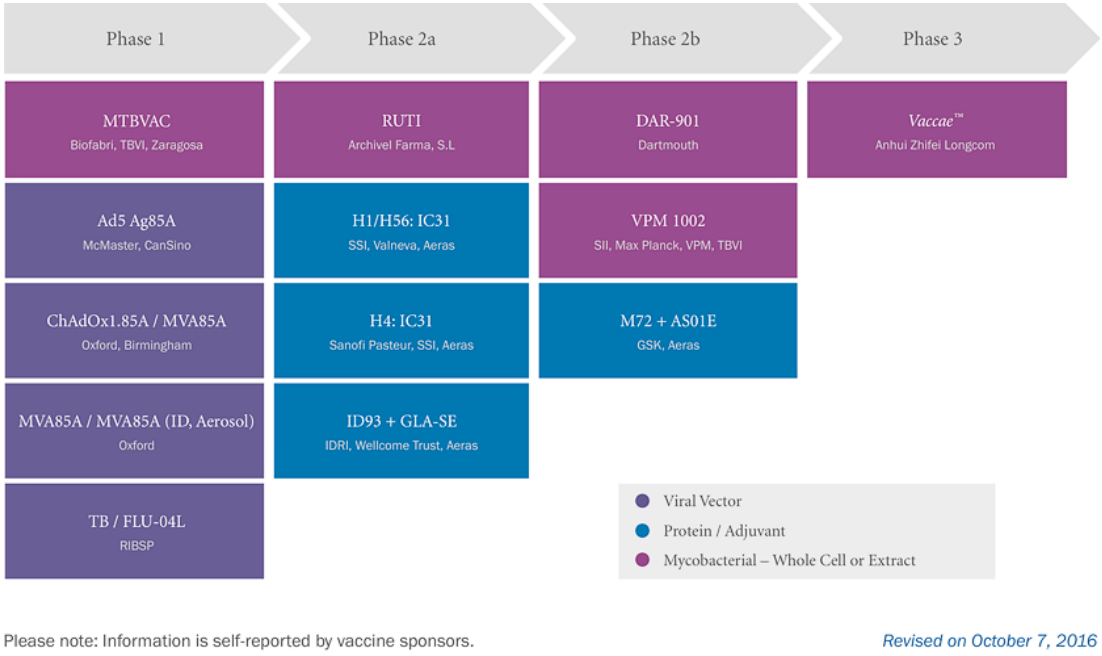


Figure 2.2: New TB vaccine pipeline as at 7th October 2016.[45] Reproduced with permission from Aeras.

Although a strong business case has been made for investment in new TB vaccines,[50] resources to develop these candidates remain limited. Progression through the pipeline is slow, in part due to limited and risk-averse allocation of funding, and in part due to relatively low incidence rates and lack of a correlate of protection necessitating large efficacy trials with long duration of follow up.

Critical research activities to overcome some of the challenges in TB vaccine development have been identified in the TB Vaccine Blueprint.[51] To support clinical trials a critical activity identified in the Blueprint was to "determine TB prevalence and incidence, select trial sites and choose target populations for TB vaccines that

result in the greatest reduction of disease".[51] This constitutes two epidemiological research needs:

1) Data-driven identification of vaccine characteristics and target populations for vaccination that could achieve the greatest reduction in disease burden.
2) A better understanding of local epidemiology to inform the selection of trial sites, recruitment populations and design clinical trials.

With several candidates progressing through phase IIB/III efficacy trials, research activities to fill these data needs are urgently required. In this thesis, I aim to use mathematical modelling to inform which vaccine characteristics and implementation strategies could help achieve greatest reduction in disease burden, and develop a novel tool to assist better understanding of local epidemiology to inform future clinical trial design and recruitment of the target populations such as those identified through mathematical modelling. In the following sections, I summarise the role and existing literature of mathematical modelling in TB vaccine development (section 2.6-2.10, and the tools available for collecting local epidemiological data (section 2.11-2.12), and identify the research gaps in these two areas leading to the specific research aims and objectives described in Chapter 1.

2.6 Informing TB vaccine Target Product Profiles

Mathematical modelling provides a data- and logic-driven framework for exploration of the epidemiological impact of varying vaccine characteristics and implementation strategies. With maximisation of the reduction of disease burden as the primary goal,[51] modelling at the country and global level can help inform decisions in vaccine development, clinical trial design including enrolment populations, and implementation strategy. This mostly occurs through the use of modelling to inform the vaccine Target Product Profile (TPP) – a summary of the desired eventual product labeling, critical for guiding efficient vaccine development strategy and ensuring development of products that are fit for purpose. TPPs may include, but are not limited to, the strategic targets for key vaccine characteristics, such as vaccine efficacy, duration of protection, safety profile and target population.

Development of a TPP requires consideration of three main factors: the technical feasibility, clinical trial feasibility and potential population-level impact of the vaccine profiles under consideration. The former are mostly informed by laboratory and clinical trial scientific knowledge. Epidemiological mathematical modelling to estimate the future impact of different vaccine profiles can provide the evidence base to identify vaccine profiles capable of achieving the required population-level impact. In TB vaccine development, mathematical modelling has been invaluable to estimate the potential future impact of new TB vaccines, forming the quantitative basis for refocusing the current TB vaccine development strategy towards vaccination of adolescents/adults.[52]

The TPP usually specifies the 'minimum' acceptable and the 'ideal' characteristics for the product in development. Modelling can inform the TPP 'minimum' characteristics by demonstrating those required to achieve the lowest acceptable level of impact, and the 'ideal' based upon the characteristics required to maximise impact, combined with consideration of technical feasibility. For TB vaccines, TPP characteristics that can be informed by modelling include duration of protection, vaccine efficacy, efficacy for prevention of infection and/or disease ('effect type'), whether the vaccine efficacy is modulated by the infection status or disease history of the vaccinated individual ('host infection status required for efficacy'), and the target population (e.g. age group) for vaccination.

The 'effect type' is defined as the natural history transition that the vaccine protects against: prevention of infection (POI), prevention of disease (POD), or both (POI&D). This is dependent upon the antigens included in the vaccine, and may also be dependent upon the vaccine platform (i.e. attenuated whole cell, subunit vaccine, viral vector). The anticipated effect type of a candidate will affect the trial endpoints during development, and the anticipated vaccine impact upon implementation. Precedent exists for either of these effect types, as although BCG is mostly considered a prevention of disease vaccine, some efficacy has also been shown against infection, with a meta-analysis of 14 studies (n=3,855) estimating BCG efficacy against infection

of 19% (95% CI, 8-29%). Therefore, BCG demonstrates that development of TB vaccines effective against infection or disease could be feasible.

The 'host infection status required for efficacy' is defined as whether the vaccine can induce a protective immune response in populations that have never been infected with *M.tb* (pre-infection), populations previously or currently infected (post-infection), or both (pre- and post-infection). Pipeline candidates have been trialed in both infected and uninfected populations, suggesting that both types of candidates are potentially in development. In terms of precedent, BCG is most effective as a pre-infection vaccine. A meta-analysis assessing BCG efficacy against pulmonary TB in school-age children estimated efficacy of 74% (95% CI: 63-81%) when only including those stringently tested for infection, whereas without stringent testing (i.e. inclusion of those likely to be latently infected) vaccine efficacy estimates declined to 45% (95% CI: -1 to 61%), with confidence intervals suggesting no significant clinical benefit.[35] It is unclear whether the mechanism of this reduced efficacy is masking or blocking by existing immune responses.

Given the potential for long latency periods of *M.tb* infection, vaccines effective post-infection could be a valuable TB prevention tool in this group at elevated risk of developing disease. Unlike diseases with short incubation periods such as measles or influenza, the duration of *M.tb* latency could allow sufficient time for identification and immunisation of infected individuals. Although a precedent for a post-infection vaccine does not exist in TB, vaccines for other long-latency diseases, such as Human Papilloma Virus, have demonstrated post-infection efficacy.[53,54] Some TB vaccine candidates contain antigens specific to the latent period (e.g. M72-AS01 vaccine).

The anticipated vaccine efficacy and duration of protection of future vaccines are unknown. However, BCG has demonstrated paediatric efficacy of 59-90% depending upon the outcome measure, therefore vaccines with efficacy within this range could be considered feasible. As discussed previously, BCG protection has been

demonstrated to last at least 10 years, but with new vaccines the type or robustness of the immune response induced could differ considerably from BCG.

Although the exact characteristics of the pipeline vaccines are unknown, there are clear differences in anticipated characteristics, as indicated by vaccine candidate antigens, trial endpoints, and trial recruitment criteria. For example, the M72-AS01 candidate contains possible latency-phase antigens, and the phase IIB study recruited Quantiferon (QFN)-positive adults (18-50 years) to study TB disease outcomes. Therefore, this vaccine is anticipated to be effective for prevention of disease (POD), in populations already latently infected with *M.tb* ('post-infection'). Conversely, the Aeras-404 candidate is currently being trialed in QFN-negative adolescents (12-17 years) to assess efficacy in preventing infection, as measured by QFN conversion. Hence this candidate is expected to be effective for prevention of infection (POI) and in populations with no history of *M.tb* infection ('pre-infection'). A summary of some of the key characteristics of the most recent trials of the candidates in phase IIB/III trials are summarized in Table 2.2 below. These trial design aspects are indicative of some of the possible anticipated characteristics and target populations for these vaccines.

Table 2.2: Key characteristics of current trial for candidates in phase IIB and III in July 2017

Vaccine candidate	Phase	Vaccine platform	BCG priming	Study outcome	Infection status of recruited population	Recruitment age group in current trial
Vaccae[46]	III	Whole cell (M.vaccae)	N/R	Disease	PSI	15-65 years
DAR-901[44]	IIB	Whole cell (M.vaccae)	Yes	Infection or disease	PRI	13-15 years
VPM-1002[49]	IIB	Whole cell (rBCG)	No	Recurrent disease	PRI (PSI planned)	Neonates (adults planned)
M72-AS01E[47]	IIB	Subunit	Yes	Disease	PSI	18-50 years

N/R: Not Reported, PRI: pre-infection, PSI: post-infection

2.7 Mathematical modelling to estimate population-level vaccine impact

The direction of impact of some of the described vaccine characteristics and implementation strategies can be logically predicted. For example, an inverse correlation would be anticipated between burden of disease and vaccine efficacy or coverage. To quantify such relationships, however, requires mathematical modelling. Such information can be used in TPP decision making to estimate vaccine efficacies required to achieve a minimum acceptable level of epidemiological impact, and for trade offs between cost of vaccine delivery and impact.

The complexities of dynamic populations, epidemiology, and the interactions between vaccine characteristics, make predicting the impact of other characteristics more challenging, especially over the extended time frames required to assess the impact of TB vaccines. For example, the relative impact of vaccines may evolve over time, or be influenced by prevalence of infection in the vaccinated population, which is dependent upon age at vaccination, contact patterns between age groups, and epidemiological trends over time. Therefore, mathematical modelling provides an appropriate framework for exploring the impact of these characteristics using an explicit and logical approach.

No systematic review of the modelling literature exploring the epidemiological impact of new TB vaccines has previously been conducted. Therefore, to provide such a summary and to identify key research gaps, I conducted a systematic review based upon the following research question: What is the estimated epidemiological impact of pipeline or theoretical human tuberculosis vaccines, as estimated by mathematical modelling? The modelling methods employed in studies identified were a secondary outcome of interest, to inform design of the subsequent research study. The PICOS framework for this research question is summarized in Table 2.3 below.

Table 2.3: PICOS framework for systematic review research question

Limit	Definition
Population	Humans, any age, any country
Intervention	Novel/theoretical/pipeline TB vaccines *Not* single efficacy BCG-only
Comparator	No intervention, currently available interventions (at current or scaled-up levels), or other theoretical interventions.
Outcome	Tuberculosis epidemiological impact (incidence, prevalence, mortality, number needed to vaccinate, cost effectiveness) *Not Mycobacterium bovis* *Not* immune outcomes
Time	No limit
Study Design	Epidemiological mathematical models *Not* within-host impact models *Not* reviews/commentaries

The systematic review was conducted in accordance with the York Centre for Reviews and Dissemination guidance for undertaking reviews in healthcare.[55] Detailed search strategies and sifting methodology are included in the published article (section 2.8) and appendix A. An adapted risk of bias tool was developed for this study (see Appendix A).

The published systematic review follows (in section 2.8). The review protocol, registered online on the PROSPERO database (reference: CRD42016033266), can be found in Appendix A.[56]

2.8 Systematic review of the TB vaccine modelling literature

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	Rebecca Claire Harris
Principal Supervisor	Richard White
Thesis Title	Informing development strategies for new TB vaccines: mathematical modelling and novel epidemiological tools

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?	Human Vaccines and immunotherapeutics		
When was the work published?	22nd July 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)	See next page

Student Signature: [redacted] **Date:** 29/9/17

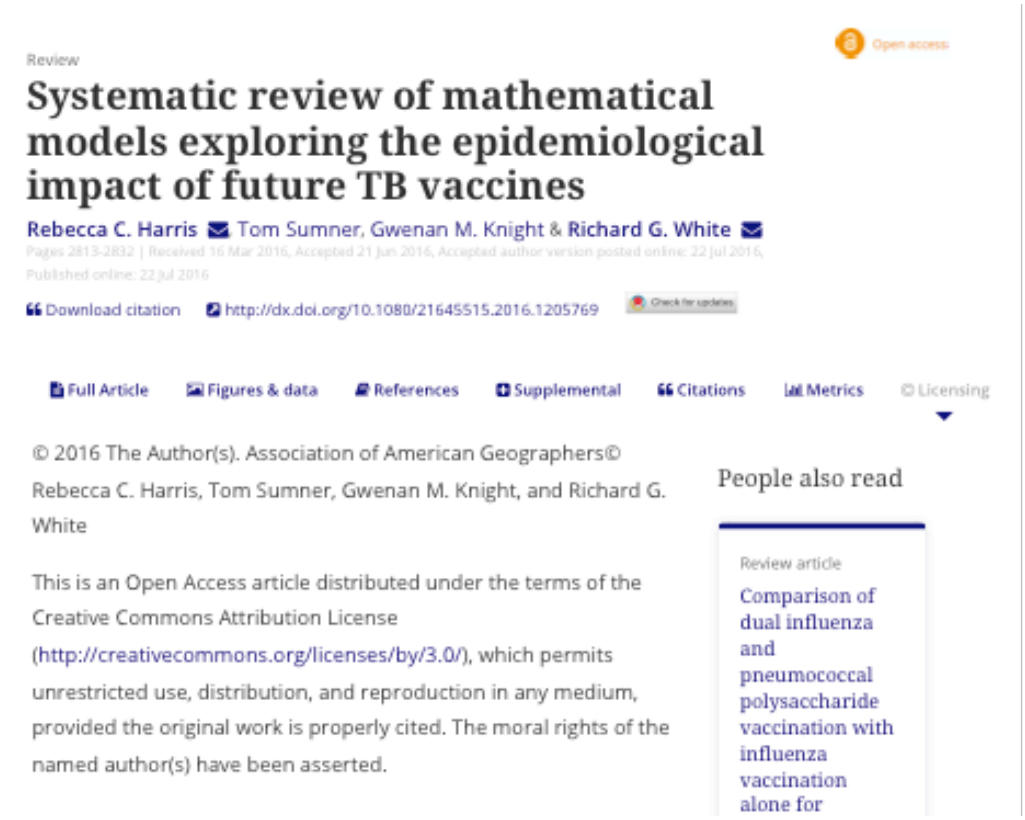
Supervisor Signature: [redacted] **Date:** 29/9/17

Paper 1: Informing development strategies for new TB vaccines: mathematical modelling and novel epidemiological tools

Authors: Rebecca C. Harris, Tom Sumner, Gwenan M. Knight, Richard G. White

Author contribution: For this systematic review, I conceptualised the study design, executed search strategies, sifting, data extraction, analysis, assessment of quality, interpretation of results and wrote the first draft of the manuscript. My supervisors, Prof. Richard White and Dr. Tom Sumner provided feedback on the design of the study. Dr Tom Sumner was the second reviewer for papers where there were uncertainties in inclusion or risk of bias assessment. Dr Gwenan Knight contributed writing to the section titled "Economic models". All authors contributed to design of the quality assessment tool, and reviewed the manuscript.

Permission from copyright holder to include this work:



Review

Systematic review of mathematical models exploring the epidemiological impact of future TB vaccines

Rebecca C. Harris, Tom Sumner, Gwenan M. Knight & Richard G. White

Pages 2813-2832 | Received 16 Mar 2016, Accepted 21 Jun 2016, Accepted author version posted online: 22 Jul 2016, Published online: 22 Jul 2016

Download citation http://dx.doi.org/10.1080/21645515.2016.1205769

Full Article Figures & data References Supplemental Citations Metrics Licensing

People also read

Review article

Comparison of dual influenza and pneumococcal polysaccharide vaccination with influenza vaccination alone for

CONTACT Rebecca C. Harris rebecca.harris@lshtm.ac.uk London School of Hygiene and Tropical Medicine, Keppel Street, London, UK; Richard G. White richard.white@lshtm.ac.uk London School of Hygiene and Tropical Medicine, Keppel Street, London, UK.

Supplemental data for this article can be accessed on the publisher's website.

HUMAN VACCINES & IMMUNOTHERAPEUTICS
2016, VOL. 12, NO. 11, 2813–2832
http://dx.doi.org/10.1080/21645515.2016.1205769

Taylor & Francis
Taylor & Francis Group

REVIEW

OPEN ACCESS

Systematic review of mathematical models exploring the epidemiological impact of future TB vaccines

Rebecca C. Harris[a], Tom Sumner[a], Gwenan M. Knight[a,b], and Richard G. White[a]

[a]TB Modelling Group, TB Centre and Centre for the Mathematical Modelling of Infectious Diseases, Faculty of Epidemiology and Population Health, London School of Hygiene & Tropical Medicine, London, UK; [b]National Institute for Health Research Health Protection Research Unit in Healthcare Associated Infection and Antimicrobial Resistance, Imperial College London, London, UK

ABSTRACT

Mathematical models are useful for assessing the potential epidemiological impact of future tuberculosis (TB) vaccines. We conducted a systematic review of mathematical models estimating the epidemiological impact of future human TB vaccines. PubMed, Embase and WHO Global Health Library were searched, 3-stage manual sifted, and citation- and reference-tracked, identifying 23 papers. An adapted quality assessment tool was developed, with a resulting median study quality score of 20/28. The literature remains divided as to whether vaccines effective pre- or post-infection would provide greatest epidemiological impact. However, all-age or adolescent/adult targeted prevention of disease vaccines achieve greater and more rapid impact than neonatal vaccines. Mass campaigns alongside routine neonatal vaccination can have profound additional impact. Economic evaluations found TB vaccines overwhelmingly cost-effective, particularly when targeted to adolescents/adults. The variability of impact by setting, age group and vaccine characteristics must be accounted for in the development and delivery of future TB vaccines.

ARTICLE HISTORY
Received 16 March 2016
Revised 13 June 2016
Accepted 21 June 2016

KEYWORDS
epidemiology; infectious disease dynamics; mathematical model; systematic review; theoretical models; tuberculosis; vaccines

Introduction

Although Bacillus Calmette–Guérin (BCG) is a longstanding cornerstone of the Expanded Programme on Immunization (EPI) schedule, the only licensed vaccine against tuberculosis (TB) disease, and one of the most widely used vaccines worldwide, it provides variable protection against pulmonary forms of tuberculosis disease and an uncertain duration of protection.[1,2] Tuberculosis is responsible for the largest number of annual deaths from a single infectious agent, estimated at 1.5 million in 2014, of which 91% were adults.[3] Of 9.6 million incident cases, 37% were undiagnosed or unreported.[3] Hence there remains a substantial unmet need for preventative measures such as new TB vaccines, particularly for protection against pulmonary disease in adult populations, which is clinically challenging to manage as well as being the source of most on-going transmission. It is a widely held view that new TB vaccines will be essential to the efforts to meet the World Health Organization (WHO) End TB Strategy 2035 goals and WHO 2050 goal of elimination of TB as a public health problem.[3-5] With 15 candidates currently in clinical trials, including one in each of phase IIb and III, the current TB vaccine pipeline is the most promising to date.[3]

The field is yet to see clinical efficacy in a novel candidate and the lack of an immunological correlate of protection for TB makes identifying promising candidates challenging, therefore clinical trials are long and costly, with limited guarantee of success. Mathematical models are invaluable tools for exploring the potential epidemiological impact of different future vaccine profiles and implementation strategies. They can inform the development of target product profiles and clinical development plans leading to vaccine candidates ready for licensure in the target populations in which they would have the greatest population-level impact. Modeling results are also important in advocating for vaccine development and investment.

The TB vaccine pipeline consists of a variety of vaccine profiles. Novel TB vaccine profiles can be categorized into four dimensions: the host infection status required for efficacy (pre-, post- or, pre- and post-infection: PRI, PSI and P&PI), an effect type (prevention of infection, disease, or infection and disease: POI, POD or POI&D), an efficacy, and a duration of protection (see Table 1 for definitions and abbreviations). Maximization of vaccine efficacy and duration of protection is obviously desirable, but given the challenges faced in TB vaccine development, a partially protective vaccine with limited duration of protection is a likely outcome. Similarly, a P&PI and a POI&D vaccine would have greatest impact, but given this may not be possible, less is known about the relative advantages of efficacy with different host infection status and effect types, particularly when considering extrinsic factors such as the age-structure

CONTACT Rebecca C. Harris rebecca.harris@lshtm.ac.uk London School of Hygiene and Tropical Medicine, Keppel Street, London, UK; Richard G. White richard.white@lshtm.ac.uk London School of Hygiene and Tropical Medicine, Keppel Street, London, UK.
Supplemental data for this article can be accessed on the publisher's website.

Table 1. Vaccine profile definitions.

Vaccine characteristic	Terminology	Definition	Abbreviation
Host infection status required for efficacy [a]	Pre-infection	Protects when delivered to uninfected populations. Does not protect when delivered to infected or previously infected populations.	PRI
	Post-infection	Protects when delivered to latent (and/or recovered) populations. Does not protect when delivered to uninfected populations.	PSI
	Pre- and post-infection	Protects when delivered to uninfected, latently infected or recovered populations	P&PI
Effect type (infection/disease transition protected against)	Prevention of infection	Effective against the acquisition of *M.tb* infection (uninfected to infected state)	POI
	Prevention of disease	Prevents progression to active disease (uninfected or infected to disease state)	POD
	Prevention of infection and disease	Prevents both infection and development of disease	POI&D
Efficacy	Vaccine efficacy	Protection provided by the vaccine. Can be "take" or "degree"	VE
Duration of protection	Duration of protection	Time during which vaccine remains efficacious. May include waning of protection	—

[a] *M.tb* exposure without infection, and immune priming with BCG or another vaccine, are not included within this definition, as exposure would not impact vaccine response, and priming could be under the control of the public health system.

and temporal trends of local epidemiology and other control measures. Therefore, modeling is a logical framework for estimating the influence of such factors on the population-level impact of future vaccines. This is important for rational development of target product profiles, minimum acceptable profiles, identifying target populations for vaccination, and estimating cost effectiveness of such vaccines.

As a growing field of research, publication of mathematical models assessing the potential impact of future TB vaccines has increased in recent years, yet no systematic review exists of this literature. Given the importance of this information for rational decision making in TB vaccine development and the strength of the current pipeline, a review of the literature was considered of importance. We therefore conducted a systematic review of published literature to answer the research question: what is the epidemiological impact of future human TB vaccines delivered to any age group when compared to no vaccination, other future vaccine profiles or other TB control interventions, as estimated using mathematical models? The aim was to provide a summary of the modeling methodology used, the characteristics of future TB vaccines explored using modeling, and the comparative epidemiological impact of such novel vaccine profiles.

Methods

PICOS framework

The PICOS framework was employed to define the review research question (Table S1). Searches were restricted to human studies. The interventions of interest were future, new, pipeline or theoretical vaccines for human tuberculosis. Papers exploring the impact of a single, defined-efficacy BCG vaccine were excluded; however, those exploring impact of a nominally BCG vaccine but with varying vaccine efficacy were considered for inclusion as these could be considered reflective of other theoretical vaccines. A broad definition was applied for the comparator of interest, therefore studies comparing new vaccines to no vaccine, BCG-only, alternative new vaccines or alternative currently available interventions were considered. Only articles reporting epidemiological outcomes, such as the impact on rates or absolute numbers of incidence, prevalence or mortality, or alternatively the number needed to vaccinate per case/death averted or cost-effectiveness measures, were included. Within-host impact models exploring immunological outcomes were excluded. The research question focused on the use of mathematical modeling as the study design. Narrative reviews and commentaries were excluded unless providing new modeling analyses or outcomes. No limits were placed upon publication dates.

Search strategy

Three electronic healthcare sources (PubMed, Embase and WHO Global Health Library) providing access to seven databases (PubMed/Medline, Embase, African Index Medicus, LILACS, SEARO Index Medicus, WPRO Index Medicus, EMRO Index Medicus) were searched. A comprehensive search strategy was developed using the defined PICOS framework, using free text and Mesh/Emtree/DeCS terms tailored by database for groups of terms covering tuberculosis, vaccines and modeling (Table S2). All searches were run with the "human" filter, and the WHO Global Health Library search was limited to regional databases to avoid duplication. No language or publication type limits were applied. Searches were conducted on the 12th January 2016.

Study selection and data extraction

This research was conducted in accordance with the York Centre for Reviews and Dissemination guidance for undertaking reviews in healthcare.[6] Three-stage manual sifting of titles, abstracts, then full texts was employed by the primary reviewer to apply the pre-defined inclusion/exclusion criteria described in Table 2. Where uncertainty with regards to inclusion at full text existed, a low threshold was used to trigger assessment by a second reviewer (TS). Reasons for exclusion at the full text stage were recorded. Data were extracted from eligible papers into a piloted, standardized Excel database by a first reviewer (RH), and fully checked by a second (TS/RW). Reference lists were hand searched and onward citation searching was conducted using Web of Science for all included studies.

Table 2. Inclusion and exclusion criteria applied in manual screening of articles.

Inclusion Criteria
Mathematical model
Systematic review of models of novel/future/hypothetical TB vaccine, or commentary adding to the analyses/interpretation of models reported elsewhere
Intervention is novel/future/hypothetical vaccine against tuberculosis or of an unspecified novel TB intervention with characteristics in-line with a vaccine
Reported outcomes are of the epidemiological impact of vaccine(s) (e.g. incidence, prevalence, mortality, number needed to vaccinate, cost effectiveness)
Exclusion Criteria
Within-host/immunological vaccine impact models
Review or commentary not adding to existing body of knowledge
TB epidemiological models not reporting impact of vaccine
TB epidemiological models reporting only interventions other than vaccines
Model only reporting on impact of BCG with single known/fixed efficacy
Disease or infection caused by *Mycobacterium bovis* or other non-*Mycobacterium tuberculosis* strain.

Qualitative synthesis of extracted data was employed to produce a narrative summary of included literature.

In accordance with PRISMA methodology, a protocol summary is registered on PROSPERO (reference: CRD42016033266) and the full protocol and PRISMA checklist are available in supplementary materials (Supplementary appendix C and Table S4).[7]

Assessment of quality

We developed a new adapted tool for assessment of modeling study quality and risk of bias (Table S3). This built upon previously published frameworks for health related modeling and health economic modeling.[8,9] By adding four more criteria to the Fone tool we allowed a more in-depth consideration of definition of model setting and population, appropriateness of modeling methodology and structure, fitting methodology, and reporting of conflicts of interest, which are essential for assessing the reproducibility of the model, alignment of the model and research question, and risks of bias.[8] To improve usability, the adapted tool presented contains the questions to be considered for each of the 14 criteria and clear guidance on the rating of zero, one or two for each criterion. If a criterion was not relevant for a particular paper, a score of one was assigned so as not to unduly bias the scores in either direction. Papers were scored 0-2 on each of 14 criteria, giving a maximum score of 28 points. A quality of "low" (<14), "medium"(14-18), "high" (19-22) or "very high" (>22) was assigned based upon the overall score.

Assessment of quality was conducted by one reviewer (RCH), but with a low threshold for examining any uncertain scores by a second reviewer (TS). The second reviewer also

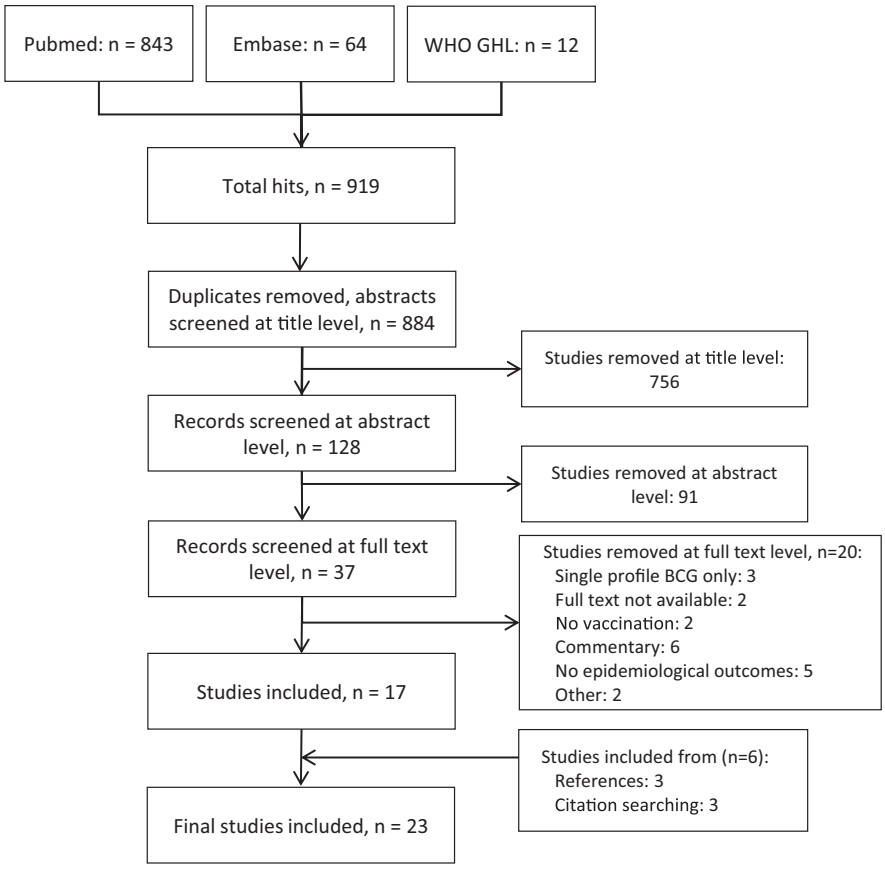


Figure 1. Summary of systematic screening of identified articles.

reviewed one paper from each quality level to check for agreement with the scoring and overall assessment of quality. Due to potential conflicts of interest, one paper was also reviewed by a third reviewer (SR).

Results

Screening process

A total of 919 records, comprising 884 unique articles after removal of duplicates, were identified (Fig. 1). Title screening removed 756 articles, and abstract screening a further 91, yielding 37 articles for review at full text. Twenty articles were excluded from the review during full text screening. The two most common reasons for exclusion were commentaries providing no additional analysis (n = 6) and articles without epidemiological outcomes (n = 5). Reference and citation searching identified 6 additional articles for inclusion, therefore 23 research articles were included in the review (Table 3). Although one BCG-based TB vaccine model exists from the 1960s,[10] two models published in 1998 were the first to explore entirely novel TB vaccine profiles,[11,12] and the majority (20/23) of included papers have been published since 2000,[4,13-31] motivated by the promising late-stage pipeline and the push to attain challenging WHO/Stop TB global TB targets.

Modeling methods (n = 23)

The included studies comprised 18 deterministic, compartmental, dynamic models constructed using difference or differential equations,[4,10-14,16-21,23-25,27-29] 3 Markov decision tree analyses,[15,22,31] one a simple mathematical model including a fixed number of transmissions per case,[30] and one statistical model.[26] A subset of these papers (n = 5) present analytical solutions of models and discuss the theoretical implications of a range of factors on the impact of vaccines.[11,13,17,25,26] These 5 papers do not provide any quantitative estimates of the epidemiological impact of novel vaccines and are not discussed further. A summary of model structures and fitting methods is provided in supplementary appendix A.

Vaccine characteristics (n = 18)

Modeled vaccine characteristics are summarized in Table 3. An anticipated vaccine efficacy in the range of 40–80% was most commonly modeled,[10,12,15,18,20-22,27-31] though some explored the public health impact of the extremes above[15,20,29] or below [10,12,22,29] this range. Vaccine efficacy was modeled as 'take' (i.e. a proportion of those vaccinated are completely protected, also known as an "all-or-nothing" vaccine) in 6 studies,[4,16,22,24,27,30] and as 'degree' (i.e., all vaccinees receive some protection from the vaccine, sometimes known as a 'leaky' vaccine) in 8 studies.[10,12,14,15,18,20,21,31] One study modeled efficacy as a combination of take and degree.[28] The type of vaccine efficacy was not possible to identify in three studies.[19,23,29] Those studies explicitly modeling vaccine coverage tended to model coverage of 80–100%,[10,12,15,20-22,27-31] though some included lower coverages.[12,20,27] Only one study assumed coverage using data reflective of country- and age-specific access to the healthcare system.[27] Some studies use a combined parameter ('proportion effectively immunized') of vaccine efficacy multiplied by coverage.[4,16,19,23,24] Where the proportion effectively immunized was not reported, we estimated the approximate proportion at five years after vaccine implementation. The scenarios of the proportion effectively immunized spanned a wide range in most studies, ranging from a lower limit of 9–38% to upper limit of 70–95%,[10,20,22,24,27,29,30,32] though some studies did employ lower upper limits,[4,12,15,16] and some explored just a single scenario of proportion effectively immunized within the range 44-70%.[19,21,23,28,31] In two studies it was not possible to estimate the proportion effectively immunized.[14,18]

Most often, duration of vaccine protection was assumed to be 10 years[15,20,22,24,27,30,31] or lifelong,[4,16,27,28] with some assuming alternative scenarios including 5, 30 and 33 years protection.[20,21,27,31] Seven articles did not report the duration of protection modeled,[10,12,16,18,19,23,29] but several appeared to use lifelong protection.[10,16,29] Waning of protection has been modeled as either exact (all depart the vaccinated state exactly at the end of duration),[4,14,16,24,27,28,31] or linear or exponential waning throughout the duration of protection.[15,21,31]

Many of the modeling studies explored the impact of multiple vaccine profiles assuming various effect types and/or host infection statuses required for efficacy. The post-infection (PSI) vaccines modeled assumed a prevention of disease (POD) vaccine effect,[4,12,14,19-21,28] or prevention of infection and disease (POI&D) effect.[4,16] Pre-infection (PRI) vaccines have been modeled assuming a prevention of infection (POI),[4,12,16,18,23,24,29] POD,[4,10,15,19,21,22,30,31] or POI&D vaccine effect.[14,20,28] Combined pre- and post-infection (P&PI) vaccines have been modeled assuming a POI&D[16] or POD[12,19,21,23,27] vaccine effect. The most frequently explored effect types were pre-infection vaccines with prevention of disease [4,10,15,19,21,22,30,31] or prevention of infection [4,12,16,18,23,24,29] effect, and post-infection vaccines with prevention of disease effect.[4,12,14,19-21,28]

Targeting of vaccination to populations with a specific host infection status (n = 12) was common in the models identified,[4,10,12,16,18-21,24,28,29,31] as was targeting to specific age groups (n = 12).[10,12,15,18,19,21,22,24,27,28,30,31] Pre-infection vaccines were most frequently targeted to neonates as they were a mostly uninfected population,[10,15,18,19,21-23,28,30,31] occasionally with an adolescent boost[21,31] or with a one-off mass campaign to all ages.[10,19,23,28] However, several studies included an analysis of the impact of pre-infection vaccines in adolescent,[24] high risk,[18] or all-ages mass vaccination campaigns or routine immunization.[12,20,21,28,29] Post-infection vaccines have mostly been modeled as delivered to all ages,[4,19-21,28] and in several studies age targeting is not stated but is thought to be delivered to all ages.[4,14,16] Pre- and post-infection vaccines have been modeled as delivered to neonates,[23,27] neonates with short term all-age mass campaigns,[12,23] routine adolescent vaccination with adult mass campaigns,[27] or delivered to all ages.[19,21]

Setting and population (n = 18)

Three were global studies,[12,14,16] two regional studies (e.g. WHO regions),[19,21] one in low- and middle-income countries,[27] four based upon hypothetical high burden settings,[4,10,20,28] three were set in South Africa,[15,22,24] one in Japan,[30] one in Zambia,[31] one in a hypothetical township,[29] one

Table 3. Summary of the 23 studies included in the review.

Author	Year	Modeling aim	Modeling Methods	Setting	Vaccine Profile: Host infection status	Effect type	Vaccine efficacy and coverage (%): Efficacy (take or degree)	Coverage	Proportion immunized [a] (approx. % at 5 yrs)	Duration of protection (waning)	Age targeting	Infection status targeting [b]	Schedule	Time horizon (years)	Outcomes [c]
Quantitative outcome studies (n=18)															
Abu-Raddad[21]	2009	Epidemiological benefits of TB	DE	SEAR (without China)	PRI	POD-f	60% (Degree)	100%	*60%*	33 yrs (0.03/yr)	Neo + Ado boost	All	Routine	35	IRR = 39% ICA = 18.2 m
					PRI	POD-fst	60%f 50%st (Degree)	100%	*60%f 50%st*	33 yrs (0.03/yr)	Neo + Ado boost	All	Routine	35	IRR = 52% ICA = 23.8 m
					PRI	POD-f	60% (Degree)	100%	*60%*	33 yrs (0.03/yr)	All	All	SM	35	IRR = 80% ICA = 68.2 m
					PSI	POD-s	50% (Degree)	100%	*50%*	33 yrs (0.03/yr)	All	L	SM	35	IRR = 37% ICA = 30.1 m
					P&PI	POD-fs	60% (Degree)	100%	*60%*	33 yrs (0.03/yr)	All	All	SM	35	IRR = 92% ICA = 80.2 m
				WPR	P&PI	POD-fs	60% (Degree)	100%	*60%*	33 yrs (0.03/yr)	All	All	SM	35	ICA = 51.5 m
				AFR	P&PI	POD-fs	60% (Degree)	100%	*60%*	33 yrs (0.03/yr)	All	All	SM	35	ICA = 47.1 m
				EMR	P&PI	POD-fs	60% (Degree)	100%	*60%*	33 yrs (0.03/yr)	All	All	SM	35	ICA = 15.4 m
				EUR	P&PI	POD-fs	60% (Degree)	100%	*60%*	33 yrs (0.03/yr)	All	All	SM	35	ICA = 10.1 m
				AMR	P&PI	POD-fs	60% (Degree)	100%	*60%*	33 yrs (0.03/yr)	All	All	SM	35	ICA = 9.1 m
Channing[22]	2014	Cost-effectiveness of adding MVA85A booster	Markov, Government perspective	South Africa	PRI	POD-d	17.3% and varied (Take)	99%	17.1% and varied	10 years	Neo BCG + 4m MVA	n/s, but likely U	Routine	10	Incremental CCA = USD 1,150 Incremental CDA = USD 284,017 Threshold VE = 41.3%
Cohen[14]	2008	Effect of strain diversity on vaccine performance	DE	Global (prevalence 220cases/100,000 population)	PRI	POI&D-isf	Calibrated so both vaccines have equal VE against strain 1(Degree)			n/s	n/s	n/s	n/s	n/s	Similar impact on quantity and distribution of strains at equilibrium. PSI demonstrated slower impact.
					PSI	POD-s				n/s	n/s	n/s	n/s	n/s	

(*Continued on next page*)

Table 3. (*Continued*)

					Vaccine Profile										
							Vaccine efficacy and coverage (%)								
Author	Year	Modeling aim	Modeling Methods	Setting	Host infection status	Effect type	Efficacy (take or degree)	Coverage	Proportion immunized [a] (approx. % at 5 yrs)	Duration of protection (waning)	Age targeting	Infection status targeting[b]	Schedule	Time horizon (years)	Outcomes [c]
Ditkowsky[15]	2014	Cost-effectiveness of BCG booster	Markov, Societal perspective	South Africa	PRI	POD-f	*70-85%*,[e] 0% VE in AIDS (Degree)	81% receiving prime and boost	*57-69%* [g]	10 years (linear over 10 years)	Neo BCG + 4m boost	n/s	Routine	10, one cohort	Infant booster with new TB vaccine less costly than BCG alone. ICA = 2,800-4,160 (40-70%) CA = USD 7.69m - 16.68m (40-70%)
Dye[23]	2000	Impact of future vaccines on TB control	DE	n/s	PRI	POI-i	n/s (n/s)	n/s	70%	n/s	All + Neo	All	1M + routine neo	25	IRR approx. 80%
					P&PI	POD-d	n/s (n/s)	n/s	70%	n/s	Neo	All	Routine	25	IRR approx. 25%
					P&PI	POD-d	n/s (n/s)	n/s	70%	n/s	All + Neo	All	1M + routine neo	25	IRR approx. 85%
Dye[16]	2008	Assessing impact and pipeline measures for elimination	DE	World, excluding sub-Saharan Africa and HIV (1030cases/ million pop)	PRI	POI-i [h]	n/s (Take)	n/s	*0-38% of uninfected*	n/s, assumed lifetime (none)	n/s	U	Continuous	43	IR (2050) approx. 10/million/yr. Greater impact than PSI when low rates of treatment of active disease.
					PSI	POI&D-is	n/s (Take)	n/s	*0-38% of latent*	Lifetime (none)	n/s	L	Continuous	43	IR (2050) approx. 100/million/yr. Greater impact than PRI when higher rates of treatment of active disease.
					P&PI	PO&ID-is[h]	n/s (Take)	n/s	*0-38% of people*	Lifetime (none)	n/s	All	Continuous	43	IR (2050) <0.2/ million/yr
Dye[24]	2013	Cost effectiveness of BCG revaccination of TST negatives in adolescence	DE	Cape town, South Africa	PRI	POI-i [h]	n/s (Take)	n/s	10-80%	10 years (exact)	Ado, HIV negative	U	Routine	Cohort lifetime	ICA = 7.5-17% (40-80% VE over cohort lifetime) Cost/DALY averted = USD 52-4,540 (80-10% VE)

Dye[4]	2013	Prospects for elimination using various control measures	DE	Typical high burden country (1100 cases/million pop/yr, CFR 16%)	PRI	POI-i	n/s (Take)	n/s	*76% at 5 years*	n/s	All	U	Continuous	35	IR (2050) = 130/million/yr
					PSI	POI&D-is	n/s (Take)	n/s	*34% and 53% at 5 years*	Lifetime	All	L	Continuous	35	IR (2050, 14%) = 20/million/yr
				South Africa (9800 cases/million/yr)	PRI	POD-d	n/s (Take)	n/s	*70% by 2050*	Lifetime	n/s	U	Continuous	25	Similar reductions to PSI profile below
					PSI	POD-d	n/s (Take)	n/s	*75% by 2035*	Lifetime	n/s	L	Continuous	25	IR (2050) = 1,400 cases/million/yr
				China	PRI	POD	n/s (Take)	n/s	n/s	n/s	n/s	U	Continuous	25	"Limited impact"
					PSI	POD	n/s (Take)	n/s	n/s	n/s	n/s	L	Continuous	25	PSI required for elimination IR (2050) < 1 case/million/yr
				India	PRI	POD	n/s (Take)	n/s	33% by2030, 100% by	n/s	n/s	U	Continuous	25	"Modest impact" by 2050
					PSI	POD	n/s (Take)	n/s	2050	n/s	n/s	L	Continuous	25	IR (2050) < 1 case/million/yr
Knight[27]	2014	Cost effectiveness of future TB vaccines	DE	Low and middle income	P&PI	POD-d	40%, 60%, 80%, reduced by 40% (10-70%) in HIV (Take)	22–99% (country specific)[g]	*9-79%*	5yr, 10yr, lifetime (exact)	Neo	All	Routine	26	ICA (10y/60%) = 0.89m Cost/DALY averted (10y/60%) = $1692 Only CE in infants if vaccine 80% VE with lifelong protection
					P&PI	POD-d	40%, 60%, 80%, reduced by 40% (10-70%) in HIV (Take)	16-99% ado, 68-91% adu (country specific)[i]	*6-79% ado, 27-73% adu*	5yr, 10yr, lifetime (exact)	Ado + Adu	All	Routine Ado + mass Adu (interval= duration)	26	ICA (10y/60%) = 17m Cost/DALY averted (10y/60%) = $149 CE threshold per dose = USD 4/20 in low/upper-middle income. All scenarios cost effective, some cost saving.

(*Continued on next page*)

Table 3. (*Continued*)

					Vaccine Profile										
							Vaccine efficacy and coverage (%)								
Author	Year	Modeling aim	Modeling Methods	Setting	Host infection status	Effect type	Efficacy (take or degree)	Coverage	Proportion immunized [a] (approx. % at 5 yrs)	Duration of protection (waning)	Age targeting	Infection status targeting[b]	Schedule	Time horizon (years)	Outcomes [c]
Lietman[28]	2000	Model used to assess impact of future TB vaccine	DE	Not clearly described, appears to start around incidence of 190cases/ 100,000pop/yr	PRI	POI&D-id	50% (Take & Degree)	88%	*44%*	Lifetime	Neo	U	Routine	40	IRR (40y) approx. 33%
					PRI	POI&D-id	50% (Take & Degree)	88%	*44%*	Lifetime	Neo + All	U	Routine +1M	40	IRR (40y) approx. 45%. Most effective of the strategies
					PRI	POI&D-id	50% (Take & Degree)	88%	*44%*	Lifetime	All	U	1M	40	IRR (40y) approx. 10%. Rebound observed as cohort age out.
					PSI	POD-s	50% (Take & Degree)	"88% … will eventually be vaccinated"	*44%, unclear timeframe*	Lifetime	All	L	SM	40	IRR (40y) approx. 45%. Fastest impact on incidence rate.
Murray[12]	1998	Impact of global control strategies	DE	Global (with 5 regional sub-models within)	PRI	POI-i	20%, 50%, 80% (Degree)	66%	*13-55%*	n/s	Appears all	U	SM	17	ICA (global)=10.5-37m ICA (Asia) = 6.6-23.2m ICA (SSA) = 3.4-12.2m
					P&PI	POD-f	20%, 50%, 80% (Degree)	Scale up to 80% over 10 years	*8-32%*[k]	n/s	Neo + All	All	Scale up mass (10 years), routine Neo after	17	ICA (global) = 16.2-51.6m ICA (Asia) = 10.1-32.3m ICA (SSA) = 5.4-17.1m
Pienaar[29]	2010	Effect of novel TB vaccine in hypothetical township	DE, household	Hypothetical township	PRI	POI	25%, 50%, 75%, 95% (n/s)	100%	*25-95%*	n/s, assumed lifelong	All	U	SM	10	Population infectious fraction reduced by approx. 7-70% after 120 months depending on VE

Rahman[30]	2001	Universal BCG (variable efficacy) for Japanese infants	Simple mathematical including transmission	Japan, hypothetical cohort	PRI	POD-d	40%, 60%, 80% (Take)	95%	*38-76%*	10 years	Neo	All	Routine	10, one cohort	CA = 111-542 (40-80% VE) CCA = USD 35,950-175,862 (80-40% VE) NNV =2,125-10,399 (80-40% VE)
Revelle[10]	1967	Optimization of TB control measures (variable efficacy of BCG)	DE, Optimization	Developing nation, high prevalence of active cases	PRI	POD-d	0%, 30%, 70% (Degree)	100%	*0-70%*	n/s, assume lifetime	Neo + All	U	Routine (10-13yrs) + 1M	20	Optimization of combination of vaccination and treatment
Rodrigues[18]	2009	Impact of vaccinating high risk groups versus uniform coverage	DE	'Resemble… developed country'	PRI	POI-i	75% (Degree)	Varied to achieve epidemiological targets	Varied	n/s	Neo	U	n/s	n/s	Targeted vaccination better than universal only when transmission is below the reinfection threshold.[l]
											High risk	U	n/s	n/s	
Tseng[31]	2011	Cost effectiveness of novel vaccines	Markov	Zambia	PRI	POD-f	70%, 0% if AIDS (Degree)	92%	*64%*, 0% if AIDS	10 years (linear)	Neo	U	Routine	30	ICA=932 CA=USD 3.6m Cost saving after 1 year
					PRI	POD-f	70%, 0% if AIDS (Degree)	92%	*64%*, 0% if AIDS	10 years (linear)	Neo+ Ado boost	U	Routine	30	ICA=1863 CA=USD 5.6m Cost saving after 6 years, greater savings than neonatal-only ≥16 years
Young[19]	2006	Estimate the impact of novel TB vaccines	DE	South Asia (200cases/100,000 population)	PRI	POD-d	n/s (n/s)	n/s	70%	n/s	Neo	U	Routine	35	IR (2050) approx. 70/100,000/yr
					PRI	POD-d	n/s (n/s)	n/s	70%	n/s	All + Neo	U	1M+ Routine	35	IR (2050) = 20/100,000/yr
					PSI	POD-d	n/s (n/s)	n/s	70%	n/s	All	L	Unclear if SM or 1M+ Routine	35	IR (2050) approx. 50/100,000/yr
					P&PI	POD-d	n/s (n/s)	n/s	70%	n/s	All	All	Unclear if SM or 1M+Routine	35	IR (2050) = 14/100,000/yr
Ziv[20]	2004	Public health impact of new TB vaccines	DE	'High burden settings' (100-200cases/100,000 population)	PRI	POI&D-ifs	50-90% (Degree)	60-90%	*30-81%*	10-30 years	All	U	1M + Routine U	40	ICA (10y) = 23%
					PSI	POD-s	50-90% (Degree)	60-90%	*30-81%*	10-30 years	All	L	1M + Routine L	40	ICA (10y) = 34% Impact of the 2 profiles becomes similar after 20-30 years.

					Vaccine Profile										
							Vaccine efficacy and coverage (%)								
Author	Year	Modeling aim	Modeling Methods	Setting	Host infection status	Effect type	Efficacy (take or degree)	Coverage	Proportion immunized [a] (approx. % at 5 yrs)	Duration of protection (waning)	Age targeting	Infection status targeting[b]	Schedule	Time horizon (years)	Outcomes [c]
Analytical studies (n=5)															
Bhunu[13]	2008	Effect of pre- and post-infection vaccines	DE (Analytical)	n/s	PRI	POD-f	100% (Take)	80%/yr	*100%*	(0.0002/yr)	n/s	U	n/s	200	Reduces infectious proportion approx. 75% compared to BCG-only, does not eliminate. Combined with treating latents and infectives, TTE = 15 yrs
					PSI	POD-s	75% (Degree)	80%/yr	*75%*	(0.0002/yr)	n/s	L	n/s	200	Combined with treating infectives, TTE = 20 years
Castillo-Chavez[11]	1998	Optimal age targeting of novel vaccines	DE (Analytical)	n/s	PRI	POI-i	0-100% (Degree)	Rate part of model optimization	n/s	n/s	1 vs. 2 age groups	U	n/s	n/s	Optimal age targeting schedule to minimize cost or effective reproduction number
Gomes [17]	2004	Assess impact of 'reinfection threshold' on VE	DE (Analytical)	n/s	P&PI	POI-i	75-80% (Degree)	95%-100%	*71-80%*	n/s (none)	Birth	n/s, but likely U	SM	400	VE≤natural immunity= impact only seen at low prevalence (i.e. below reinfection threshold). VE > natural infection immunity =impact seen over wider range of prevalence (i.e., reinfection threshold increased).
Gomes [25]	2007	Impact of reinfection on post-infection interventions	DE (Analytical)	n/s	PSI	POD-s	n/s (Degree)	n/s	Varies by L or recovered population	n/s	n/s	L & Recovered	SM	100	Bistable dynamics of magnitude of impact depending on whether prevalence above/below a threshold.

Hawn[26]	2014	How *M.tb* transmission affects whether VE is readily observed or masked	Statistical (Analytical)	n/s	PRI	POI-i	60% (Degree)	n/s	n/s	n/s	n/s	n/s	n/s	n/s	Fewer high-dose exposures more likely to attenuate vaccine efficacy than larger number of low dose exposures

1M: One-time mass campaign;

[a]Proportion immunized is the proportion of the population protected, defined as coverage times vaccine efficacy. Where not reported in the article, we have calculated this value (indicated by italics) by multiplying vaccine efficacy by coverage (or where vaccination rates are given, by estimating coverage at 5 years after vaccine introduction), though it should be noted that this does not account for waning.

Ado: Adolescent;

Adu: Adult;

[b]In population vaccinated, "all" refers to vaccination of all infection status populations except those with active disease.

[c]here large volumes of data were available, key outcomes of interest were reported for each vaccine type.

CA = Cost averted;

CCA = Cost per case averted;

CDA = Cost per death averted;

CE: cost-effective;

d: protection against progression to disease;

DA: Deaths averted;

DE: compartmental, deterministic, dynamic, difference or differential equations;

[e]Calculated, as BCG assumed 50% efficacious and boost VE 40-70% relative to BCG alone.

f: protection against fast progression to disease;

[g]Calculated proportion immunized does not account for waning, however waning at 5 years in this study will be significant, therefore calculated proportion immunized will be an overestimate.

[h]Vaccine leads to transition directly from uninfected to recovered with no possibility of developing disease.

i: protection against infection;

ICA: Incident cases averted;

IR: Incidence rate;

IRR: Incidence rate reduction;

[j]Infant coverage equivalent to DTP3 coverage, adolescent coverage equivalent to school attendance at 10yrs, mass coverage 20% below regional average of rubella campaigns.

[k]It was assumed that the scale up of coverage to 80% over 10 years was linear, therefore at 5 years assume 40% coverage.

[l]Duration of disease assumption very short (1 week).

L: latently infected;

Markov: Markov decision tree analysis;

Neo: Neonatal;

NNV = Number needed to vaccinate per case averted;
n/s: not stated;

POD: prevention of disease;

POI: prevention of infection;

POI&D combined prevention of infection and disease;

PRI: Pre-infection,

PSI: Post-infection,

P&PI: combined pre- and post-infection;

s: protection against progression to disease from slow latent state;

SM: sustained mass;

t: reduction in infectiousness when vaccinated;

TTE: Time to eradication;

U: Uninfected.

in a developed country,[18] and one where the setting was not clearly stated.[23] Two of the afore-mentioned studies also reported on country- (China, India and South Africa) or region-level models in addition to the main model reported.[4,12]

Only two of the 18 studies consider heterogeneous social mixing patterns: one included mixing by HIV status,[12] and one household model considers different mixing patterns during community, diurnal interactions and familial interactions at night.[29]

A small number of studies included risk groups within the modeled population. HIV status is largely neglected in TB vaccine modeling, a surprising fact given its importance as a driver of the TB epidemic in Africa and parts of Asia, but perhaps linked to this population being largely excluded from vaccine trials. One model targeted vaccination exclusively to the HIV-negative population.[24] Only four models explicitly include HIV stratification.[12,15,27,31] Two of the Markov decision tree models included the impact of HIV on TB natural history parameters, and assumed that the vaccine was equally efficacious in early-stage HIV infection and HIV uninfected individuals, but had zero efficacy in patients with AIDS.[15,31] Two globally-focused dynamic models included HIV strata, one of which did not alter vaccine efficacy in the HIV stratum,[12] and the other accounted for immunocompromise through reduced vaccine efficacy in HIV-infected patients.[27] Two of the models with HIV strata did not explicitly include antiretroviral therapy (ART) in the model.[12,31] The Ditkowsky study assumed 58% ART coverage which provided a 75% and 9.8% decrease in HIV-related annual mortality for those with early HIV and clinical AIDS, respectively, and a 61% reduction in the risk of progression to clinical AIDS (with associated higher risks of TB disease).[15] In the Knight model, receipt of ART doubled life expectancy, decreased rates of progression to TB disease or death, and halved the reduction in vaccine efficacy experienced due to HIV infection.[27] Aggressive ART scale up was assumed between 2012 and 2020, increasing from the 2009 coverage value in each country by half the difference between the 2009 value and 100%.[27]

Epidemiological impact of future TB vaccines (n = 18)

Fourteen studies had the epidemiological impact[4,12,16,18-21,23,28,29] and/or cost effectiveness [15,22,27,31] of future TB vaccine candidates as a primary focus. Three studies explored the impact of a variable efficacy BCG vaccination.[10,24,30] One strain competition model explored impact of differential vaccine effectiveness by strain.[14] Details of the vaccine profiles and outcomes discussed are summarized in Table 3.

i. Age targeting: Neonatal versus all ages, adolescents or adults

Only one study offered a clear comparison of targeting a given vaccine to neonates compared to adolescents/adults.[27] Two studies model the impact of vaccines delivered to neonates compared to delivery to all ages.[21,28]

In the study by Knight et al., implementation of routine adolescent vaccination with periodic mass adult campaigns was found to have a much greater impact than routine neonatal vaccination with a prevention of disease vaccine in low- and middle-income countries across the 2024-50 time horizon.[27] For example, a 60% efficacious vaccine providing 10 years protection could prevent 17 million (range 11–24m) cases between 2024-2050 when delivered to adolescent/adults compared to just 0.89 million (range 0.42–1.58m) when delivered to neonates.[27] The vaccine coverage differed in each age targeting scenario, based upon data from current vaccination campaigns relevant to those populations and so might be considered a more realistic reflection of achievable coverage. There are likely greater cost and logistical implications of delivering an adolescent/adult vaccine than a routine neonatal vaccine, however in this study it was found that due to the greater impact achieved when vaccinating older age groups, all of the vaccine profiles explored in this age group had a cost effective price per vaccine dose, whereas for neonatal vaccination some of the shorter durations of protection and lower vaccine efficacies were not considered cost effective. It was noted in the article that some conclusions may be timeframe dependent, as those receiving neonatal vaccination with long durations of protection would only just be reaching the age of high TB risk in 2050.

As would be expected, vaccinating all ages had a greater impact (80% incidence rate reduction over a 35-year time horizon) compared to vaccinating neonates with an adolescent boost (39% incidence rate reduction) with a POD vaccine in the Abu-Raddad study.[21] Though it should be noted that the number vaccinated would be much greater in the all-ages scenario. In the other neonatal vs. all-ages study,[28] routine neonatal vaccination was compared to a one-off mass campaign for uninfected individuals of all ages with a POI&D vaccine. In this study, mass vaccination of all ages initially provides greater epidemiological impact than neonatal vaccination. As the vaccinated cohort ages out the incidence rates rebound, whereas with sustained routine neonatal vaccination incidence rates continue to decline, and provide greater impact than a one-off mass campaign after approximately 20 years.[28]

No studies were found comparing vaccinating neonates to all ages or adults with prevention of infection vaccines, or of combining such age targeting with targeting of post-infection populations.

ii. Addition of mass campaigns or boosters to routine neonatal vaccination

Three studies found that the addition of one-off mass vaccination campaigns on top of neonatal routine vaccination could have profound effects on population-level impact.[19,23,28] In a South-East Asian study exploring impact of a pre-infection prevention of disease vaccine over a 35 year time horizon, an incidence rate reduction of approximately 65% was predicted with a neonatal vaccine, compared to >90% when adding a one-off mass campaign at launch.[19] In another study, addition of a one-off mass vaccination campaign to the routine neonatal program (with 70% of the population protected in each group) increased the reduction in incidence rate from around a quarter to approximately 85% over 25 years.[23] In the Lietman study, addition of a one-off mass campaign to routine neonatal vaccination greatly increased the short-term impact of the program.[28]

One study assessed the impact of an adolescent booster in addition to neonatal vaccination, which in a Markov model

was found to double the number of cases averted compared to neonatal vaccination alone.[31]

iii. Host infection status required for efficacy: Pre-infection, post-infection or pre- and post-infection

Eight studies report a comparison of vaccines effective in different host infection statuses,[4,12,14,16,19-21,28] of which six report clear comparisons between vaccines effective pre- versus post-infection.[4,16,19-21,28]

Exploring the impact on disease incidence, three studies suggest post-infection vaccines would have the greatest impact,[4,20,28] two suggest pre-infection would provide greatest impact,[19,21] and one suggests that either could have greater impact, dependent upon the rate of treatment of active disease.[16] The six studies allowing comparison are described below.

iii.a. Post-infection vaccines leading to greater impact on disease incidence (n = 3). The Ziv, Lietman and Dye (2013) models report post-infection vaccines as providing greater impact on disease incidence than pre-infection vaccines.[4,20,28] Ziv et al., compared the effect of POD vaccines targeted to uninfected (pre-infection) or infected (post-infection) populations in a hypothetical high burden setting with 28-50% latently infected and 40-60% of active cases treated and cured, and reported that post-infection vaccines would have a greater and more rapid effect on cumulative number of incident cases than pre-infection vaccines.[20] After 10 years, post-infection vaccines prevented 34% of cumulative cases compared to 23% by pre-infection vaccines.[20] Given 28-50% of the population were latently infected, the number vaccinated with the post-infection vaccine will be lower in most scenarios, therefore greater impact on overall disease incidence is achieved despite being given to fewer individuals. Age targeting, vaccine efficacy and coverage of target population were identical between scenarios and both were assumed to prevent development of disease, though in the pre-infection scenario the vaccine was additionally assumed to have POI activity, and therefore was a POI&D vaccine.

Similarly, in the Lietman study the pre-infection vaccine was POI&D whereas the post-infection vaccine was POD-only, but again, when comparing otherwise identical vaccines, post-infection vaccination provided the largest and fastest impact on the disease incidence rate.[28] In a background of 50% successful cure of active disease, comparing 88% coverage of latents with a post-infection vaccine, to vaccinating 88% of newborns with a pre-infection vaccine, impact was observed more rapidly with the post-infection vaccine, and 60 years later the disease rates with a post-infection vaccine were still lower.[28] The sizes of the latent and newborn populations are not reported, but they are unlikely to be equal, therefore the number of people effectively vaccinated and therefore the number needed to vaccinate per case averted in each of these scenarios could be very different. Even if routine neonatal pre-infection vaccination was preceded by a one-off mass campaign, the initial impact was still not as large as the post-infection vaccine, but in this scenario after 40 years the overall incidence rate reduction was similar.

In a study by Dye et al., in a "typical high burden country" with 110 cases per 100,000 population per year and 65% case detection, of which 70% are cured, effective immunization of 25%/year of the uninfected population (pre-infection) with a POI vaccine reduced disease incidence rates to 130/million population/year, whereas as little as 14%/year effective immunization of latents (post-infection) with a POD vaccine reduced the incidence much further to 20/million population/year.[4]

In all three studies reporting greater impact of post infection vaccines, the proportion assumed effectively immunized was lower in the post-infection vaccine target group.[4,20,28] Therefore, even with lower coverage, post-infection achieved greater impact than pre-infection vaccines.

iii.b. Pre-infection vaccines leading to greater impact on disease incidence (n = 2). In the Abu-Raddad and Young models,[19,21] pre-infection vaccines demonstrated greater impact on disease incidence than post-infection vaccines. In the Abu-Raddad model, campaigns targeted at latently infected individuals (post-infection) compared to mass vaccination campaigns using pre-infection or P&PI vaccines, predicted incidence rate reductions of 37%, 80% and 92% over a 35 year time horizon, respectively.[21] The Young study was also a South Asian model, which compared the impact of otherwise-identical PRI, PSI and P&PI vaccines implemented in 2015.[19] The model estimated an incidence rate of approximately 20, 50, and 14 per 100,000 population per year in 2050, indicating that the most effective vaccine in this scenario would also be a pre-infection or P&PI vaccine.[19] Though it should be noted that there was some ambiguity in the article as to the vaccine schedule, therefore there may be confounding due to differences in schedules modeled. Both the Abu-Raddad and Young models assumed that the pre-infection vaccines were prevention of disease vaccines, so would not have had a direct effect on infection.

iii.c. Research indicating a possible reason for such contrasting results (n = 1). A 'global' study excluding sub-Saharan Africa found that at low active disease treatment rates pre-infection vaccines providing lifetime protection had greater impact on disease incidence rates than equivalent targeting of post-infection vaccines.[16] However, when treatment rates of active disease were increased, reducing the proportion of disease from recent infection compared to reactivation, the post-infection vaccine was reported to have greater impact on disease incidence. We discuss this more below.

iii.d. Other factors influencing comparison of impact of pre- or post- infection vaccines. Although the proportions effectively immunized (coverage multiplied by vaccine efficacy) did not differ markedly between studies favoring pre- or post-infection vaccines, the proportion of the population latently infected will co-determine the number of people that would receive either a pre- or post-infection vaccine, and could be an important cause of differences. However, only two of these six studies reported the assumed infection prevalence.[16,20]

Time horizon could be important for this comparison. As demonstrated in the Ziv and Lietman models,[20,28] even though post-infection vaccines may have greater initial impact, the

Table 4. Quality assessment of included modeling studies.

Author	Year	Aims and objectives	Setting and population	Intervention/ comparators	Outcome measures	Model structure and time horizon	Modeling methods	Parameters, ranges and data sources	Assumptions explicit and justified	Quality of data and uncertainty and/or sensitivity analyses	Method of fitting	Model validation	Presentation of results and uncertainty	Interpretation and discussion of results	Funding source and conflicts of interest	Final Score (/28)	Rating
Abu-Raddad	2009	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Partial	Yes	No	Partial	Yes	Partial	23	Very high
Bhunu	2008	Yes	No	Partial	Yes	Yes	Yes	Partial	Partial	No	Partial	No	Partial	Yes	Partial	16	Medium
Castillo-Chavez	1998	Yes	Partial	Partial	Yes	Yes	Yes	Partial	Partial	Partial	Partial	Yes	Yes	Partial	Partial	20	High *
Channing	2014	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Partial	No	Yes	Partial	Yes	24	Very high
Cohen	2008	Yes	Partial	Partial	Yes	Partial	Yes	Partial	Yes	Yes	Partial	No	Partial	Partial	Yes	19	High
Ditkowsky	2014	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Partial	No	Yes	Yes	Yes	25	Very high
Dye	2000	Partial	Partial	Partial	Partial	Partial	Partial	Partial	Partial	No	Partial	No	Partial	Partial	No	11	Low
Dye	2008	Yes	Yes	Partial	Yes	Partial	Yes	Partial	Partial	Yes	Yes	No	Partial	Yes	Partial	20	High
Dye	2013a	Yes	Yes	Yes	Yes	Partial	Partial	Yes	Yes	Yes	Yes	No	Yes	Yes	Partial	23	Very high
Dye	2013b	Yes	Yes	Partial	Yes	Partial	Partial	Partial	Partial	Partial	Yes	No	Partial	Partial	Yes	18	Medium
Gomes	2004	Yes	Partial	Partial	Partial	Partial	Yes	Partial	Yes	Partial	Partial	No	Partial	Yes	Partial	17	Medium*
Gomes	2007	Yes	Partial	Partial	Partial	Partial	Yes	Yes	Yes	Partial	Partial	No	Partial	Partial	Partial	17	Medium*
Hawn	2014	Yes	Partial	Partial	Yes	Partial	Partial	Partial	Partial	Partial	Yes	No	Yes	Partial	Yes	18	Medium*
Knight	2014	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Partial	Yes	Yes	No	Yes	Yes	Yes	25	Very high
Lietman	2000	Yes	No	Yes	Yes	Partial	Yes	No	Partial	No	No	No	Partial	Yes	Partial	14	Medium
Murray	1998	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Partial	Yes	No	Yes	Yes	No	23	Very High
Pienaar	2010	Yes	Yes	Partial	Yes	Partial	Yes	Yes	Yes	No	Partial	No	Partial	Partial	Partial	18	Medium
Rahman	2001	Yes	Yes	Yes	Yes	Yes	Partial	Yes	Partial	Yes	Partial	No	Yes	Yes	No	21	High
Revelle	1967	Yes	Partial	Partial	Yes	Yes	Yes	Yes	Yes	Partial	Partial	No	Yes	Partial	Partial	20	High
Rodrigues	2009	Yes	Partial	Partial	Yes	Partial	Yes	Partial	Yes	Partial	Partial	No	Partial	Yes	Partial	18	Medium*
Tseng	2011	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Partial	Yes	Partial	No	Yes	Yes	Yes	24	Very high
Young	2006	Yes	Partial	Partial	Yes	Partial	Partial	Partial	Partial	No	No	No	Partial	Partial	No	12	Low
Ziv	2004	Yes	Partial	Yes	Yes	Yes	Partial	Partial	Yes	Yes	No	No	Yes	Partial	Yes	20	High
Median score		2	1	1	2	1	2	1	2	1	1	0	1	2	1	20	High

*Analytical modeling papers.

impact of the vaccine types sometimes converge over longer time horizons due to the greater time lag before impact is observed with pre-infection vaccines. However, all six studies explore outcomes over a narrow horizon of 35-43 years and therefore between-study differences in the impact of pre- vs. post-infection vaccines do not appear attributable to time horizon.[4,16,19-21,28] Durations of protection, and effect type for post-infection or pre-infection vaccines all vary between models, but were also somewhat similar in both groups.

iv. Disease stage protected against: POI, POD or POI&D

Dye et al. (2000) compared POI and POD vaccines with 70% effective immunization in a mass campaign followed by neonatal vaccination. In this study, the incidence rate reduction was greater with the POD vaccine than the POI vaccine over a 25 year time horizon.[23] In the Murray et al. global model, sustained mass vaccination with 66% coverage of uninfected individuals for a POI vaccine prevented fewer cases (10.5m-37.0m cases averted with 20-80% vaccine efficacy over a 17 year time horizon) than a POD vaccine with mass campaigns scaling up to 80% coverage over 10 years followed by annual neonatal vaccination (16.2m-51.6m cases averted); a trend also reflected in the regional estimates presented.[12] However, it must be noted that there are difficulties in identifying whether these are equivalent comparisons given the differences in schedule, size and infection status of population receiving vaccine and coverage.

Several models explored POI or POD vaccines compared to a POI&D vaccine. As would be expected since it is a combination of mechanisms, a number of the studies found the POI&D vaccine to be most effective.[4,14] Those finding more impact on disease with the POD vaccine than the POI&D vaccine were likely confounded by favorable targeting of the vaccines to latently infected individuals in the POD scenario.[20,28]

v. Time horizon of impact

The WHO/Stop TB global targets aim to reduce TB incidence rates to 10 cases per 100,000 population per year by 2035, and less than 1 case per million population per year by 2050, termed as tuberculosis elimination.[3,5] Several of the studies explored the potential contribution of future TB vaccines to achieving the Stop TB 2050 target, but all were published prior to release of the WHO End TB strategy containing the 2035 goals, so none directly address these targets. In two papers by Dye et al., it was shown that pre-infection POI vaccines to interrupt transmission were unlikely to reach elimination, as even with complete transmission interruption in 2015 an incidence of >100/million/year would still be expected by 2050.[4,16] In the Dye model, a post-infection vaccine giving full and permanent protection to 14% of the latently infected population per year was estimated to reduce incidence to 20/million/year by 2050, and if combined with treatment of active disease, or a pre-infection vaccine, would achieve elimination by 2050.[4] In country-specific models, results suggested that a post-infection vaccine (or mass preventative therapy) could achieve elimination targets in China and India.[4] These models suggest that a post-infection vaccine, or either vaccine effect type combined with treatment of active cases, may be capable of reaching the WHO 2050 elimination goals.[4] However, similar models for South Africa and the USA suggested that elimination was not considered feasible in these settings with novel vaccines.[4] Results from both the Young and Knight models indicated that none of the scenarios explored would achieve incidence rates lower than around 10-20/100,000/year by 2050.[19,27]

vi. Settings

There was clear variability in vaccine epidemiological impact by setting. Knight et al. demonstrated a larger proportion and absolute number of cases would be averted by vaccines in low-income countries than middle-income countries.[27] In Japan, 95% coverage of a BCG-like vaccine with 40-80% vaccine efficacy was considered not cost effective as it could only avert 10-47 cases per 100,000 children vaccinated over a 10-year horizon.[30] Whereas with 92% coverage of a 70% POD vaccine in the Zambian setting, cases averted were much higher at 199 cases per 100,000 vaccinated at birth;[31] though it should be noted that the Zambian model analyzed a longer 30-year time horizon. Three models explored the cost effectiveness of adding booster vaccines in South Africa, but outcomes are not comparable with the Japanese and Zambian models.[15,22,24] In two regional models, for each of the vaccine profiles explored the greatest numbers of cases and deaths avoided were in Asia/Western Pacific, followed by sub-Saharan Africa (see Table 3).[12,21] Much lower absolute numbers of cases were avoidable in the other regions, with the next largest impact observed in the Eastern Mediterranean region where cases averted were around a third of those in sub-Saharan Africa.[21]

vii. Economic models

A small number of published studies evaluated the economics of future TB vaccines (n = 7/18),[10,15,22,24,27,30,31] with the majority (n = 6/7) evaluating cost effectiveness of new vaccines (either future or BCG based) using either static,[15,22,30,31] or dynamic models.[24,27] Optimization techniques were also used to consider costs associated with different control schemes.[10] With the exception of one model exploring a very low efficacy vaccine and one exploring infant vaccination in a low burden setting,[22,30] those evaluating cost effectiveness found new TB vaccines to be an overwhelmingly cost effective intervention, whether from the health system[24,27] or societal perspective.[15,27,31] More than half of the models included societal costs, highlighting that much of the economic burden of TB disease falls on the TB patient due to the long durations of treatment. The populations considered in these economic models were either theoretical[10] or setting specific (South Africa,[15,22,24] Zambia[31] or Japan[30]) with only one model considering multiple settings (low- and middle-income countries).[27] Cost effectiveness of vaccines was highly dependent on vaccine characteristics such as efficacy (with lower efficacy linked to higher costs),[7,15,22,24,30] setting-specific burden of disease,[30] and economic considerations such as discount rate,[22,30,31] and time horizon.[31]

It is important to know the minimum acceptable vaccine efficacy for designing clinical trials and when making implementation decisions. Two South African Markov model threshold analyses of new vaccine boosters demonstrated that the booster strategy would be more effective, and either cost neutral or more expensive if the booster vaccine demonstrated a vaccine efficacy of approximately 40% or above.[15,22] The Knight

et al. study provided contour plots demonstrating the vaccine efficacies and durations of protection for which the vaccines would have an above-zero price at which the vaccine would be cost effective.[27] All vaccines in the ranges vaccine efficacy 20-100% and duration of 5 years to lifelong were cost effective, and even cost saving above 20% vaccine efficacy and 10 years duration, when delivered to adolescents and adults.[27] However, the neonatal vaccination program was only cost effective at higher values of these vaccine characteristics, with no cost effective vaccine price for a region of the plane where vaccine efficacy and/or duration were relatively low.

Studies identified post-hoc (n = 2)

For completeness, one unpublished study and one study published after the review search date were identified. The study published in March 2016 explores the impact of spatial 'hotspot' targeting compared to random allocation of an adult POD vaccine in Gujarat, India.[33] This model did not compare vaccine effect types or host infection status or age targeting, but is the first to explore spatial targeting of vaccines, and suggests spatial targeting could increase impact by 17% in the base case scenario in this setting.[33]

The unpublished study is an exploration of the epidemiological impact of different future TB vaccine profiles when targeted to adolescents (15 year olds) compared to older adults (60 year olds) in China.[34] This model is a dynamic transmission model incorporating heterogeneous social mixing patterns by age, age-specific natural history parameters, and fitted to age-stratified epidemiological data, and will contribute to the literature on vaccine age targeting and impact of pre- versus post-infection vaccines.

Quality assessment (n = 23)

Using the adapted quality assessment tool, scores ranged from 11 to 25 out of 28 (Table 4). Two were considered low quality,[19,23] eight of medium quality,[4,13,17,18,25,26,28,29] six high quality,[10,11,14,16,20,30] and seven of very high quality.[4,12,15,21,22,27,31] The median score was 20/28, equivalent to high quality. As the tool was not perfectly suited to analytical papers, the sensitivity of the median score was assessed by excluding analytical papers, and the overall score did not change. The major gaps observed were definitions of the population and intervention, fitting methodology, uncertainty and sensitivity analyses, data sources, and conflicts of interest and funding. Comprehensive uncertainty and sensitivity analyses were lacking in many studies. Further discussion is included in the supplementary appendix B.

Discussion

Modeling the epidemiological impact of future TB vaccines is a relatively young but growing field, used to inform rational decision making in portfolio strategy, prioritization of resources, and global target setting. To date, the literature remains divided as to whether vaccines effective pre-infection or post-infection would provide the greatest epidemiological impact. However, all-age or adolescent/adult targeted prevention of disease vaccines achieve greater and more rapid epidemiological impact than neonatal vaccines. Mass campaigns or boosters added to routine neonatal vaccination can have profound additional epidemiological impact. With the exception of one very low efficacy vaccine and one low burden setting, economic evaluations found new TB vaccines to be overwhelmingly cost effective, particularly when targeted to adolescent/adult age groups. The variability of impact by setting, age group vaccinated, vaccine characteristics and time frame, must be taken into account in the development and delivery of future vaccines.

Given the importance of indirect effects in this research question, the majority of the 23 included studies captured transmission through development of dynamic models. Some potential geographical bias was observed, as several either excluded sub-Saharan Africa or were based upon an Asia-like epidemic. Although the avoidable burden of disease may be higher in Asia, the high disease rates and slower progress toward global targets in the African region may mean that this region would be in greatest need of new interventions such as vaccines to meet the WHO 2035 and 2050 targets.[3] Epidemiological outcomes were indicative of considerable heterogeneity by setting, highlighting the importance of setting-specific modeling for decision making. Only four of the models incorporated HIV strata,[12,15,27,31] yet globally 12% of all TB cases are HIV co-infected, of which three-quarters are found in the African Region.[3] It is currently unknown what effect HIV-infection may have on vaccine efficacy, but it is generally thought that it may be lower in this population. Therefore, models without HIV structure may overestimate vaccine impact, and models including this structure are a useful tool to conduct sensitivity analyses around HIV-related assumptions.

Historically, novel TB vaccine development has focused primarily on vaccination in infancy. However, more recently a shift in thinking has led to the prioritization of older age groups in clinical trials and development plans. For prevention of disease vaccines, this is supported by the modeling literature, which demonstrates that all-age or adolescent/adult targeted POD vaccines achieve greater and more rapid epidemiological impact than neonatal vaccines. This is epidemiologically consistent given that in many settings adults comprise the majority of the disease burden and their primarily pulmonary disease is a greater source of *M.tb* transmission than the extra-pulmonary disease prevalent in children,[3] therefore delivery of POD vaccines directly to the adult population provides relatively immediate impact, whereas there is usually a lag of 10-20 years before any impact of long (>10 years) duration of protection neonatal vaccination can be seen.

From an implementation perspective, neonatal and adolescent vaccines could potentially be incorporated into existing delivery platforms, but developing a platform for delivery of adult vaccination could have serious resource implications. Although this is an important consideration, one model has demonstrated that adolescent/adult vaccine targeting would be more cost effective than neonatal vaccination with a POD vaccine.[27] Alternatively, several models suggest that addition of one-off mass campaigns for all ages at initiation of routine neonatal vaccination could have a profound effect on the

population-level impact.[19,23,28] No models were identified comparing age targeted POI vaccines or combining age targeting with targeting to post-infection populations. It was also noticed by the authors that none of the models explored targeted vaccination of older adults or the elderly, which is surprising given that several high burden countries are undergoing population aging (e.g., China, Indonesia) and the higher risk of developing active disease in this age group.

Our finding that the modeling literature was equivocal as to whether post-infection[4,20,28] or pre-infection[19,35] vaccines would have greatest epidemiological impact is interesting and may have important implications. Given the complexity of the studies, it is not possible to confidently identify the study differences that explain these diverse findings. However, there are 2 factors that we believe are likely to be most important. Firstly, the underlying epidemiology of the population and, in particular, factors affecting the proportion of disease due to primary disease and rapid progression versus reactivation or relapse, are likely to be influential. This was best illustrated in Dye and Williams,[16] in this study the authors increased rates of treatment of active disease, which resulted in a switch from greater impact from a pre-infection vaccine to greater impact from a post-infection vaccine (Dye and Williams, figure 6 panels a and b – by comparing new TB cases per million in 2050 in panel a versus b at increasing treatment rates per TB case). Increasing the treatment of active disease reduced transmission and the proportion of disease arising from primary infection (versus reactivation). We recreated this by coding up this model (not shown) showing that, for the same overall disease burden, when a greater proportion of disease was due to reactivation the post-infection vaccines were predicted to have the greater impact across the full range of rates of treatment of active disease explored. Secondly, the proportion of the population latently infected will co-determine the number of people that would benefit from either a pre- or post-infection vaccine. However, only two of these six models report the infection prevalence.[16,20] Another factor that could help explain this difference is that most of the six models compared assume no relapse from the recovered class, but in the two models assuming relapse,[4,19] it is possible that relapse rates and assumptions as to whether this group is protected by post-infection vaccination, could affect the relative impact of pre- versus post-infection vaccines. Further, although no major differences were identified here, time horizon, vaccine efficacy, and duration of protection assumptions could also potentially influence pre- vs. post-infection outcomes. Given the difficulties we have had in identifying the reasons why these models are making conflicting predictions, there is a clear need that these assumptions are reported carefully in future vaccine modeling studies, and if this question is important for vaccine development planning, there may be a need for a controlled modeling study focusing on this question.

From the vaccine development perspective, although a vaccine effective both pre- and post-infection would be the ideal scenario, pipeline candidates with either pre- or post-infection efficacy could have value, and modeling could be used to assess their relative value in different settings. Further, if a candidate could potentially exhibit different efficacy in infected versus uninfected populations, efficacy estimates from trials could be confounded by the balance between primary and reactivation disease in the trial setting. Or worse, if the recruited population is limited to either infected or uninfected individuals it is conceivable that a candidate's development could be discontinued due to poor efficacy in one population without knowing whether the candidate would have shown better efficacy in the other group. Therefore, if proven safe and immunogenic in both infected and uninfected populations, ideally both should be recruited into clinical trials and the study powered to estimate vaccine efficacy separately for those IGRA or TST positive versus negative at recruitment to improve generalizability to other settings. However, this could make trials infeasibly large, in which case enrolling both uninfected and infected populations, but powering the trial on the primary endpoint in one population, and looking for trend, safety and immunogenicity in the other population as secondary endpoints may be preferable to using a combined endpoint.

Upon implementation, sustained campaigns specifically targeting either uninfected (with a pre-infection vaccine) or latently infected populations (with a post-infection vaccine) may not be feasible. To identify such populations, TST or IGRA testing would be required. Such tests come with cost and organizational implications, and neither are perfect tests for latency or the absence of infection.[36] Blanket vaccination to ensure the target population is captured would be an alternative, but empirical data and modeling would be needed to assess the costs of vaccine wastage, and consideration given to the ethics of vaccinating individuals unable to derive direct benefit from vaccination.

Regarding prevention of disease versus prevention of infection vaccines, although several studies include vaccines with these different effect types, there tend to be other simultaneous changes in the vaccine profiles or targeting, such as age or infection status targeting, which confound the comparison of the impact of these two vaccine types. However, overall the studies suggest that prevention of disease vaccines tend to have a quicker and greater epidemiological impact than prevention of infection vaccines over the time horizons explored.

Theoretically, it is possible that there could be a genetic predisposition responsible for both ability to control *M.tb* latent infection and to respond to a vaccine. For a POD vaccine, this would become apparent in efficacy trials, as there would be no impact of such a vaccine on the disease endpoint. However, for POI vaccine studies with an infection endpoint, such a scenario could reduce infection rates, but have little impact on population level burden of disease as the vaccine may not be effective in those individuals most likely to progress to disease if infected. None of the POI vaccine models explored the potential scenario where vaccine efficacy is linked to likelihood to progress to disease, but this could be an interesting avenue for future research.

Although the relative impact of vaccine profiles to one another is informative for rational development of portfolio strategy, the absolute impact of such programs is important for understanding the potential role of such new technologies in achieving global targets and for advocacy for investment. Novel TB vaccines have the potential to provide an important contribution toward achieving the WHO 2035 and 2050 goals. Yet

given the ambitious nature of the 2050 targets, even novel vaccines may require synergistic pairing with other interventions to achieve elimination in most modeled scenarios. Due to the sizeable global pool of latent infection, even a complete transmission block may not achieve elimination because of the continued burden from reactivation disease. Therefore, prevention of reactivation disease through vaccination or preventative treatment of latently infected individuals will be essential to elimination strategy.

Cost must also be considered when planning implementation of new TB vaccines, therefore health economic models will be essential. With the exception of one low-efficacy vaccine study and one low burden setting,[22,30] the studies identified found new TB vaccines to be an overwhelmingly cost effective intervention. The results of threshold analyses are highly context dependent. However, in one analysis of low and middle income settings, vaccines targeted to adolescents and adults were shown to be cost effective as low as 20% vaccine efficacy and five years duration of protection, whereas infant vaccines required higher efficacies and longer durations to cross this threshold over the time horizons considered.[27]

An adapted tool was developed to assess quality and risk of bias of included studies for the purposes of this review to systematically assess reporting, methodological and risk of bias factors. The majority of included papers were scored as medium or high quality. The major gaps observed highlight for future studies the importance of thorough reporting and the conduct of comprehensive uncertainty and sensitivity analyses. It is hoped that this quality assessment tool will be of broad use in future systematic reviews assessing epidemiological models of other interventions and diseases.

The main limitation of this review was the conduct of independent sifting and data extraction by a single reviewer. The authors recognize that sifting by two independent reviewers remains the gold standard for systematic reviews.[6] However, due to resource constraints this was not possible, but a very low threshold was applied for directing sifting queries to a second reviewer. In addition, the review was first conducted in 2014 and then repeated in 2016, therefore duplication of sifting by the same primary reviewer was expected to reduce the likelihood of missing relevant literature. It was found that the study quality assessment tool developed was not as well suited to assessment of analytical models as several domains were not applicable, leading to a higher likelihood of scoring 'medium' quality.

Several research gaps were identified in this analysis of the available literature. The lack of a clear explanation for the polarization of outcomes for pre-infection and post-infection vaccines is troubling, therefore a model to explore which key determinants within the model impact these outcomes would be an important addition to the literature. None of the models presented explicitly explored the potential impact of targeting vaccines to older adult or elderly populations. Such a model would be pertinent for a country such as China, which has high disease burden, an aging population, and has only been briefly explored in one sub-model in the literature. Future vaccines could be important in tackling multi-drug resistance disease through prevention of transmission or disease, yet drug resistance was not explored in any of the models identified. Few of the models included non-random social mixing patterns, and none considered the potential impact of evolving mixing patterns on the impact of vaccines. Some studies have explored the epidemiological impact of vaccine age targeting in sub-Saharan Africa; however, these models were either missing HIV structure, did not explore reduction of vaccine efficacy in HIV-infected individuals, only considered vaccination of uninfected populations, or were static models. Given HIV co-infection and high forces of *M.tb* infection are fundamental to the epidemic in many sub-Saharan African countries, there is a need for a comprehensive model incorporating these important elements.

Conclusion

Mathematical modeling has been used to understand how the epidemiological impact of future vaccines could be altered by vaccine characteristics, vaccine age targeting, and epidemiological setting. It has also proved important for exploring the potential role of new vaccines for achieving the WHO 2050 goal of tuberculosis elimination. Such modeling should be integral to the development of future TB vaccines, informing rational decision making by cross-product bodies, academia, industry and policy makers for the development, investment and implementation of pipeline vaccines.

Abbreviations

ART	Antiretroviral therapy
BCG	Bacillus Calmette-Guérin
CI	Confidence intervals
EPI	Expanded Programme on Immunization
M.tb	*Mycobacterium tuberculosis*
PICOS	Population, Intervention, Comparator, Outcome, Study design
POD	Prevention of Disease
POI	Prevention of Infection
POI&D	Prevention of Infection and Disease
P&PI	Pre- and post-infection
PRI	Pre-infection
PSI	Post-infection
SSA	sub-Saharan Africa
TB	Tuberculosis
VE	Vaccine efficacy
WHO	World Health Organization

Disclosure of potential conflicts of interest

RCH provides consultancy for GSK Vaccines on work unrelated to the topic of this review. No other authors report potential competing interests.

Acknowledgments

We thank Sophie Rhodes (LSHTM) for independent review of risk of bias scoring of the Knight et al. paper.

Funding

RCH is funded by the UK Medical Research Council (MRC) under the LSHTM MRC vaccines scholarship program. RGW is funded by the UK Medical Research Council (MRC) and the UK Department for International Development (DFID) under the MRC/DFID Concordat agreement that is

also part of the EDCTP2 program supported by the European Union (MR/J005088/1) and the Bill and Melinda Gates Foundation (TB Modeling and Analysis Consortium: OPP1084276, and #OPP1110334), and UNITAID (4214-LSHTM-Sept15; PO #8477-0-600). TS is funded by the Bill and Melinda Gates Foundation (#OPP1110334). GK is funded by the National Institute for Health Research (NIHR) Health Protection Research Unit in Healthcare Associated infection and Antimicrobial resistance at Imperial College London in partnership with Public Health England. The views expressed are those of the author and not necessarily those of the NHS, the NIHR, the Department of Health, or Public Health England.

Author contributions

Conceived the study: RGW and RCH. Conceived and designed the protocol: RCH, RGW, TS. Execution of search strategy, sifting and data extraction: RCH. Development of quality assessment tool: RCH, RGW, TS, GK. Risk of bias assessment: RCH, TS. Manuscript preparation: RCH, RGW, TS, GK.

References

[1] Abubakar I, Pimpin L, Ariti C, Beynon R, Mangtani P, Sterne JA, Fine PE, Smith PG, Lipman M, Elliman D, et al. Systematic review and meta-analysis of the current evidence on the duration of protection by bacillus Calmette-Guerin vaccination against tuberculosis. Health Technol Assessment (Winchester, England) 2013; 17:1-372, v-vi.

[2] Mangtani P, Abubakar I, Ariti C, Beynon R, Pimpin L, Fine PEM, Rodrigues LC, Smith PG, Lipman M, Whiting PF, et al. Protection by BCG against tuberculosis: a systematic review of randomised controlled trials. Clin Infect Dis 2014; 58(4):470-80; PMID:24336911; http://dx.doi.org/10.1093/cid/cit790

[3] World Health Organization. Global Tuberculosis Report. 2015. Geneva, Switzerland, 2015. http://www.who.int/tb/publications/global_report/en/ Accessed: 30th January 2016.

[4] Dye C, Glaziou P, Floyd K, Raviglione M. Prospects for tuberculosis elimination. Annu Rev Public Health 2013; 34:271-86; PMID:23244049; http://dx.doi.org/10.1146/annurev-publhealth-031912-114431

[5] World Health Organization. The End TB Strategy. World Health Organization, 2015. http://www.who.int/tb/post2015_strategy/en/ Accessed: 30th January 2016.

[6] Centre for Reviews and Dissemination. CRD's guidance for undertaking reviews in health care. https://www.york.ac.uk/media/crd/Systematic_Reviews.pdf Accessed: 1st December 2015.

[7] Harris RC, White RG, Sumner T, Knight GM. Systematic review of models exploring the epidemiological impact of novel TB vaccines. PROSPERO 2016:CRD42016033266 http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42016033266 Accessed: 19th January 2016.

[8] Fone D, Hollinghurst S, Temple M, Round A, Lester N, Weightman A, Roberts K, Coyle E, Bevan G, Palmer S. Systematic review of the use and value of computer simulation modelling in population health and health care delivery. J Public Health 2003; 25:325-35; http://dx.doi.org/10.1093/pubmed/fdg075

[9] Jaime Caro J, Eddy DM, Kan H, Kaltz C, Patel B, Eldessouki R, Briggs AH. Questionnaire to Assess Relevance and Credibility of Modeling Studies for Informing Health Care Decision Making: An ISPOR-AMCP-NPC Good Practice Task Force Report. Value in Health 2014; 17:174-82; PMID:24636375; http://dx.doi.org/10.1016/j.jval.2014.01.003

[10] ReVelle CS, Lynn WR, Feldmann F. Mathematical models for the economic allocation of tuberculosis control activities in developing nations. Am Rev Respirat Dis 1967; 96:893-909; PMID:6059199

[11] Castillo-Chavez C, Feng Z. Global stability of an age-structure model for TB and its applications to optimal vaccination strategies. Math Biosci 1998; 151:135-54; PMID:9711046; http://dx.doi.org/10.1016/S0025-5564(98)10016-0

[12] Murray CJ, Salomon JA. Modeling the impact of global tuberculosis control strategies. Proc Natl Acad Sci U S A 1998; 95:13881-6.

[13] Bhunu CP, Garira W, Mukandavire Z, Magombedze G. Modelling the effects of pre-exposure and post-exposure vaccines in tuberculosis control. J Theoret Biolo 2008; 254:633-49; PMID:18644386; http://dx.doi.org/10.1016/j.jtbi.2008.06.023

[14] Cohen T, Colijn C, Murray M. Modeling the effects of strain diversity and mechanisms of strain competition on the potential performance of new tuberculosis vaccines. Proc Natl Acad Sci U S A 2008; 105:16302-7.

[15] Ditkowsky JB, Schwartzman K. Potential cost-effectiveness of a new infant tuberculosis vaccine in South Africa–implications for clinical trials: a decision analysis. PloS one 2014; 9:e83526; PMID:24454706; http://dx.doi.org/10.1371/journal.pone.0083526

[16] Dye C, Williams BG. Eliminating human tuberculosis in the twenty-first century. J R Soc Interface 2008; 5:653-62; PMID:17690054; http://dx.doi.org/10.1098/rsif.2007.1138

[17] Gomes MGM, Franco AO, Gomes MC, Medley GF. The reinfection threshold promotes variability in tuberculosis epidemiology and vaccine efficacy. Proc R Soc B: Biol Sci 2004; 271:617-23; http://dx.doi.org/10.1098/rspb.2003.2606

[18] Rodrigues P, Margheri A, Rebelo C, Gomes MG. Heterogeneity in susceptibility to infection can explain high reinfection rates. J Theoret Biolo 2009; 259:280-90; PMID:19306886; http://dx.doi.org/10.1016/j.jtbi.2009.03.013

[19] Young D, Dye C. The development and impact of tuberculosis vaccines. Cell 2006; 124:683-7; PMID:16497578; http://dx.doi.org/10.1016/j.cell.2006.02.013

[20] Ziv E, Daley CL, Blower S. Potential public health impact of new tuberculosis vaccines. Emer Infect Dis 2004; 10:1529-35; PMID:15498152; http://dx.doi.org/10.3201/eid1009.030921

[21] Abu-Raddad LJ, Sabatelli L, Achterberg JT, Sugimoto JD, Longini Jr IM, Dye C, Halloran ME. Epidemiological benefits of more-effective tuberculosis vaccines, drugs, and diagnostics. Proc Natl Acad Sci U S A 2009; 106:13980-5; PMID:19666590; http://dx.doi.org/10.1073/pnas.0901720106

[22] Channing L, Sinanovic E. Modelling the cost-effectiveness of a new infant vaccine to prevent tuberculosis disease in children in South Africa. Cost Effect Resource Allocation 2014; 12:1-9; PMID:24405884; http://dx.doi.org/10.1186/1478-7547-12-20

[23] Dye C. Tuberculosis 2000-2010: control, but not elimination [The Comstock Lecture]. Int J Tuberculosis Lung Dis 2000; 4:S146-S52.

[24] Dye C. Making wider use of the world's most widely used vaccine: Bacille Calmette-Guerin revaccination reconsidered. J R Soc Interface 2013; 10:20130365; PMID:23904584; http://dx.doi.org/10.1098/rsif.2013.0365

[25] Gabriela MGM, Rodrigues P, Hilker FM, Mantilla-Beniers NB, Muehlen M, Cristina Paulo A, Medley GF. Implications of partial immunity on the prospects for tuberculosis control by post-exposure interventions. J Theoret Biolo 2007; 248:608-17; PMID:17669435; http://dx.doi.org/10.1016/j.jtbi.2007.06.005

[26] Hawn TR, Day TA, Scriba TJ, Hatherill M, Hanekom WA, Evans TG, Churchyard GJ, Kublin JG, Bekker LG, Self SG. Tuberculosis vaccines and prevention of infection. Microbiol Mol Biol Rev 2014; 78:650-71; PMID:25428938; http://dx.doi.org/10.1128/MMBR.00021-14

[27] Knight GM, Griffiths UK, Sumner T, Laurence YV, Gheorghe A, Vassall A, Glaziou P, White RG. Impact and cost-effectiveness of new tuberculosis vaccines in low- and middle-income countries. Proc Natl Acad Sci U S A 2014; 111:15520-5; PMID:25288770; http://dx.doi.org/10.1073/pnas.1404386111

[28] Lietman T, Blower SM. Potential impact of tuberculosis vaccines as epidemic control agents. Clin Infect Dis 2000; 30 Suppl 3:S316-22; PMID:10875909; http://dx.doi.org/10.1086/313881

[29] Pienaar E, Fluitt AM, Whitney SE, Freifeld AG, Viljoen HJ. A model of tuberculosis transmission and intervention strategies in an urban residential area. Computat Biol Chem 2010; 34:86-96; PMID:20381428; http://dx.doi.org/10.1016/j.compbiolchem.2010.03.003

[30] Rahman M, Sekimoto M, Takamatsu I, Hira K, Shimbo T, Toyoshima K, Fukui T. Economic evaluation of universal BCG vaccination

of Japanese infants. Int J Epidemiol 2001; 30:380-5; PMID:11369746; http://dx.doi.org/10.1093/ije/30.2.380

[31] Tseng CL, Oxlade O, Menzies D, Aspler A, Schwartzman K. Cost-effectiveness of novel vaccines for tuberculosis control: a decision analysis study. BMC Public Health 2011; 11:55; PMID:21269503; http://dx.doi.org/10.1186/1471-2458-11-55

[32] Rahman M, Sekimoto M, Takamatsu I, Hira K, Shimbo T, Toyoshima K, Fukui T. Economic evaluation of universal BCG vaccination of Japanese infants. Int J Epidemiol 2001; 30:380-5; PMID:11369746; http://dx.doi.org/10.1093/ije/30.2.380

[33] Shrestha S, Chatterjee S, Rao KD, Dowdy DW. Potential impact of spatially targeted adult tuberculosis vaccine in Gujarat, India. J R Soc Interface 2016; 13(116). pii: 20151016; http://dx.doi.org/10.1098/rsif.2015.1016

[34] Harris R, Sumner T, Knight GM, White R. Future trends in TB epidemiology in China and the potential impact of novel age-targeted TB vaccines: A modelling study (SOA-620-06). 46th Union World Conference on Lung Health. Cape Town, South Africa, 2015.

[35] Abu-Raddad LJ, Sabatelli L, Achterberg JT, Sugimoto JD, Longini IM, Jr., Dye C, Halloran ME. Epidemiological benefits of more-effective tuberculosis vaccines, drugs, and diagnostics. Proc Natl Acad Sci U S A 2009; 106:13980-5.

[36] 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; 2013:11.

2.9 Systematic review update

The systematic review searches were updated on the 19th July 2017 to identify any additional published literature. Two new studies published in 2017 were identified, one modelling neonatal versus all-age vaccination strategies in China,[57] and another regional-level model exploring infant versus adolescent (15 year olds) vaccination.[58] The methods and results of these studies are summarised in Table 2.4, and explored in depth in the discussion in Chapters 3 and 4.

Table 2.4: Summary of the two additional new TB vaccine modelling studies identified in the systematic review 2017 update

Author	Year	Modelling aim	Modelling Methods	Setting	Vaccine Profile						Age targeting	Infection status targeting [b]	Schedule	Time horizon (years)	Outcomes [c]
					Host infection status	Effect type	Vaccine efficacy and coverage (%)			Duration of protection					
							Efficacy (take or degree)	Coverage	Proportion immunised [a] (approx. % at 5 yrs)						
Liu[57]	2017	Epidemiological impact of 'pulsed' mass vaccination strategy compared to neonatal vaccination	DE	China	n/s	POI-i	100% (take)	95%	95%	L/L	Neo	n/s	Routine	18 (intro 2018)	Cannot achieve 2035 EndTB goals
								Neo: 70-95% Mass: 10-40%	Neo: 70-95% Mass: 10-40%		Neo + All		Routine + mass every 3-6 years		Can achieve 2035 goals with 70% neonatal coverage and 25% 5-yearly mass coverage. Could be achieved sooner with more frequent campaigns or higher coverage. With 80% neonatal and 3-yearly 30% mass coverage, EndTB goals met in 2030
Arregui[58]	2017	Age dependent TB patterns and epidemiological impact of vaccines	DE	Global (5 regions)	P&PI (not recovered)	POI-i	80% (take)	100%	80%	L/L	Neo	U&L (not recovered)	Routine	26	Adolescent vaccine consistently averted more cases than neonatal over this time period. Largest number of cases averted 2025-50 was in African region (AFR), smallest number in Western Pacific region (WPR). Biggest difference between vaccination strategies was in WPR (81%), smallest difference was in AFR (53%). Estimate from graph approx. 80m cases averted with adolescent vaccination and approx 30m with neonatal vaccination
											15 year olds				

DE: compartmental, deterministic, dynamic, difference or differential equation; L/L: Lifelong

2.10 Modelling literature summary

In summary, the existing body of mathematical modelling literature does not provide a comprehensive and comparable exploration of the population-level epidemiological impact of the vaccine characteristics and implementation strategies required to inform vaccine TPPs and clinical trial design. In particular, questions remain with regards to the relative impact of pre- versus post-infection vaccines, and of vaccines with varying combinations of efficacy against infection and disease, in the context of different durations of protection. Also, no models explore the impact of vaccinating older adults or the elderly, which is a potentially epidemiologically important population over the time horizons of TB vaccine modelling.

With phase III trial results in China due within the next year, such models are needed for estimation of the potential future impact of such vaccines, and planning for future implementation. China is the third largest contributor to global TB incidence and a rapidly ageing population, yet no modelling literature exists exploring the potential impact of targeting older adults with new TB vaccines, and the China TB vaccine modelling literature explores an extremely limited subset of vaccine types with insufficient demographic and epidemiological age complexity.[31,57]

Therefore, in the next section of this thesis (chapters 3 and 4) I develop a dynamic transmission model fitted to age-structured demographic and epidemiological data from China, to explore the future dynamics of the TB epidemic in China and the potential impact of varying vaccine characteristics in this setting (Chapter 3, objectives 2a-c). To explore the impact of implementation strategies, this model is adapted to explore the impact of the current strategic focus on adolescent vaccination compared to targeting older adults (Chapter 4, objective 5).

Additional research questions remain with regards to the impact of vaccines in high HIV-prevalence settings and the impact of new TB vaccines on multi-drug resistance. These, however, were beyond the scope of this thesis, and are discussed in further work (Chapter 6).

2.11 Translating TB vaccine modelling into trial design

The exploration of the TB vaccine modelling literature identified adolescents and adults as an important age group for vaccine targeting and has shown that prevention of disease vaccines are likely to provide the greatest and most rapid impact. The results from the existing modelling literature, in conjunction with the modelling results that will be presented in chapters 3 and 4, provide a clear indication of the importance of developing vaccines for key target populations and are suggestive of prevention of disease outcomes as the most epidemiologically important endpoint. These target populations and desired outcomes must then be translated in to appropriate design of clinical trials. As highlighted by the two epidemiological research needs identified from the TB Vaccine Blueprint (see Box 1, Chapter 1), mathematical modelling can inform the ideal characteristics and populations, but then epidemiological data is needed to select research sites and identify disease hotspots in the planned enrolment population, to inform sample size calculations and recruitment plans.

Spatial variation in the burden of tuberculosis (TB) is well recognised at national and regional levels as a result of government surveillance systems. However, such heterogeneity is often poorly understood at the local level, such as between areas within a city. A small number of studies have demonstrated such local variations exist for TB disease in sub-Saharan Africa. For example, in Cape Town, South Africa, 39 enumerator districts (across two communities) with an average population of 900 inhabitants, reported adult TB notification rates ranging between 0 and 2,847/100,000 adults per year.[59] Similar variations were reported in a settlement-level study in the Gambia, where rates in settlements of 4,000-68,000 people ranged from 48 to 239/100,000 population/year.[60] The potential for such large local variations highlights the need for local-level burden of disease information in the planned vaccine trial enrolment population.

Given the importance of prevention of disease vaccine trials, epidemiological data regarding TB disease are required, including spatial granularity and stratified by key

population characteristics of interest for vaccine development (e.g. age and HIV status). Incidence rate data in the planned recruitment population are the gold standard for informing sample size calculations, as they are directly indicative of the likely rate of accrual of endpoints in the trial. Where incidence data are not available, notification rate or prevalence data can also provide an indication of likely rates to de-risk sample size calculations. In the next section, I explore the available options for collecting such data and the strengths and limitations of those tools.

2.12 Available studies and tools for collecting spatial data for recruitment populations

Several existing study designs, databases and tools are available for collecting spatial data to inform design of clinical trials to recruit the populations identified in the vaccine TPP. These include already available data sources such as National Tuberculosis Programme (NTP) or Health and Demographic Surveillance Site (HDSS) data, prospective data collection through initiation of a new epidemiological study, or the possibility of adapting available spatial map tools to assist data collection. The advantages and disadvantages of these approaches are detailed in Table 2.5, and summarized briefly below.

Existing studies exploring spatial distribution of TB cases mostly use either HDSS data or Global Positioning Systems (GPS) to record coordinates at the patient's place of residence. There are 54 HDSS sites globally,[61] but such sites are a limited resource and not all sites have sufficient incidence of TB disease or clinical trial capacity for a TB vaccine clinical trial. Measuring GPS coordinates at the home is a resource-intensive and logistically complex method of collecting spatial data.[59,62,63] Due to these resource and logistical limitations, such detailed prospective TB case mapping in non-enumerated populations is uncommon.

Patients' addresses are widely recorded at TB registration in the NTP, but these data are rarely used for spatial analysis and in areas without official address systems, the accuracy of such descriptions is often low. This was demonstrated by a study

comparing verbal descriptions in the clinic register to GPS readings taken at the home.[64] Two study staff were unable to categorise the patient's cluster (approx. 1,300 population) of residence from verbal descriptions for 31.8% and 59.7% of the 129 participants, and agreement between the study staff was only 61.5%.[64] This demonstrated the severe limitations of retrospective spatial mapping with national registry data.

The gold standard for informing clinical trial sample size calculations and identifying where to recruit the enrolment population would be a prospective incidence study. However, there is rarely enough time when initiating clinical trials to allow for the time needed to conduct an incidence study. Prevalence surveys may be faster, but are not as directly applicable to sample size calculations as incidence or notification rates, and still require substantial resources. Both incidence and prevalence prospective studies are expensive and time consuming, so may delay the start of a clinical trial. Often there is insufficient time and resource available to conduct such prospective studies.

Table 2.5: Summary of existing studies or tools that could be employed to collect spatial TB burden data

Tool/Study	Description	Strengths	Limitations
National Tuberculosis Programme data	Routinely collected data when new patient registers for treatment with the TB programme	• Readily available, continuously collected • Most high burden settings routinely collect these data • Aggregated data usually report paediatric and adult burden separately, and registers usually record age. • Usually aggregated data are stratified by type of TB (pulmonary/extrapulmonary)	• Lack of spatial granularity as clinics may cover broad areas, and in low income settings without address systems they may only record verbal descriptions of location of residence. • Misses those cases that do not seek care • Demonstrated in Blantyre, Malawi, to be difficult to identify area of residence from the verbal descriptions.[64]
Health and Demographic Surveillance Site (HDSS) data	Enumerated HDSS sites can link cases reported in clinics to GPS-captured place of residence	• Provides spatial data to the household level • Usually good coverage and continuously collected • Often collect large amounts of demographic and additional risk factor data as part of the HDSS survey rounds • Usually sites with capacity to conduct clinical trials	• Very limited number of sites, not always sufficiently high TB burden. • Misses those cases that do not seek care • Expensive to put enumeration system in place, so only low cost if piggy-backed to a larger research study.
Prospective study (incidence or prevalence)	Design and development of an incidence or prevalence study to prospectively collect the required data	• Can collect data in the recruitment population of interest and in locations able to conduct a clinical trial and with known regional high burden of disease. • A well-executed study should capture all cases (i.e. no issues of missing those not seeking care) • Outcomes can be designed to match the planned study endpoints	• Very expensive • Time consuming to set up and conduct study. Incidence studies in particular may need long durations of follow up. There may not be sufficient time or budget ahead of a clinical trial to allow for this. • Is a one off study, is not continuously updated so data become out of date very quickly.

GPS collection of place of residence	Using GPS handsets to collect coordinates at the home of patients registering for NTP treatment	• Provides precise location of place of residence	• Time and resource heavy • Complicated logistics of ensuring availability of field staff to return to home of all patients immediately after registration, or difficulties of following up at a later date. • If conducted by enrolling patients initiating treatment at the NTP, would also miss those not seeking care
Open Street Map	A free open-access online pictographic map system.	• Flexible tool where structures and POIs can be added in areas where there are insufficient data • Free and open access	• Requires internet connection to update maps • Pictographic representation of roads and structures can be difficult to use, especially for populations with low map literacy. • Currently no system in place for collecting sensitive patient data.
Google Maps	A free open-access online satellite image-based map system.	• High resolution maps in a variety of views/formats • Can zoom to see more/less detail • Can see topography and details in satellite images • Can add on points of interest • Free and open access	• Assumes map literacy • Assumes sufficient points of interest to orientate within map • Requires high speed and reliable mobile data, which is often not available or very expensive in setting of interest for TB vaccine efficacy trials, or computer with good connection • Currently no system in place for using Google maps to collect sensitive patient data
Mapbook[64]	A series of paper maps with points of	• Mapped points of interest identified by the community improve ease of use of maps in this	• Only maps patient to a cluster of approximately 1,300 people, rather than to the street or household level.

	interest and study clusters marked, used in clinics to identify whether patients are from cluster	community without address system and with low map literacy • Can see topography and details in satellite images • Demonstrated to be highly accurate when used for patients starting ART treatment in Blantyre, Malawi.	• Only identifies residence for patients in 28 study clusters • Paper based maps and data collection, not automated, slow to update, and not very easy to share data • Cannot zoom in to see more detail

A new approach is needed to be able to prospectively collect place of residence, ideally integrated in the NTP to keep costs and logistical issues to a minimum. Several tools are available that could be used to collect place of residence at time of registration for TB care. These include open street maps, Google maps and the Blantyre Map book. As described in Table 2.5, these tools can provide high resolution identification of place of residence, and in most cases are flexible to be able to add additional points of interest to help users orientate themselves in the map. However, populations in low income areas with high burden of disease may have low map literacy, therefore such tools may not allow ease of use. Mobile data connectivity is also an issue in many low income areas with high burden of disease. Both open street map and Google maps require internet connectivity and use large volumes of mobile data to load the maps, therefore in many low income settings these map systems are not sufficiently reliable, or too expensive to run. There may also be challenges in using this open access software to collect personally identifiable information for TB patients, as such data must be collected, transferred and stored in a highly secure manner to protect patient data.

The Map book system was designed to overcome some of these issues, by using easy to use paper maps annotated with locally known points of interest, and with paper-based collection of cluster residency.[64] This is feasible for a limited number of small areas, but in itself is not a scaleable approach.

The need for a new, low-cost, easy to use and fast to implement tool for collection of spatial epidemiological data to inform TB vaccine clinical trial design was identified, ideally bringing together the low cost, ease of use and community engagement of the Map book system, but the scaleability, accuracy, and real time data collection of electronic platforms.

In chapter 5 of this thesis, I describe the development and evaluation of ePAL (electronic PArticipant Locator) - a new app designed for rapid, low cost collection of the location of the TB patient's place of residence as an integrated part of the

National Tuberculosis Programme. In Chapter 5, I provide spatial mapping of tuberculosis notifications overall and in potential TB vaccine trial recruitment populations from the first 12 months of data collection from the implementation of ePAL in Blantyre, Malawi.

2.13 Epidemiological tools and data to inform the development of TB vaccines

In this chapter, I have described the public health need for new TB vaccines and challenges in their development. From these challenges, two clear epidemiological research needs have been identified. Firstly, the need for mathematical modelling of a comprehensive range of vaccine characteristics in China, in addition to a consideration of the impact of age targeting of vaccination to adolescents versus older adults, to inform maximisation of the potential future burden of pipeline vaccines. Secondly, the need for a new, low-cost, rapidly implementable tool for spatial mapping of TB cases, to allow data-informed sample size calculations, ideally from areas with hotspots in the modelling-informed target populations.

In the research presented in this thesis, I developed a dynamic transmission model to explore the vaccine characteristics and implementation questions identified in the literature review (Chapters 3 and 4), and a novel tool for spatial mapping of the residence of TB notifications (Chapter 5).

2.14 Chapter 2 References

1. Harris RC, Sumner T, Knight GM, White RG. Systematic review of mathematical models exploring the epidemiological impact of future TB vaccines. *Hum Vaccin Immunother* 2016; **12**(11): 2813-32.
2. U.S. Centres for Disease Control and Prevention. Core Curriculum on Tuberculosis: What the Clinician should know (Chapter 2). 2013 2013. https://www.cdc.gov/tb/education/corecurr/pdf/chapter2.pdf (accessed 9th June 2017).
3. Zumla A, Raviglione M, Hafner R, Fordham von Reyn C. Tuberculosis. *New England Journal of Medicine* 2013; **368**(8): 745-55.
4. 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): 275-80.
5. Narasimhan P, Wood J, MacIntyre CR, Mathai D. Risk Factors for Tuberculosis. *Pulmonary Medicine* 2013; **2013**: 11.
6. Bates MN, Khalakdina A, Pai M, Chang L, Lessa F, Smith KR. Risk of tuberculosis from exposure to tobacco smoke: a systematic review and meta-analysis. *Archives of internal medicine* 2007; **167**(4): 335-42.
7. Lin HH, Ezzati M, Murray M. Tobacco smoke, indoor air pollution and tuberculosis: a systematic review and meta-analysis. *PLoS medicine* 2007; **4**(1): e20.
8. Lonnroth K, Williams BG, Stadlin S, Jaramillo E, Dye C. Alcohol use as a risk factor for tuberculosis - a systematic review. *BMC Public Health* 2008; **8**: 289.
9. Escombe AR, Moore DAJ, Gilman RH, et al. Upper-Room Ultraviolet Light and Negative Air Ionization to Prevent Tuberculosis Transmission. *PLoS medicine* 2009; **6**(3): e1000043.
10. Vynnycky E, Fine PE. Lifetime risks, incubation period, and serial interval of tuberculosis. *American journal of epidemiology* 2000; **152**(3): 247-63.
11. Esmail H, Barry CE, Young DB, Wilkinson RJ. The ongoing challenge of latent tuberculosis. *Philosophical Transactions of the Royal Society B: Biological Sciences* 2014; **369**(1645): 20130437.
12. Vynnycky E, Fine PE. The natural history of tuberculosis: the implications of age-dependent risks of disease and the role of reinfection. *Epidemiology and infection* 1997; **119**(2): 183-201.
13. Ferebee SH. Controlled chemoprophylaxis trials in tuberculosis. A general review. *Bibliotheca tuberculosea* 1970; **26**: 28-106.
14. Sutherland I. The ten-year incidence of clinical tuberculosis following "conversion" in 2550 individuals aged 14 to 19 years. The Hague, The Netherlands: KNCV, 1968.
15. Flynn JL, Chan J. Tuberculosis: Latency and Reactivation. *Infection and Immunity* 2001; **69**(7): 4195-201.
16. World Health Organization. Global Tuberculosis Report 2016. 2016. http://www.who.int/tb/publications/global_report/en/ (accessed 21st December 2016).
17. Schaaf HS, Collins A, Bekker A, Davies PDO. Tuberculosis at extremes of age. *Respirology* 2010; **15**(5): 747-63.

18. Tiemersma EW, van der Werf MJ, Borgdorff MW, Williams BG, Nagelkerke NJD. Natural History of Tuberculosis: Duration and Fatality of Untreated Pulmonary Tuberculosis in HIV Negative Patients: A Systematic Review. *PloS one* 2011; **6**(4): e17601.
19. World Health Organization. Global Tuberculosis Report 2016. 2016. http://apps.who.int/iris/bitstream/10665/250441/1/9789241565394-eng.pdf?ua=1 (accessed 10th May 2017).
20. Houben RMGJ, Dodd PJ. The Global Burden of Latent Tuberculosis Infection: A Re-estimation Using Mathematical Modelling. *PLoS medicine* 2016; **13**(10): e1002152.
21. Lönnroth K, Castro KG, Chakaya JM, et al. Tuberculosis control and elimination 2010-50: cure, care, and social development. *The Lancet* 2010; **375**(9728): 1814-29.
22. 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 infectious diseases* 2016; **16**(1): 71.
23. WHO. Global Tuberculosis Report 2013. 2013. http://www.who.int/tb/publications/global_report/en/.
24. WHO. Global Tuberculosis report 2012. 2012.
25. World Health Organization. Global Tuberculosis Report 2015, 2015.
26. World Health Organization. The End TB Strategy (WHO/HTM/TB/2015.19). 2015. http://www.who.int/tb/End_TB_brochure.pdf?ua=1 (accessed 3rd January 2017).
27. Dye C, Garnett GP, Sleeman K, Williams BG. Prospects for worldwide tuberculosis control under the WHO DOTS strategy. *Lancet (London, England)* 1998; **352**(9144): 1886-91.
28. Dye C, Lönnroth K, Jaramillo E, Williams BG, Raviglione M. Trends in tuberculosis incidence and their determinants in 134 countries. *Bulletin of the World Health Organization* 2009; **87**(9): 683-91.
29. Marais BJ, Raviglione MC, Donald PR, et al. Scale-up of services and research priorities for diagnosis, management, and control of tuberculosis: a call to action. *Lancet (London, England)* 2010; **375**(9732): 2179-91.
30. Fletcher HA, Schrager L. TB vaccine development and the End TB Strategy: importance and current status. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 2016; **110**(4): 212-8.
31. Dye C, Glaziou P, Floyd K, Raviglione M. Prospects for tuberculosis elimination. *Annu Rev Public Health* 2013; **34**: 271-86.
32. Abu-Raddad LJ, Sabatelli L, Achterberg JT, et al. Epidemiological benefits of more-effective tuberculosis vaccines, drugs, and diagnostics. *Proceedings of the National Academy of Sciences* 2009; **106**(33): 13980-5.
33. Balcells ME, Thomas SL, Godfrey-Faussett P, Grant AD. Isoniazid Preventive Therapy and Risk for Resistant Tuberculosis. *Emerging Infectious Diseases* 2006; **12**(5): 744-51.
34. World Health Organization. Reported estimates of BCG coverage. 2015. http://apps.who.int/immunization_monitoring/globalsummary/timeseries/tscoveragebcg.html (accessed 23rd December 2015).

35. Mangtani P, Abubakar I, Ariti C, et al. Protection by BCG Vaccine Against Tuberculosis: A Systematic Review of Randomized Controlled Trials. *Clinical Infectious Diseases* 2014; **58**(4): 470-80.
36. Abubakar I, Pimpin L, Ariti C, et al. Systematic review and meta-analysis of the current evidence on the duration of protection by bacillus Calmette-Guerin vaccination against tuberculosis. *Health technology assessment (Winchester, England)* 2013; **17**(37): 1-372, v-vi.
37. Trunz BB, Fine P, Dye C. Effect of BCG vaccination on childhood tuberculous meningitis and miliary tuberculosis worldwide: a meta-analysis and assessment of cost-effectiveness. *Lancet (London, England)* 2006; **367**(9517): 1173-80.
38. Fifteen year follow up of trial of BCG vaccines in south India for tuberculosis prevention. Tuberculosis Research Centre (ICMR), Chennai. *The Indian journal of medical research* 1999; **110**: 56-69.
39. Karonga Prevention Trial Group. Randomised controlled trial of single BCG, repeated BCG, or combined BCG and killed Mycobacterium leprae vaccine for prevention of leprosy and tuberculosis in Malawi. *Lancet (London, England)* 1996; **348**(9019): 17-24.
40. Stein SC, Aronson JD. The occurrence of pulmonary lesions in BCG-vaccinated and unvaccinated persons. *Am Rev Tuberc* 1953; **68**(5): 695-712.
41. The Global Advisory Committee on Vaccine Safety. Global Advisory Committee on Vaccine Safety, 29–30 November 2006. *Weekly Epidemiological Report* 2007; **2017**(4th March).
42. von Reyn CF, Mtei L, Arbeit RD, et al. Prevention of tuberculosis in Bacille Calmette-Guerin-primed, HIV-infected adults boosted with an inactivated whole-cell mycobacterial vaccine. *AIDS (London, England)* 2010; **24**(5): 675-85.
43. von Reyn CF. Inactivated whole cell NTM vaccine for the prevention of tuberculosis SRL 172 DAR-901. Presented at 2016 TBVI Symposium, Les Diablarets, Switzerland. 2016. http://www.tbvi.eu/wp-content/uploads/2016/02/Presentation-Fordham-von-Reyn.pdf (accessed 26th July 2017).
44. Dartmouth-Hitchcock Medical Center. DAR-901 TB Booster Vaccine to Prevent TB in Adolescents (DAR-PIA). 2016. https://clinicaltrials.gov/ct2/show/NCT02712424 (accessed 26th July 2017).
45. Aeras. The Aeras Annual Report 2015. 2015. http://www.aeras.org/annualreport2015 (accessed 20th January 2017).
46. Anhui Zhifei Longcom Biologic Pharmacy Co. Phase III Clinical Study of Efficacy and Safety of Vaccae™ to Prevent Tuberculosis. 27th December 2016. https://clinicaltrials.gov/show/NCT01979900 (accessed 3rd January 2017).
47. GlaxoSmithKline. Study to Evaluate the Efficacy of GlaxoSmithKline (GSK) Biologicals' Candidate Tuberculosis (TB) Vaccine in Adults (NCT01755598). May 2017 2012. https://clinicaltrials.gov/show/NCT01755598 (accessed 3rd May 2017).
48. Aeras. A Randomized, Placebo Controlled, Partially Blinded Phase II Study to Evaluate Safety, Immunogenicity, and Prevention of Infection With Mycobacterium Tuberculosis of AERAS-404 and BCG Revaccination in Healthy Adolescents (040-404). 23rd January 2017 2014. https://clinicaltrials.gov/ct2/show/NCT02075203 (accessed 26th July 2017).
49. Serum Institute of India Pvt. Ltd. Study to Evaluate the Safety and Immunogenicity of VPM1002 in Comparison With BCG in HIV-exposed/-Unexposed

Newborn Infants in South Africa. 22nd November 2016 2015. https://clinicaltrials.gov/ct2/show/NCT02391415 (accessed 26th July 2017).
50. Aeras and TBVI. TB Vaccine Research and Development: A Business Case for Investment. http://www.aeras.org/pdf/TB_RD_Business_Case_Draft_3.pdf (accessed 14th January 2017 2017).
51. Brennan MJ, Thole J. Tuberculosis vaccines: a strategic blueprint for the next decade. *Tuberculosis (Edinburgh, Scotland)* 2012; **92 Suppl 1**: S6-13.
52. Knight GM, Griffiths UK, Sumner T, et al. Impact and cost-effectiveness of new tuberculosis vaccines in low- and middle-income countries. *Proc Natl Acad Sci U S A* 2014; **111**(43): 15520-5.
53. Wheeler CM, Skinner SR, Del Rosario-Raymundo MR, et al. Efficacy, safety, and immunogenicity of the human papillomavirus 16/18 AS04-adjuvanted vaccine in women older than 25 years: 7-year follow-up of the phase 3, double-blind, randomised controlled VIVIANE study. *The Lancet Infectious Diseases* 2016; **16**(10): 1154-68.
54. Skinner SR, Szarewski A, Romanowski B, et al. Efficacy, safety, and immunogenicity of the human papillomavirus 16/18 AS04-adjuvanted vaccine in women older than 25 years: 4-year interim follow-up of the phase 3, double-blind, randomised controlled VIVIANE study. *The Lancet* 2014; **384**(9961): 2213-27.
55. Centre for Reviews and Dissemination. CRD's guidance for undertaking reviews in health care. https://www.york.ac.uk/media/crd/Systematic_Reviews.pdf (accessed 5th March 2016).
56. Harris RC, White RG, Sumner T, Knight GM. Systematic review of models exploring the epidemiological impact of novel TB vaccines. PROSPERO 2016:CRD42016033266 2016. http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42016033266 (accessed March 7th 2016).
57. Liu S, Li Y, Bi Y, Huang Q. Mixed vaccination strategy for the control of tuberculosis: A case study in China. *Math Biosci Eng* 2017; **14**(3): 695-708.
58. Arregui S, Sanz J, Marinova D, et al. A data-driven model for the assessment of age-dependent patterns of Tuberculosis burden and impact evaluation of novel vaccines. *bioRxiv* 2017: Online first.
59. 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): 433-7.
60. Touray K, Adetifa IM, Jallow A, et al. Spatial analysis of tuberculosis in an urban west African setting: is there evidence of clustering? *Trop Med Int Health* 2010; **15**(6): 664-72.
61. INDEPTH. INDEPTH Network: Member HDSSs. 2017. http://www.indepth-network.org/member-centres (accessed 7th August 2017).
62. Tadesse T, Demissie M, Berhane Y, Kebede Y, Abebe M. The clustering of smear-positive tuberculosis in Dabat, Ethiopia: a population based cross sectional study. *PloS one* 2013; **8**(5): e65022.
63. Munch Z, Van Lill SW, Booysen CN, Zietsman HL, Enarson DA, Beyers N. Tuberculosis transmission patterns in a high-incidence area: a spatial analysis. *Int J Tuberc Lung Dis* 2003; **7**(3): 271-7.

64. MacPherson P, Choko AT, Webb EL, et al. Development and validation of a global positioning system-based "map book" system for categorizing cluster residency status of community members living in high-density urban slums in Blantyre, Malawi. *American journal of epidemiology* 2013; **177**(10): 1143-7.

CHAPTER 3: Epidemiological impact of TB vaccine characteristics

Summary of Chapter 3

In this chapter I describe the development of a mathematical model of age-stratified TB demographics and epidemiology in China. Firstly, modelled epidemiological projections to 2050 without introduction of new TB vaccines are described (*objective 2a*). Using this calibrated epidemiological model, the population-level epidemiological impact of a wide range of vaccine characteristics was explored in depth, and incidence and mortality results in terms of absolute and rate reductions in incidence and mortality are reported (*objective 2b and c*).

I developed the modelling research in this chapter as described below. The research question was identified based upon research gaps identified from the literature review in Chapter 2, and through discussions with Willem Hanekom at the Bill and Melinda Gates Foundation (part-funder of the work in this chapter). Collaborators at Aeras, WHO and TBVI also contributed to scoping the research question and outcomes. The methodology was developed by myself with input from Professor Richard White, Dr Tom Sumner and Dr Gwenan Knight. I collected the data for parameterisation and calibration of the model from secondary sources. The core model structure was based upon a model developed by Knight et al.,[1] but with major adaptations to capture the demographics and epidemiology in China and a wide variety of vaccine characteristics and implementation strategies, including 'leaky' mechanisms of vaccine efficacy against infection and disease, incorporation of social mixing patterns, and development of model outcomes (e.g. incidence rate reduction, NNV). I carried out the model calibration, with guidance from Dr Tom Sumner. Interpretation of results and writing of the chapter were my own with feedback from my supervisors.

The results presented in this chapter have been submitted as an internal report to the Bill and Melinda Gates Foundation (BMGF), and presented to the WHO new TB vaccine advisory group (24th May 2017)

CHAPTER 3: Epidemiological impact of TB vaccine characteristics

3.1 The need to inform TB vaccine development in China

As discussed in Chapter 2, China is the third largest contributor to global incident TB, reporting 918,000 (range: 788,000-1,060,000) new cases and 376,000 (range: 35,200-41,500) TB deaths in 2015.[2] In addition, China is on the WHO multidrug resistant TB high burden list.[2] It is anticipated that new vaccines would likely have similar prophylactic efficacy against drug sensitive and drug resistant TB, therefore a new vaccine could be an important part of preventative measures against drug sensitive, and primary and acquired drug resistant TB.

China is also a WHO high burden country with 10% of the current population aged 60 years or above.[3] The proportion of the population constituted by the elderly (≥65 years) is anticipated to almost triple between 2010 and 2050, to 23.9% of the total population In China.[4]

Although the population proportion aged 60 years and above is highest in developed countries, it has been estimated that by 2050, nearly 80% of the world's older population will be living in less developed regions.[3] Three of the WHO's high TB burden countries already estimate at least 1 in 10 of their population are aged at least 65 years (China 10%, Thailand 10%, and Russia 13.6%).[2,5] With the ongoing demographic ageing occurring globally, including in many countries with high burden of TB disease, the results of such modelling in China could be indicative of possible future trends in other ageing countries (e.g. Thailand). A non-vaccine model has estimated that in the future there is the potential for large influence of reactivation disease in China's future epidemic.[6]

Given China's important contribution to global TB burden, and the fact that there is currently a TB vaccine candidate in phase III trials in China,[7] there are surprisingly few

research studies using modelling to estimate the potential impact of new TB vaccines.[8] The majority of modelling papers exploring potential future TB vaccine impact are in global or India-like settings,[8] where the epidemic remains younger adult and new infection driven. Modelling in China's epidemiological setting is needed to understand the characteristics required to maximise the potential impact of new vaccines in this setting, by informing target product profiles, clinical development plans, candidate stage gating and implementation strategies.

There are currently 13 candidates in the clinical development pipeline for new TB vaccines,[9] representing a variety of antigens delivered through different vaccine modalities (e.g. whole cell, adjuvanted protein, viral vector), with the potential for relatively diverse vaccine characteristics and indications.[10] However, scientific challenges in the current TB vaccine development environment, such as the complexity of the natural history of TB disease and the lack of an immune correlate of protection, mean that in it is unlikely that early successful candidates will achieve all desirable vaccine characteristics. Therefore, not only is modelling required to help maximise vaccine impact by informing development strategy, modelling can also help identify the characteristics required to achieve the lowest acceptable level of impact to inform the 'minimum' profile in the TPP, which is important for informing sample size calculations and stage-gating decisions between trial phases.

There is a clear need for TB vaccine modelling in China to inform development and possible implementation of the candidate currently in trials, and of other pipeline candidates that could potentially be developed in this high burden country.

Therefore, modelling was conducted to fill the most pressing research gap for TPP development, exploring vaccine efficacy against infection, vaccine efficacy against disease, duration of protection, and the relative impact of pre- versus post-infection vaccines. This research was also an opportunity to improve on the modelled age specificities of the model, such as age structured demographics, age-specific contact patterns, and historical trends by age in burden of disease, aiming to improve the

future baseline projections in this first piece of research in chapter 3, and to allow a more detailed exploration of age-targeting of vaccines in chapter 4. The evidence generated by this work can be used to inform evidence-based funding allocation, efficient vaccine development, and maximization of future vaccine impact.

3.2 Aims and objectives

The research presented in this chapter contributes towards the first aim of the thesis, and fulfils thesis objectives 2a-c identified in Chapter 1:

2. a) Develop a mathematical model calibrated to epidemiological and demographic temporal and age distribution trends in China, accounting for uncertainty in natural history parameters, to predict the evolution of the TB epidemic over the 2050 time horizon.

   b) Using the calibrated China TB model, simulate the introduction of new TB vaccines to investigate the population-level epidemiological impact of varying TB vaccine characteristics identified in objective 1.

   c) Identify the combination of vaccine characteristics that would be most likely to deliver pre-specified minimum incidence rate reductions compared to the no new vaccine baseline in 2050.

3.3 Methods

3.3.1 Model structure

An age-structured dynamic compartmental transmission model (Figure 3.1) described by a series of difference equations (section 3.6.1) was developed and calibrated to epidemiological and demographic data from China to estimate TB burden in the baseline scenario (no new vaccine intervention) up to 2050, and a series of vaccination scenarios implementing new TB vaccines 2025-2050. The model structure is outlined below. The model was programmed in R.[11]

Consistent with other models in the literature,[8] the natural history of tuberculosis was described by five epidemiological states: uninfected (S), latently infected (L), infectious active TB disease (I), non-infectious active TB disease (NI), and recovered from active TB disease (R) (Figure 3.1). The uninfected population was defined as those never infected with *Mycobacterium tuberculosis,* latent as the population remaining infected without developing clinical symptoms, infectious disease as those with bacteriologically-positive pulmonary TB, non-infectious as patients with bacteriologically-negative pulmonary TB or extra-pulmonary TB, and recovered as those who have either received successful treatment or naturally recovered from active disease. Each infection state was represented by an unvaccinated and a vaccinated stratum, and age was modelled in single years from 0-100 years to provide granularity in demographic, natural history and vaccination parameters.

Newborns entered the system as uninfected at rate B[*k*], where *k* is the calendar year. Transmission to uninfected individuals occurred at rate λ, dependent upon the age-wise proportion of the population with infectious disease at a given time step, the age-specific contact patterns ($\eta[m, y]$) between the infectious and disease-free populations in that time step, and the probability of transmission per respiratory contact (*z*). Of the uninfected population becoming infected, a proportion (*p*) experienced fast progression directly to active disease and the remainder (1-*p*) entered the latent state. Latently infected individuals could develop active disease

through 'slow progression' of the existing infection (*v*) or 'fast progression' upon reinfection (λ*px*), where *x* represents protection against development of reinfection disease due to the immune response to existing infection. For all new active cases, regardless of whether fast or slow progressors, a proportion (*f*) developed bacteriologically positive active disease, and the remainder (1-*f*) developed bacteriologically-negative 'non-infectious' disease. Individuals with non-infectious TB could progress to infectious disease at rate *w*.

As TB patients become rapidly non-infectious once initiating appropriate treatment,[12] the modelled case detection (CDR) and effective treatment (CoT) of new active cases moved the population successfully detected and treated directly to the recovered state. For bacteriologically-negative disease, the case detection ratio (CDR) was scaled down by a factor *e*, as detection of bacteriologically negative and extrapulmonary disease is lower.[13,14] Undetected cases entered the relevant active disease state and became prevalent cases. Prevalent cases could be removed by natural cure (*n*) to the recovered state, or removed from the system by TB death (u_i/u_{ni}) or all-cause mortality (*u*). Individuals with non-infectious disease could convert to infectious disease at rate *w,* at which point there was another opportunity for case detection at the bacteriologically-positive disease case detection rate given the likely alteration in symptoms.

The population in the recovered state could be reinfected at rate λx to develop primary active disease (*p*) or enter the latent pool (1-*p*). Relapse (*r*) from the recovered state to return to the active disease state occurred at higher rates than the latent population due to the higher risk of developing disease in this population.

All-cause mortality (*u*) was age-specific and occurred at the same rates across all natural history states. TB-related mortality occurred in infectious or non-infectious disease states at rates u_i and u_{ni}, respectively, scaled by overall (*rmortTB*) and age-specific calibration factors (*uiscale*).

Each state was represented by a vaccinated and unvaccinated stratum; therefore, vaccination was represented by a transition to the vaccinated stratum ($\theta_S, \theta_L \text{ and } \theta_R$). Vaccine efficacy was modelled as "leaky", therefore all infection and disease states were possible in the vaccinated stratum, but vaccination altered the relevant natural history parameters to reduce progression to infection and/or disease. Vaccine efficacy against infection (*effI*) was modelled by multiplying the infection parameter (λ) by 1-*effI*. Vaccine efficacy against disease (*effD*) was applied in the same manner to development of disease parameters, which were the proportion developing primary active disease (*p*), and risk of reactivation from the latent (*v*) and recovered (*r*) classes. Therapeutic vaccination was not considered, so transition from unvaccinated to vaccinated strata was only possible for the susceptible, latent and recovered populations ($\theta_S, \theta_L \text{ and } \theta_R$, respectively). Theta was equivalent to vaccine coverage in the given age group and year. Individuals departed the vaccinated stratum either through all-cause mortality (*u*) or reaching the end of the duration of protection and returning to the unvaccinated stratum (*d)*.

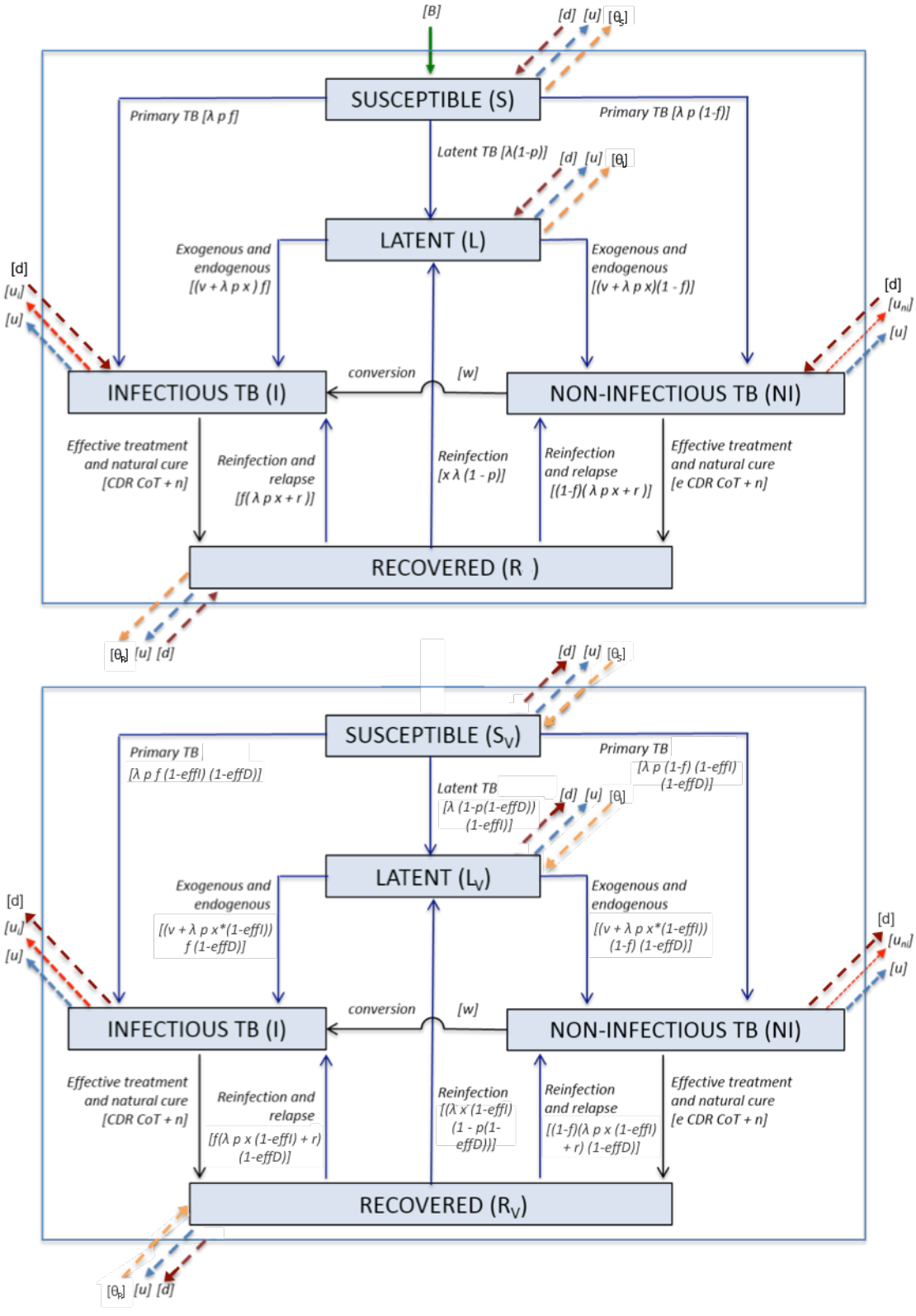


Figure 3.1: Model structure, showing the unvaccinated natural history stratum (top) and the vaccinated stratum (bottom). Figure adapted with permission from Knight et al. (2014) [1]

3.3.2 Model Parameters

Key model parameters are summarised below. A full description and justification of choice of parameter ranges are provided later in this chapter (section 3.6.2).

Natural history parameter ranges for calibration, detailed in Table 3.7 were based upon available literature, age-stratified where indicated by available data or biological plausibility. Social mixing between age strata was based upon social mixing pattern data from a study in Southern China.[15] Treatment success was based upon historical data, and assumed to plateau beyond 2011 due to high success rates achieved (95%).[16] To minimise short-term fluctuations in the WHO-reported case detection ratio (CDR) data,[17] a generalised logistic function was manually fitted to the WHO CDR estimates for 1990-2010.[17] BCG delivery was assumed to remain constant, so was not explicitly modelled. Scaling factors for the contact matrix, overall TB mortality, age-wise TB mortality, and case detection ratio by age were included as calibration parameters.

3.3.3 Model calibration

A two-stage calibration process was employed, with the first to calibrate demographics, then a second for burden of TB disease trends. Details of calibration data and justification for choice of methods are provided in section 3.6.3.

In the first stage, manual calibration of population size at initiation, birth rate and death rate parameters, was used to fit the model to age-stratified (0-14, 15-54, 55-64 and ≥65years) UN population division estimates for 2010 and projections for 2050.[18]

In the second stage, the model was calibrated to the bacteriologically-positive TB prevalence rate (≥15, 15-29, 30-59 and ≥60years) in 2000 and 2010;[19] TB notification rate (all-age, 0-14, 15-54, 55-64, and ≥65 years) in 2010;[20] and TB mortality rate (all-age, 0-14, 15-59 and ≥60 years) in 2010.[21] Country-specific age-stratified incidence

rates are not reported by the WHO, therefore the model was fitted to the WHO 2010 all-age incidence rate for China.[17] The model was compared to, but not fitted to, LTBI survey data for China, due to uncertainties in the sensitivity and specificity of the LTBI tests and representativeness of the LTBI survey sites.[22]

Calibration to these 18 country-specific epidemiological data sets was achieved using an adaptive Approximate Bayesian Computation-Markov Chain Monte Carlo (ABC-MCMC)-based method using a modified *easyABC* package.[23,24] First, the model was run for one million randomly generated parameters sets and the likelihood calculated. The 20 highest likelihood parameter sets were then used as seeds for an adaptive ABC-MCMC. Sequential parallel ABC-MCMC chains were employed, seeding with acceptances from the previous chains and adapting the acceptance criterion until a full model fit to all 18 data ranges was achieved. From the 2,152 model fits achieved in the final set of parallel chains, 1,000 parameter sets were randomly selected to provide median estimates and uncertainty ranges reflective of parameter uncertainty in the model.

3.3.4 Modelled intervention scenarios

The key vaccine characteristics and implementation assumptions are summarized in Table 3.1, and described in full below.

Table 3.1: Vaccine characteristics and implementation assumptions

Characteristic/Assumption	*Value*
Intrinsic vaccine characteristics	
Efficacy for prevention of infection (VE-POI, *effI*)	Range: 0-100% Interval: 10% (i.e. 0%, 10%, 20% etc.)
Efficacy for prevention of disease (VE-POD, *effD*)	Range: 0-100% Interval: 10%
Duration of protection (*D*)	Range: 2yrs-lifetime Values: 2yrs, 3, 5, 7, 10, 15, 20, 25, lifelong Waning: Exact, no Immunosenescent waning
Population protected by vaccine	Pre- and post-infection (P&PI), pre-infection (PRI) or post-infection (PSI)
Type of protection	Leaky
Implementation assumptions	
Epidemiological setting	China
Year of vaccine introduction	2025
Time horizon of vaccination scenarios	2025-2050
Vaccination schedule	Annual routine vaccination of 9 year olds from 2025 onwards; Mass vaccination of all aged ≥10 years in 2025 & mass campaigns at 10-yearly interval or duration of protection, whichever longer (scenario analysis with 5-yearly mass campaigns)
Coverage (*ϑ[k,j]*)	Routine: 80% Mass: 70%

Based upon the literature review in chapter 2, four intrinsic vaccine characteristics were varied in this research: vaccine efficacy for prevention of infection (VE-POI),

vaccine efficacy for prevention of disease (VE-POD), efficacy by host infection status (P&PI, PRI and PSI), and duration of protection (Figure 3.2).

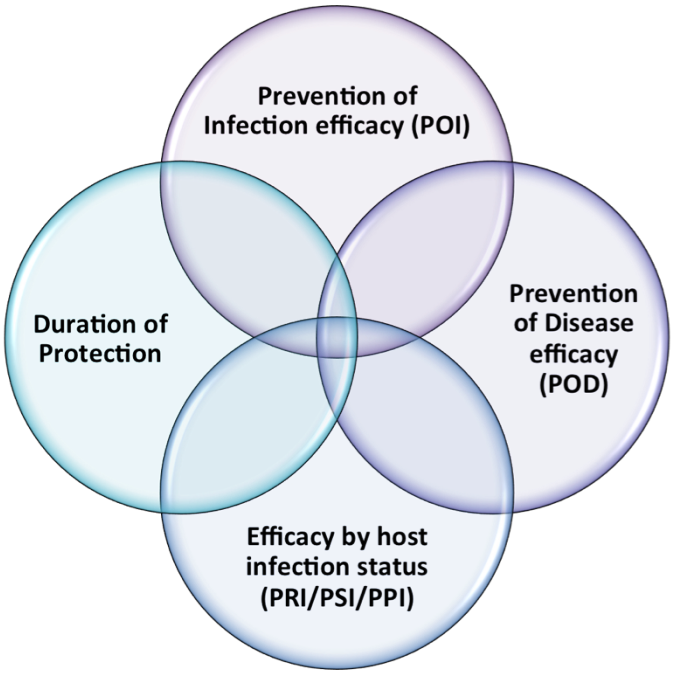


Figure 3.2: Four main new TB vaccine characteristics varied in the model

Both POI and POD vaccine efficacies were varied from 0-100% in 10% intervals and in combination with each other (i.e. 121 vaccine efficacy combinations). Efficacy was modelled as 'leaky' protection (otherwise known as 'degree' protection). With this type of protection, vaccine efficacy was implemented as a percentage reduction in the relevant natural history parameters. Prevention of infection vaccine efficacy was applied to the infection parameter, and prevention of disease efficacy was applied to development of disease parameters, which are the proportion developing primary active disease, and risk of reactivation from the latent and recovered classes. In this 'leaky' protection model, the vaccinated population could develop infection and disease, but at a lower rate than the unvaccinated population. The modelled vaccines were not assumed to affect infectiousness of disease, severity of disease, case fatality ratio, likelihood of natural cure, or case detection.

The primary vaccine outcomes were reported for vaccines effective both pre- and post-infection (P&PI), modelled as vaccine 'take' possible in susceptible, latently infected and recovered populations. As secondary analyses, pre-infection (PRI) vaccines effective only in susceptible populations, and post-infection (PSI) vaccines only effective in latently infected and recovered populations, were modelled.

Impact of duration of protection on the potential population-level impact of new vaccines is important information to inform the TPP. Duration of protection was assumed to be 2, 3, 5, 7, 10, 15, 20, 25 years or lifelong. Duration of protection was assumed to be exact, with no Immunosenescent waning in elderly populations.

Vaccine implementation in China was assumed from 2025 to 2050. Routine TB vaccination was delivered to nine year olds, assumed to co-deliver with HPV vaccine as part of the school-based vaccination platform. Although HPV is not currently implemented in China, WHO recommends delivery to ages 9-13 years,[25] and vaccination at the lower end of this range is anticipated long-term.[26] Mass campaigns vaccinating adolescents and adults (>9 years old) were initiated in 2025, and modelled with a frequency of the duration of protection or 10 years, whichever was longer. Although this covers a wider age group than might be anticipated once the vaccine is available, this ensures the peak ages of infection and disease will be covered by this vaccination campaign. A 10-year interval between mass campaigns was considered the most feasible, informed by stakeholder experience of other campaigns.[27,28]

As there is no existing HPV vaccination in China to inform coverage of routine vaccination,[29] the modelled coverage of routine vaccination was a conservative assumption (80%), based upon a combination of gross secondary school enrolment ratio from China (94.3%) and the HPV coverage achieved in 9 year olds in South Africa (87%).[30,31] Coverage of mass vaccination was modelled at 70%, based upon expert opinion and the lower end of Menafrivac mass campaign data from SSA (70-98% of

1-29 year olds) given the relatively low routine elderly influenza vaccination coverage in China (36-49%).[32-35]

3.3.5 Outcomes

In order to inform TB vaccine TPPs, the research question focused on identifying the vaccine characteristics most likely to deliver pre-specified ranges (e.g. 20-29%, 30-39% etc.) of population-level reduction in TB incidence rate in 2050. Therefore, the primary outcomes of interest were the incidence rate reduction in 2050 with new vaccine compared to no new vaccine (presented as heat maps), and the combinations of characteristics required to achieve pre-specified ranges of incidence rate reduction (presented as example TPP tables).

The percentage incidence rate reduction compared to the no new vaccine baseline in 2050 was estimated for all evaluated values of VE-POI, VE-POD and duration of protection, with the median and range estimated for each from the 1000 calibrated parameter sets.

The definition of "lowest acceptable level of impact" may vary depending upon the burden of disease, other care and prevention options available, and cost of the vaccine. China's high DOTs coverage and treatment success leaves minimal opportunities for reducing burden with currently available alternatives, therefore even relatively small reductions in disease burden through vaccine programmes could be of value. Discussion with stakeholders identified 20-29% incidence rate reduction in 2050 compared to no new vaccine baseline as the absolute minimum level of impact from a new TB vaccine that could potentially be of interest (Personal communication, Willem Hanekom, 15th December 2016).[36] This was equivalent to reducing incidence from a modelled no new vaccine median baseline in 2050 of 34 cases per 100,000 population, down to 24-27 cases per 100,000 population.

From initial exploratory modelling, the shortest duration of protection able to achieve this minimum impact of 20-29% IRR was identified (5 years protection, when

mass campaigns were 10-yearly). In addition, the duration of protection above which minimal additional epidemiological gains could be achieved was also identified (10 years protection with 10-yearly mass campaigns). These durations were used to limit the number of scenarios presented in later analyses. To communicate the correlation between VE-POI and VE-POD, examples were also given at the extremes of each vaccine efficacy range. These were calculated for the vaccine scenarios using the model outputs from all 1000 fitted runs.

Although 20-29% was the absolute lowest acceptable incidence rate reduction, the level of incidence rate reduction required for the "minimum" vaccine profile in the TB vaccine TPP is not yet confirmed. Therefore, the median and ranges of VE-POI and VE-POD characteristics required to achieve incidence rate reductions in 10% increments from 20% to 79% were investigated.

In clinical trials, the trial outcome is usually either *M.tb* infection or TB disease. Rarely is a trial designed or powered to measure both. Therefore, if a desired population level impact is the goal, if measuring just one of these endpoints, it is helpful to understand any risks of making stage-gating decisions based upon a single endpoint. To give confidence that a vaccine would achieve the required level of impact, the measured efficacy outcome could be benchmarked against the modelled median for that efficacy in a 'worst case' scenario for the two unmeasured characteristics. The 'worst case' scenario was defined as zero vaccine efficacy against the unmeasured endpoint and the lowest acceptable duration (5 years).

Primary outcomes (for vaccines effective both pre-and post-infection):

1. Percentage population-level TB incidence rate reduction in vaccine scenarios compared to no new vaccine baseline in 2050 (informs heat maps).
2. Identify shortest duration of protection able to deliver at least 20-29% IRR with 10-yearly mass campaigns; and identify the duration of protection above which minimal additional epidemiological gains can be achieved (informs TPP table).
3. For the two durations of protection identified in primary outcome 2: Calculate median values and ranges for VE-POI and VE-POD required to achieve a population-level reduction in TB incidence rate in 2050 of 20-80% in 10% intervals (informs TPP table).

Secondary outcomes:

1. Cumulative TB cases and TB deaths averted during 2025-2050
2. Mortality rate reduction in 2050 compared to no new vaccine baseline
3. Incidence rate reduction in 2035 compared to no new vaccine baseline and cumulative cases averted 2025-2035
4. Maximum achievable and range of incidence rate reduction for vaccines with efficacy POI- or POD-only (as distinct from efficacy in both).
5. Minimum efficacy that would be required in one of the vaccine efficacy characteristics, to be confident that a vaccine would achieve a population-level reduction in TB incidence rate of 20-80% in 10% intervals in 2050 if the duration of protection and the unreported vaccine efficacy were at a minimum (5 years and 0%, respectively).
6. Primary outcomes repeated for vaccines effective either pre-infection or post-infection.

3.3.6 Scenario analyses

In the main analysis, mass campaigns were implemented every 10 years based upon expert opinion (Personal Communication, teleconference, 28th October 2016, BMGF, WHO, Aeras and TBVI). However, in some settings, or if the vaccine were shown to be highly effective but with short duration of protection, there may be sufficient political will to deliver mass campaigns more frequently. A scenario analysis was conducted to explore the primary outcomes with frequency of mass campaigns shortened to the shortest frequency considered feasible, which was every five years (Personal Communication, teleconference, 28th October 2016, BMGF, WHO, Aeras and TBVI).

3.4 Results

3.4.1 Baseline (no new vaccine) model

The baseline model fit is shown in Figure 3.3 to Figure 3.7. In the calibrated model, the incidence rate in China was estimated to continue to decline from 56.4 (UR: 51.6-59.5) per 100,000 in 2025 to 33.8 (UR: 27.8-37.9) per 100,000 population in 2050 (Figure 3.3). During the same period, mortality rates were predicted to decline from 2.0 (UR: 1.2-3.2) per 100,000 in 2025 to 1.1 (UR: 0.6-2.0) per 100,000 population (Figure 3.4).

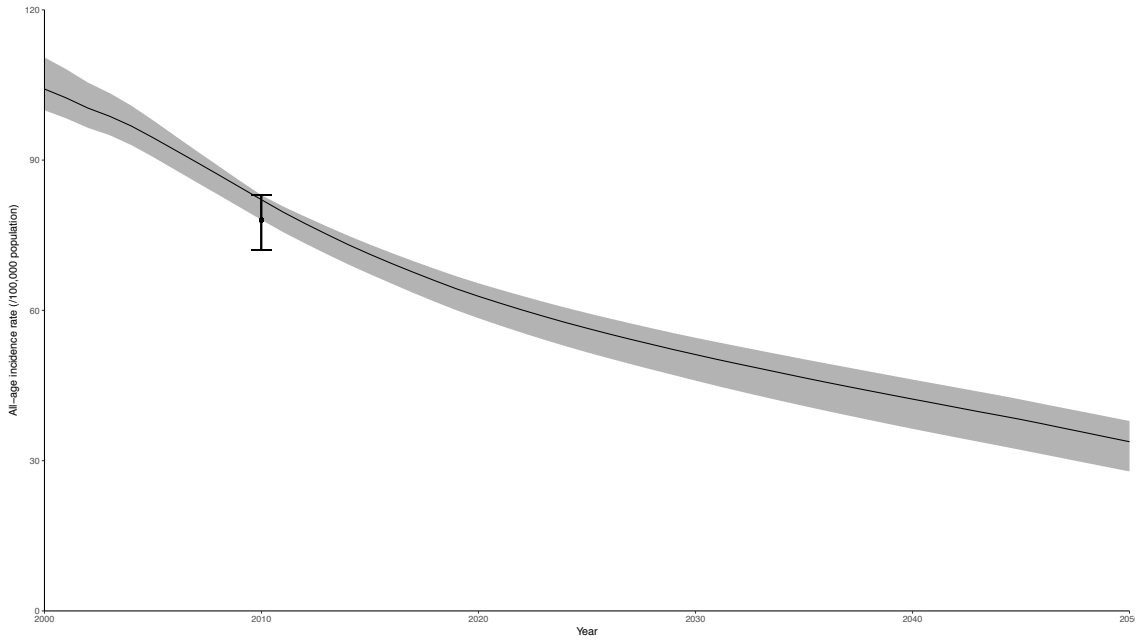


Figure 3.3: Modelled all-age incidence rate per 100,000 population 2000-2050. Black circle and vertical bar represents WHO estimates and ranges, solid horizontal line and shaded area represent modelled median and uncertainty range.

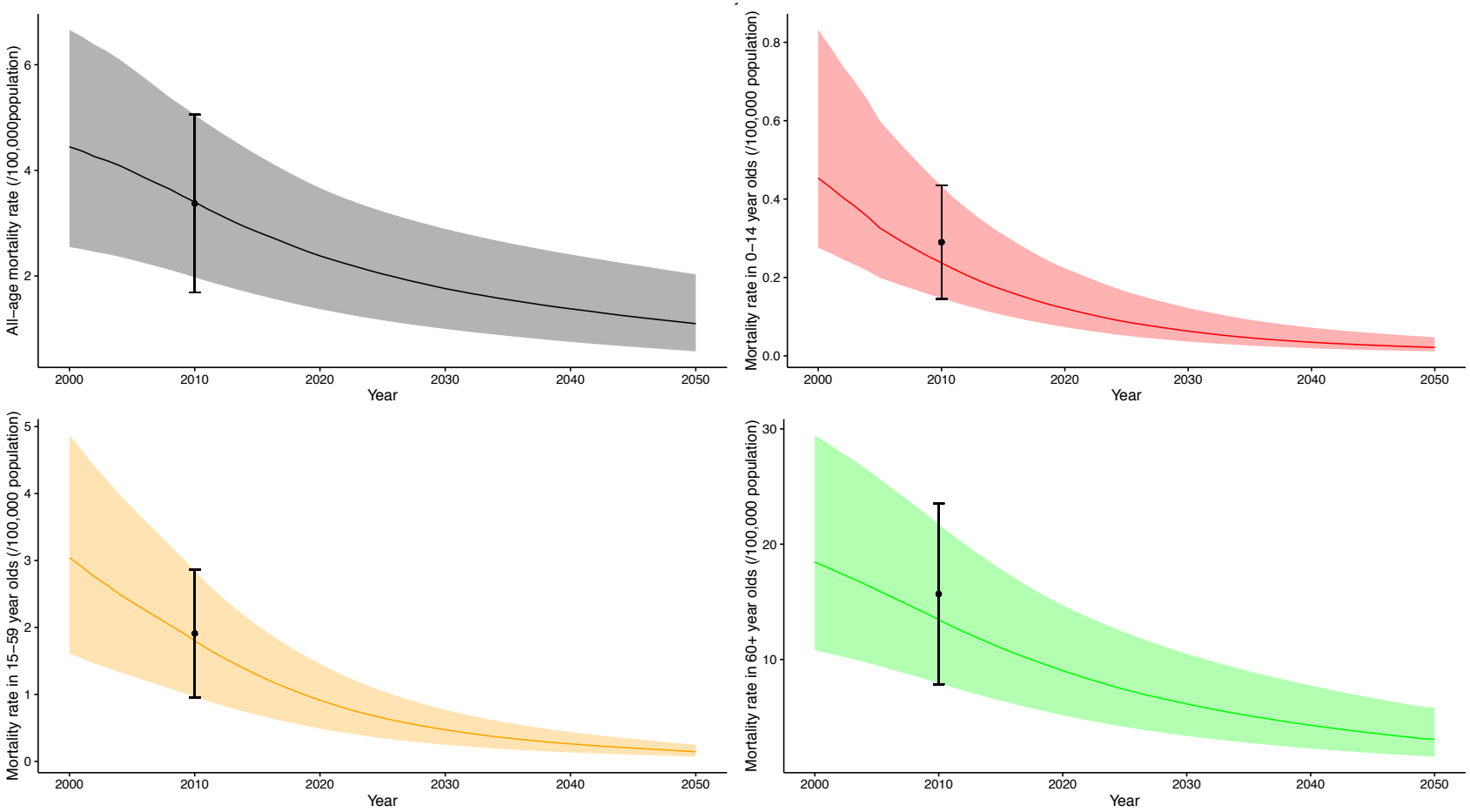


Figure 3.4: Modelled mortality rate per 100,000 population 2000-2050 for all-age population (top left, black), 0-14 year olds (top right, red), 15-59 year olds (bottom left, yellow), and ≥60 year olds (bottom right, green). Black circles and vertical bars represent empirical calibration data and estimated ranges, solid horizontal lines and shaded areas represent modelled medians and uncertainty ranges.

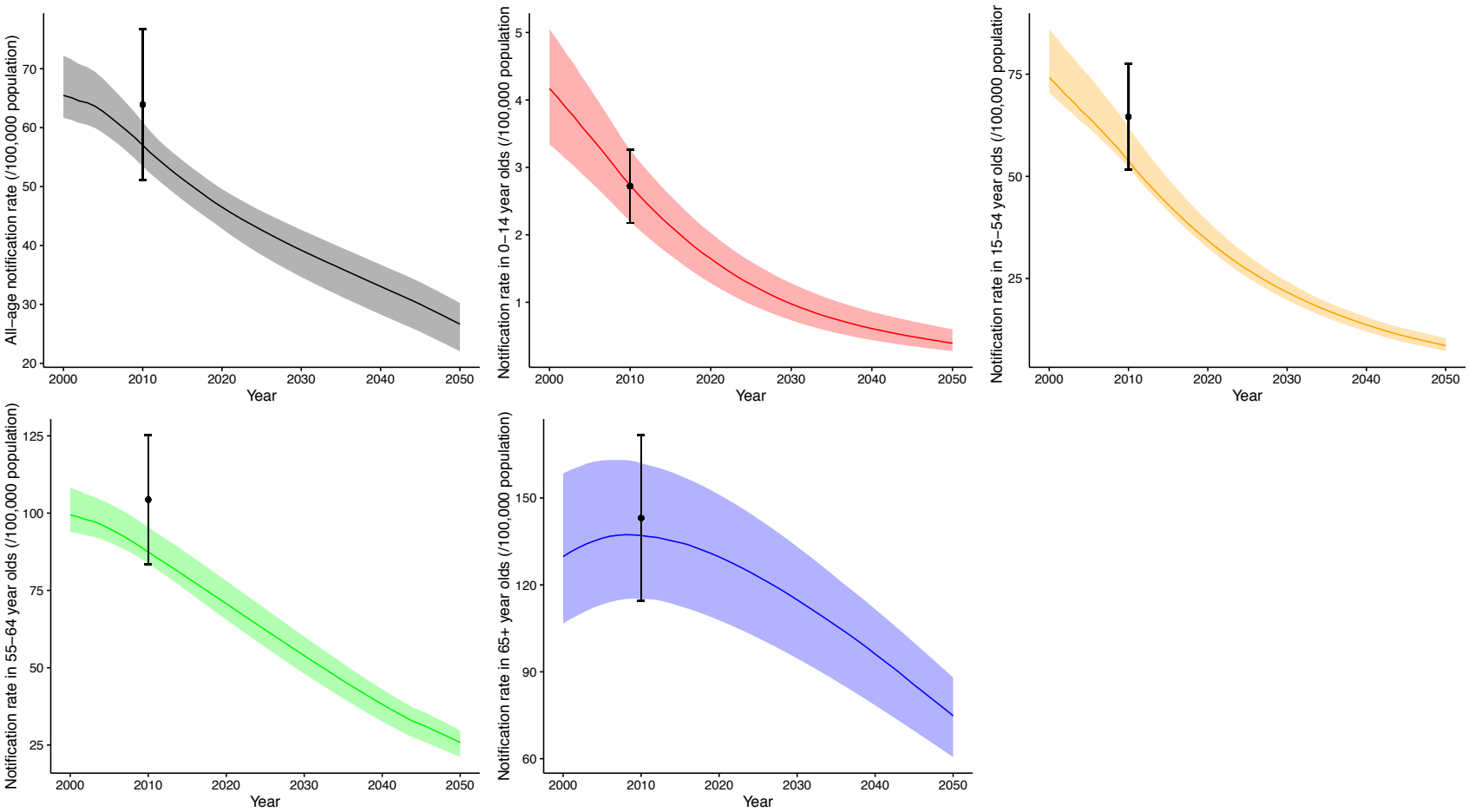


Figure 3.5: Modelled notification rate per 100,000 population 2000-2050 for all-age population (top left, black), 0-14 year olds (top centre, red), 15-54 year olds (top right, yellow), 55-64 year olds (bottom left, green), and ≥65 year olds (bottom centre, blue). Black circles represent WHO data and vertical bars the estimated ranges, solid horizontal lines and shaded areas represent modelled medians and uncertainty ranges.

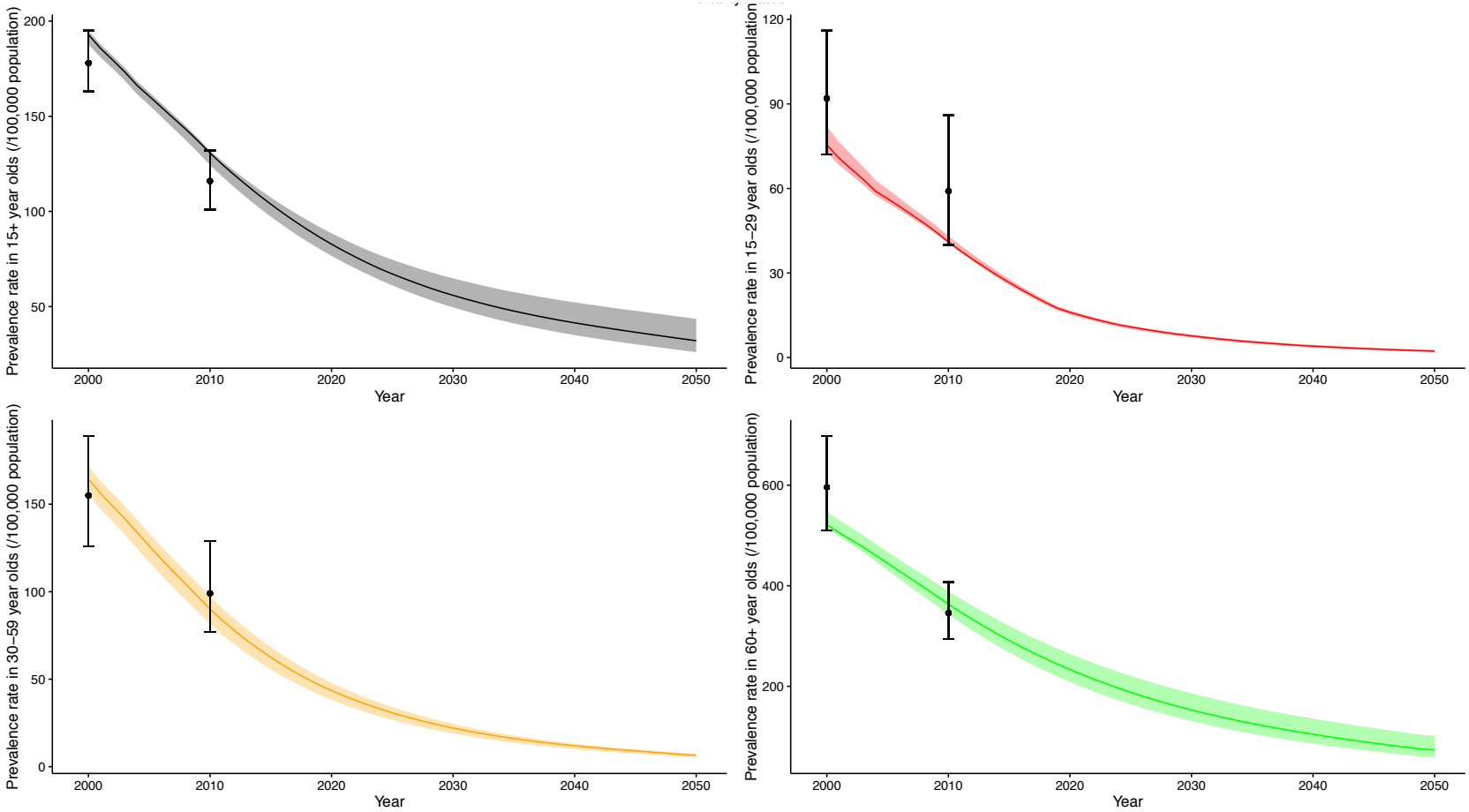


Figure 3.6: Modelled bacteriologically-positive tuberculosis prevalence rate calibration in 2000 and 2010 for ≥15 year olds (top left, black), 15-29 (top right, red), 30-59 year olds (bottom left, yellow), and ≥60 year olds (bottom right, green). Black circles and bars represent empirical calibration data, solid horizontal lines and shaded areas represent modelled medians and uncertainty ranges.

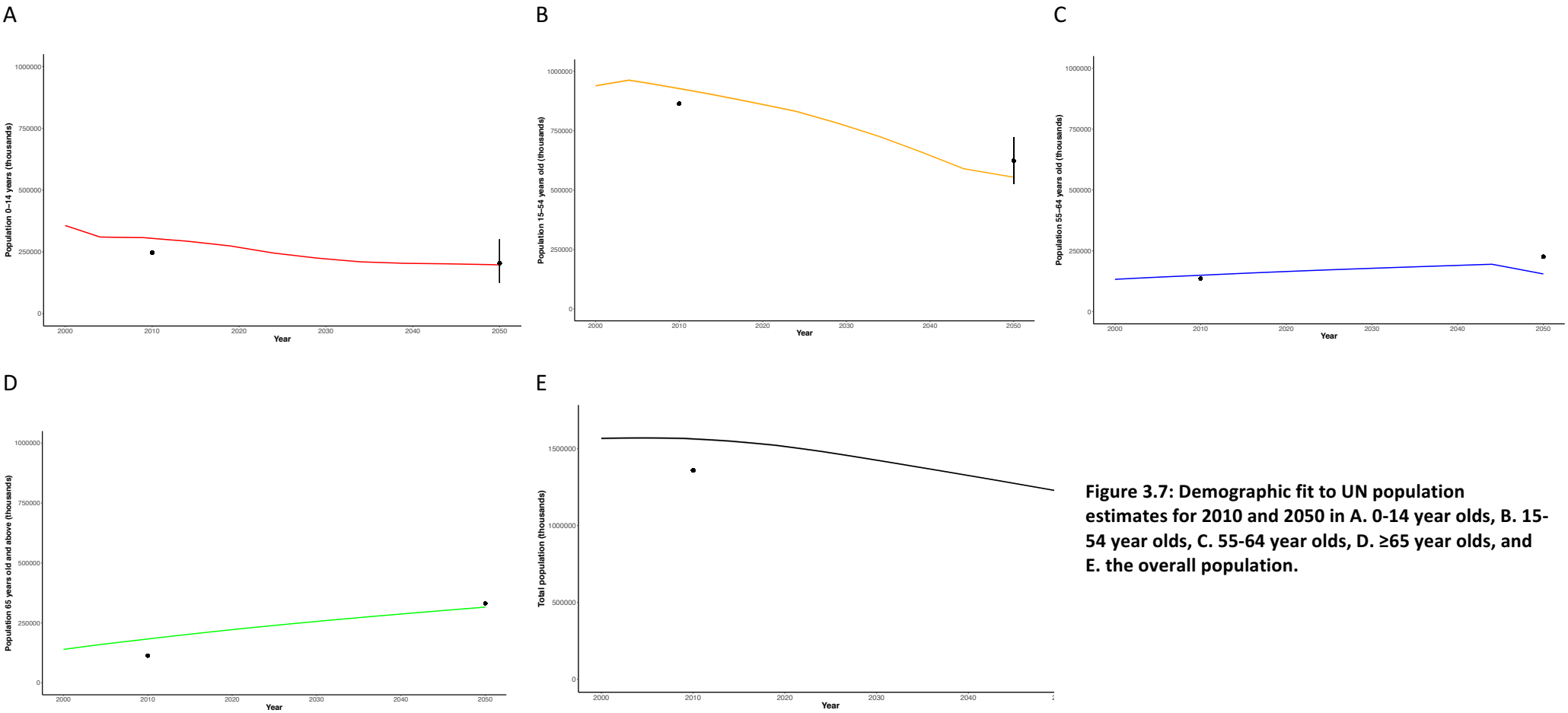


Figure 3.7: Demographic fit to UN population estimates for 2010 and 2050 in A. 0-14 year olds, B. 15-54 year olds, C. 55-64 year olds, D. ≥65 year olds, and E. the overall population.

Although not specifically calibrated to all six age groups or the 2000 and 2025 time points, in Figure 3.8 a comparison of the modelled population to the UN population estimates and projections confirms the validity of the modelled demographics.[4] Population ageing is clearly visible over the modelled time frame.

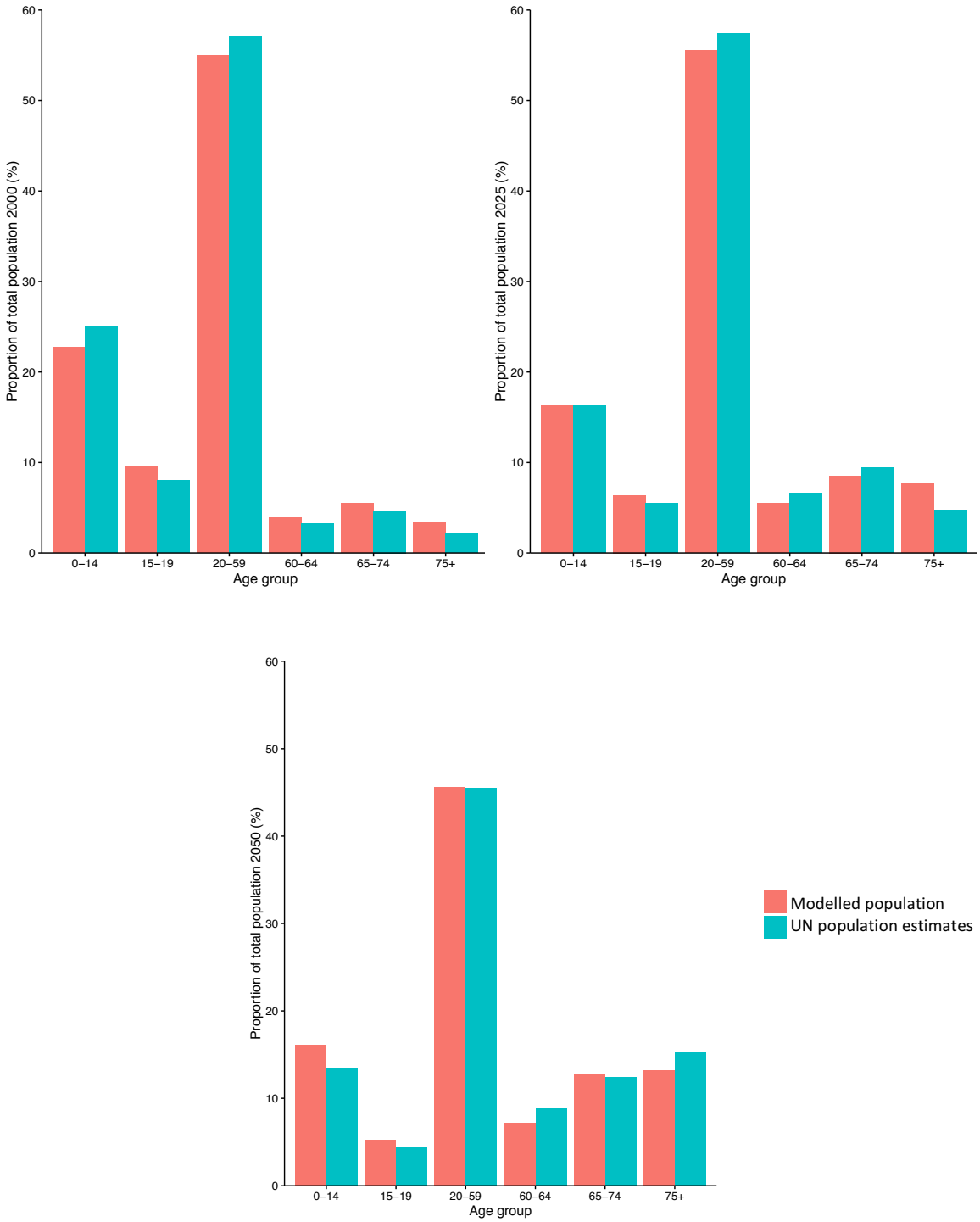


Figure 3.8: Comparison between age-stratified modelled (red) and UN population estimates (blue) for 2000 (top right), 2025 (top left) and 2050 (bottom).

The model was also not specifically calibrated to prevalence of infection data, but a comparison of the age-stratified modelled 2013 prevalence of infection was found to very closely match published IGRA data from a study in the same year (Figure 3.9).[22] The overall prevalence of infection in the population was 15·7% (UR: 12·8-18·7), ranging from 1·3% (UR: 1·0-1·6) in 5-9 year olds, to 37·0% (UR: 31·1-42·6) in ≥70 year olds.

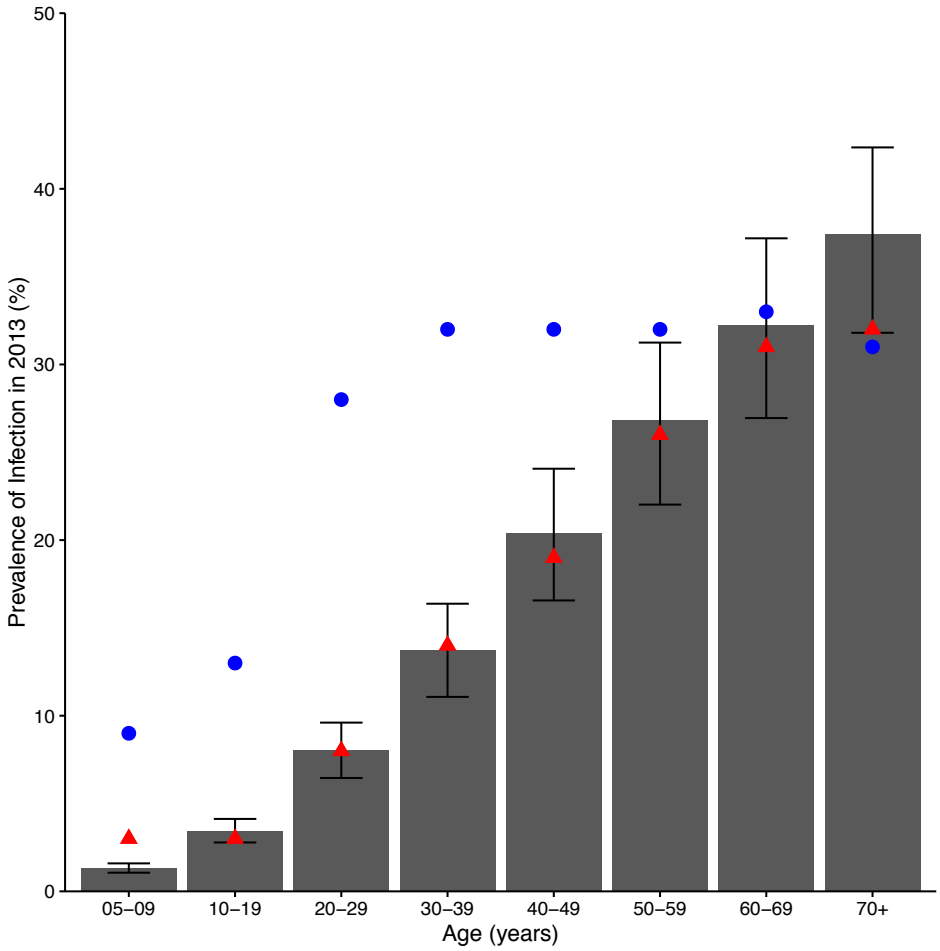


Figure 3.9: Age-stratified latent M.tb infection prevalence in 2013, modelled and empirical data. Grey bars and error bars denote modelled median estimates and uncertainty ranges. Empirical data from Gao et al. (2013) are TST prevalence as blue circles and Quantiferon as red triangles.[22]

Results from the epidemiological model are explored further in Chapter 4. However, important to the interpretation of the vaccine-related results in this chapter is the finding that China was demonstrated to be undergoing a transition from a transmission-driven to a reactivation-driven epidemic (Figure 3.10). In 2000, 28·6% (UR: 21·9-32·5) of all cases were estimated to be a result of reactivation, increasing to 75·1% (UR: 66·8-80·7) by 2050. A change in age distribution of cases was also observed in the model, due to the coupling of high but declining transmission rates with an ageing population in China. In 2000, 75·8% (UR: 71·3-81·0) in adolescents and adults (15-64 years), whereas by 2050 the vast majority occurred in the elderly (≥65 years: 74·5%; UR: 70·2-78·6). In the year of vaccine introduction, less than 1% of all new cases were estimated to occur in the vaccine-ineligible age group (<9 years, not shown).

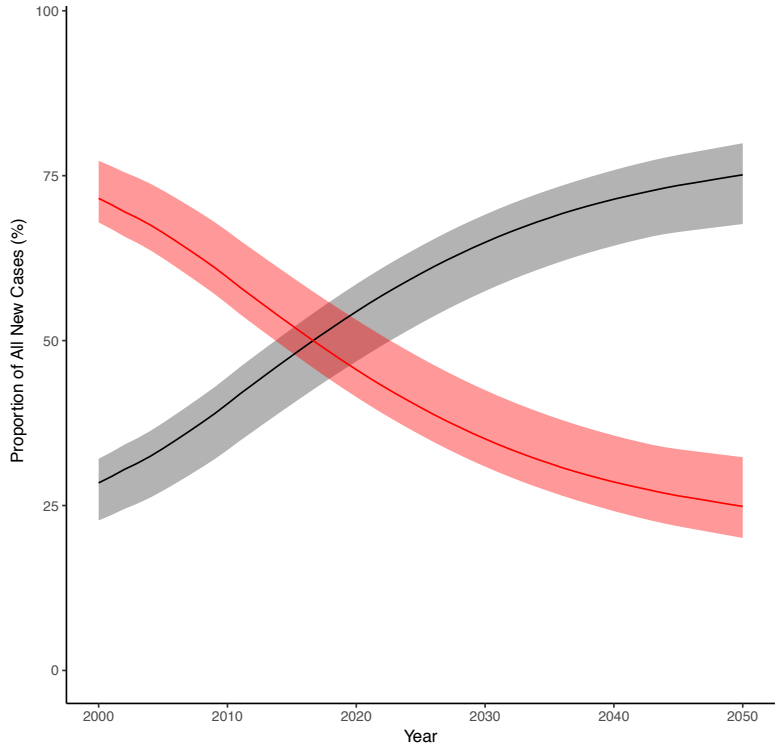


Figure 3.10: Proportion of all new active cases emerging from new infections (red) versus reactivation of existing infection (black) from 2000-2050. Lines are median values and shaded areas represent the uncertainty ranges.

3.4.2 New TB Vaccine Scenarios

3.4.2.1 Incidence rate reduction with P&PI vaccine compared to no new vaccine baseline in 2050

The predicted reduction in incidence rate in 2050 compared to no new vaccine, for P&PI vaccines of two years to lifelong duration of protection, by percentage vaccine efficacy against disease, and percentage vaccine efficacy against infection, are shown in Figure 3.11 below.

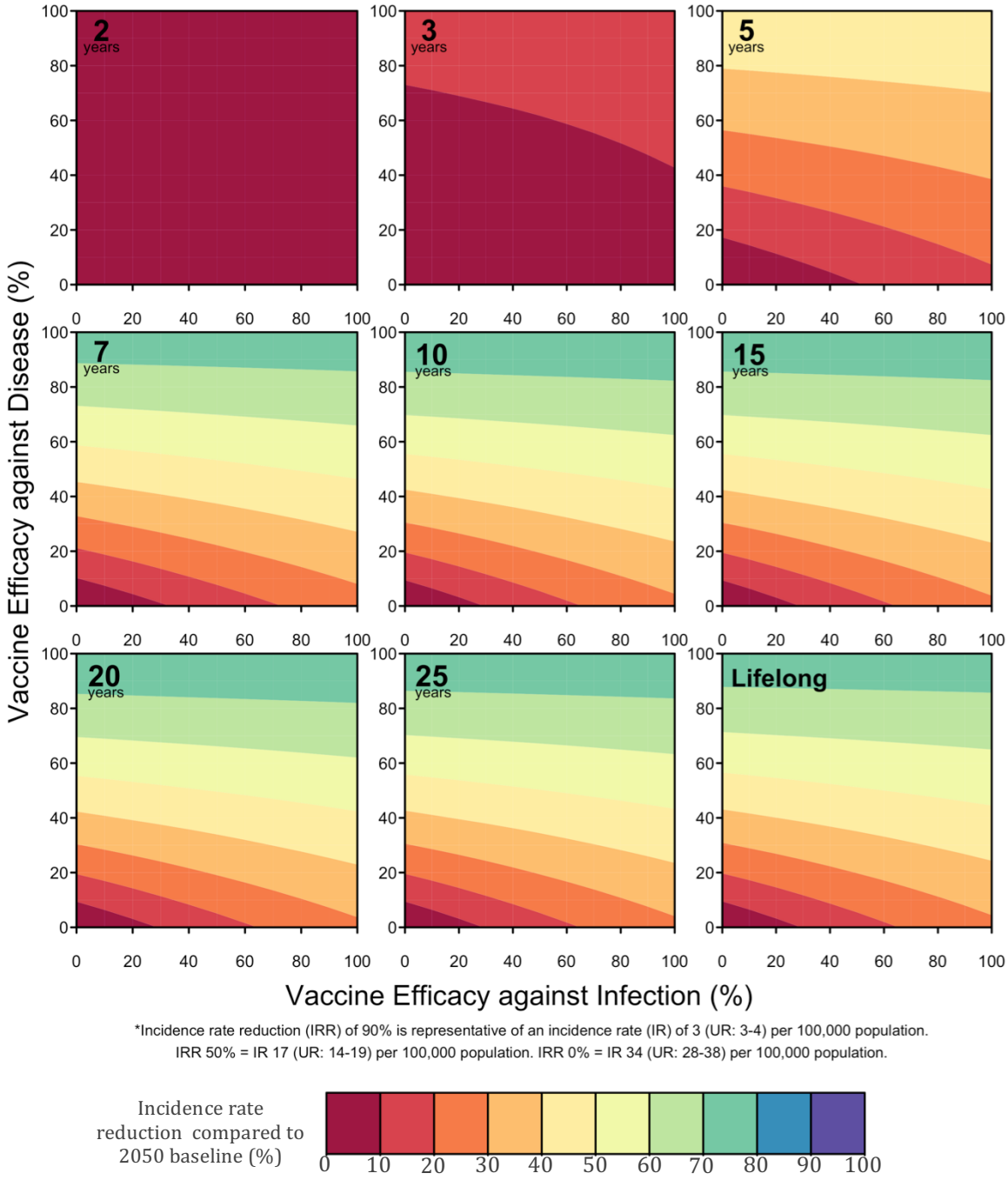


Figure 3.11: Median incidence rate reduction (%) for pre- and post-infection vaccine compared to no new vaccine baseline in 2050, by percentage efficacy against infection (x-axis), percentage efficacy against disease (y-axis), and duration of protection (as indicated in top left corner of panel), all with 10-yearly mass vaccination campaigns.

With the profiles and implementation assumptions explored, new TB vaccines implemented in 2025 reduced incidence rates compared to a no new vaccine baseline by a maximum of 79% (UR: 77-81%) by 2050, reducing incidence to 7 (UR: 6-8) cases per 100,000 population (100% VE against infection and disease, with 10 years protection) (Figure 3.11, centre panel). A vaccine with 20% efficacy against infection and disease providing 2 years duration of protection could achieve as little as 3% (UR: 2-3%) incidence rate reduction (IRR) in 2050 (Figure 3.11, top left panel).

For the absolute minimum acceptable impact of a 20-29% incidence rate reduction (Figure 3.11, dark orange contour, equivalent to an incidence rate of 27-30 per 100,000 population in 2050) to be achieved, a duration of protection of at least five years would be required assuming mass vaccination every 10 years. If at least a 50-59% IRR were required, at least seven years duration of protection would be required (Figure 3.11, light yellow contour). With two years duration of protection, less than 10% IRR would be achievable, regardless of vaccine efficacies (Figure 3.11, top left panel). With 5 years duration of protection, a vaccine completely protective against infection and disease could achieve 49% (UR: 47-53%) IRR (Figure 3.11, top right panel), and with 10 years duration of protection up to 79% (77-81%) IRR (Figure 3.11, centre panel). With a duration of protection above the maximum frequency of mass vaccination (i.e. above 10 years), a minimal difference in impact was observed compared to 10 years protection. Based upon these results, vaccines with 5 years (minimum acceptable impact) and 10 years (duration above which achieve minimal epidemiological gains) duration of protection with 10-yearly mass campaigns were explored in the subsequent analyses of primary outcomes.

For a pre- and post-infection (P&PI) vaccine, Figure 3.11 shows that the reduction in incidence is mostly dependent on the vaccine efficacy for prevention of disease, with limited sensitivity to VE-POI. Although increasing VE-POI does increase the incidence rate reduction in 2050, greater gains in impact are achieved with equivalent percentage increases in VE-POD. For example, with 10 years duration of protection and 60% vaccine efficacy against both infection and disease, an IRR of 57% (UR: 54-60%) would be expected; increasing VE-POD to 80% would increase the IRR to 68%

(UR: 66-71%), whereas increasing VE-POI to 80% would only provide a marginal improvement to an IRR of 58% (UR: 55-61%).

3.4.2.2 Median values for VE-POI and VE-POD of protection required to achieve a given population-level reduction in TB incidence rate

To achieve the minimum acceptable 20-29% incidence rate reduction in 2050 with 5 years duration of protection and 10-yearly mass campaigns, median vaccine efficacy for prevention of disease (VE-POD) and vaccine efficacy for prevention of infection (VE-POI) were 40% and 60%, respectively (Table 3.2, column 2). Therefore, a vaccine with 5 years protection, 40% VE-POD and 60% VE-POI would be very likely to provide this level of impact. However, these vaccine characteristics are highly correlated and therefore the ranges are wide (0-60% for VE-POD and 0-100% for VE-POI), and include all possible combinations that could potentially provide this level of impact. For example, for a vaccine effective against infection only (i.e. VE-POD=0%), 90-100% protection would be required to achieve this impact, whereas with a vaccine with no protection against infection (i.e. VE-POI=0%), 40-60% VE-POD would be sufficient to achieve this level of impact (Table 3.3).

Exploring the slightly higher incidence rate reduction of 30-39% in 2050, median vaccine characteristics for VE-POI and VE-POD were both 60% (Table 3.2, column 3). Comparison with the medians for 20-29% incidence rate reduction demonstrated that an increase in protection against disease was most likely to provide the additional impact. However, such impact is achievable even with no vaccine efficacy against infection (lower limit of VE-POI = 0%), as long as sufficient vaccine efficacy against disease (60-80%) is achieved. Conversely, if VE-POI were 100%, it could be coupled with a VE-POD of as low as 30% (lower limit of VE-POD = 30%, column 3) to achieve this level of impact. There was no 5-year duration scenario with less than 30% VE-POD that could achieve this level of impact. The maximum reduction possible with a 5-year duration vaccine with 10-yearly mass vaccination would be a 50-59% IRR in 2050 (Table 3.2, column 5), which would require 100% VE-POD, and any level of protection against infection as the outcome is not very sensitive to VE-POI. In Table

3.2, the VE-POI appears to oscillate with increasing required impact, however, this is because the impact is driven by VE-POD, so VE-POI can vary a little around 50-60%, with the exception of the highest impact, which ideally required a slightly higher VE-POI (70%) to be achievable.

With a duration profile of 10 years with 10-yearly mass campaigns, an incidence rate reduction of 20-29% was achieved with a median VE-POD of 10% (range 0-30%) and VE-POI 50% (range: 0-100%) (Table 3.2). The median VE-POD was substantially lower than with 5 years duration of protection. The maximum achievable incidence rate reduction with 5 years protection was 50-59%, whereas with 10 years protection as high as 70-79% could be achieved, though would require a completely disease-blocking vaccine (i.e. VE-POD=100%).

These median characteristics and ranges can inform the target product profile for new TB vaccines. Once the required levels of epidemiological impact (minimum and possibly ideal) have been agreed, the median characteristic values required for those levels of impact would be indicative of the typical combined characteristic values at which the incidence rate reduction could be achieved. The ranges provide the scope of possible values of the characteristics that could achieve the given level of impact, and are important for trial design and endpoint selection, though should be interpreted with caution, as a correlation exists between these characteristics.

In Table 3.4, an example is provided of possible TPP text, including these medians, ranges, and annotation referring to the heat maps and examples demonstrating the correlation between the vaccine characteristics. The example TPP text is provided for 20% to 79% (in 10% steps) incidence rate reduction in 2050 compared to no new vaccine baseline.

Table 3.2: Median and range of vaccine efficacies required to achieve 20-80% incidence rate reduction (in 10% steps) in 2050 compared to no new vaccine baseline for 5 and 10 year durations of protection and 10-yearly mass vaccination campaigns

Vaccine efficacy	Median VE to achieve 20-29% TB IRR 2050 (range)	Median VE to achieve 30-39% TB IRR 2050 (range)	Median VE to achieve 40-49% TB IRR 2050 (range)	Median VE to achieve 50-59% TB IRR 2050 (range)	Median VE to achieve 60-69% TB IRR 2050 (range)	Median VE to achieve 70-79% TB IRR 2050 (range)
5 year duration vaccine (with 10-yearly mass campaigns)						
VE-POD	40% (0-60%)	60% (30-80%)	90% (60-100%)	100% (100-100%)	n/a	n/a
VE-POI	60% (0-100%)	60% (0-100%)	50% (0-100%)	60% (0-100%)	n/a	n/a
10 year duration vaccine (with 10 yearly mass campaigns)						
VE-POD	10% (0-30%)	30% (0-40%)	40% (20-50%)	60% (40-70%)	80% (60-80%)	100% (100-100%)
VE-POI	50% (0-100%)	60% (0-100%)	50% (0-100%)	60% (0-100%)	50% (0-100%)	70% (0-100%)

Table 3.3: Median and range of vaccine efficacies required to achieve 20-50% incidence rate reduction (in 10% steps) in 2050 compared to no new vaccine assuming 10-yearly mass vaccination campaigns and that the two unmeasured characteristics (duration and one vaccine efficacy) are at a minimum

Characteristic	Characteristics to achieve 20-29% IRR in 2050		Characteristics to achieve 30-39% IRR in 2050		Characteristics to achieve 40-49% IRR in 2050		Characteristics to achieve 50-59% IRR in 2050	
	POI as trial outcome	POD as trial outcome	POI as trial outcome	POD as trial outcome	POI as trial outcome	POD as trial outcome	POI as trial outcome	POD as trial outcome
VE-POI	100% (90-100%)	0%*	Not achievable	0%*	Not achievable	0%*	Not achievable	0%*
VE-POD	0%*	50% (40-60%)	-	70% (60-80%)	-	90% (80-100%)	-	100% (100-100%)
Duration	5yrs*	5yrs*	5yrs*	5yrs*	5yrs*	5yrs*	5yrs*	5yrs*

*Characteristic assumed at minimum as not measured in clinical trial. 5 years duration is the lowest capable of achieving at least 20-29% IRR in 2050.

Table 3.4: Suggested TPP values for VE-POI, VE-POD and duration of protection required to achieve 20-29%, 50-59% and 70-79% incidence rate reduction in 2050 with 10-yearly mass campaigns. Other ranges are available in Appendix B.

Vaccine characteristic	To achieve 20-29% TB IRR in 2050		
	Shortest duration	Revaccination duration	Annotation
Duration of protection	5 years	10 years	Consult heat maps for full results. The medians of VE-POI and VE-POD are indicative of the minimum profile considered highly likely to provide the minimum required epidemiological impact (20-29% TB IRR) at the given duration of protection. Ranges should be interpreted with caution, as an interaction exists between these characteristics, therefore the minimums of each range combined would not provide sufficient impact. For example, with 5 years duration and a VE-POD 0%, VE-POI would need to be 90-100%. With VE-POI 0%, a VE-POD of 40-60% would be required.
Median VE-POD (range)	40% (0-60%)	10% (0-30%)	
Median VE-POI (range)	60% (0-100%)	50% (0-100%)	

Vaccine characteristic	To achieve 50-59% TB IRR in 2050		
	Shortest duration	Revaccination duration	Annotation
Duration of protection	5 years	10 years	Consult heat maps in appendix for full results. The medians of VE-POI and VE-POD are indicative of the minimum profile considered highly likely to provide the minimum required epidemiological impact (50-59% TB IRR) at the given duration of protection. Ranges should be interpreted with caution, as an interaction exists between these characteristics, therefore the minimums of each range combined would not provide sufficient impact. For example, with 5 years duration a POI-only vaccine is not feasible, but a POD-only vaccine would be feasible with VE-POD 100%.
Median VE-POD (range)	100% (100-100%)	60% (40-70%)	
Median VE-POI (range)	60% (0-100%)	60% (0-100%)	

Table 3.4 **continued…**

Vaccine characteristic	To achieve 70-79% TB IRR in 2050		
	Shortest duration	**Revaccination duration**	**Annotation**
Duration of protection	5 years	10 years	Consult heat maps in appendix for full results. The medians of VE-POI and VE-POD are indicative of the minimum profile considered highly likely to provide the minimum required epidemiological impact (70-79% TB IRR) at the given duration of protection. Ranges should be interpreted with caution, as an interaction exists between these characteristics, therefore the minimums of each range combined would not provide sufficient impact. This level of impact is not possible with a 5-year duration vaccine. However, with 10 years duration a VE-POD of 100% would be required.
Median VE-POD (range)	n/a	100% (100-100%)	
Median VE-POI (range)	n/a	70% (0-100%)	

3.4.2.3 Mortality rate reduction compared to no new vaccine baseline in 2050

Modelled estimates of mortality rate reduction (MRR) in 2050 compared to the no new vaccine baseline were similar to the incidence rate reduction estimates in terms of rate reduction and trends (Figure 3.12). The MRRs for 2-5 years duration of protection are greater than the equivalent incidence rate reductions, whereas for seven years and above the MRR is marginally lower than the incidence rate reduction. With two years duration of protection and very high vaccine efficacy against disease it is possible to achieve a MRR of 22% (UR: 20-23%), and with 5 years protection up to 55% (UR: 53-58%).

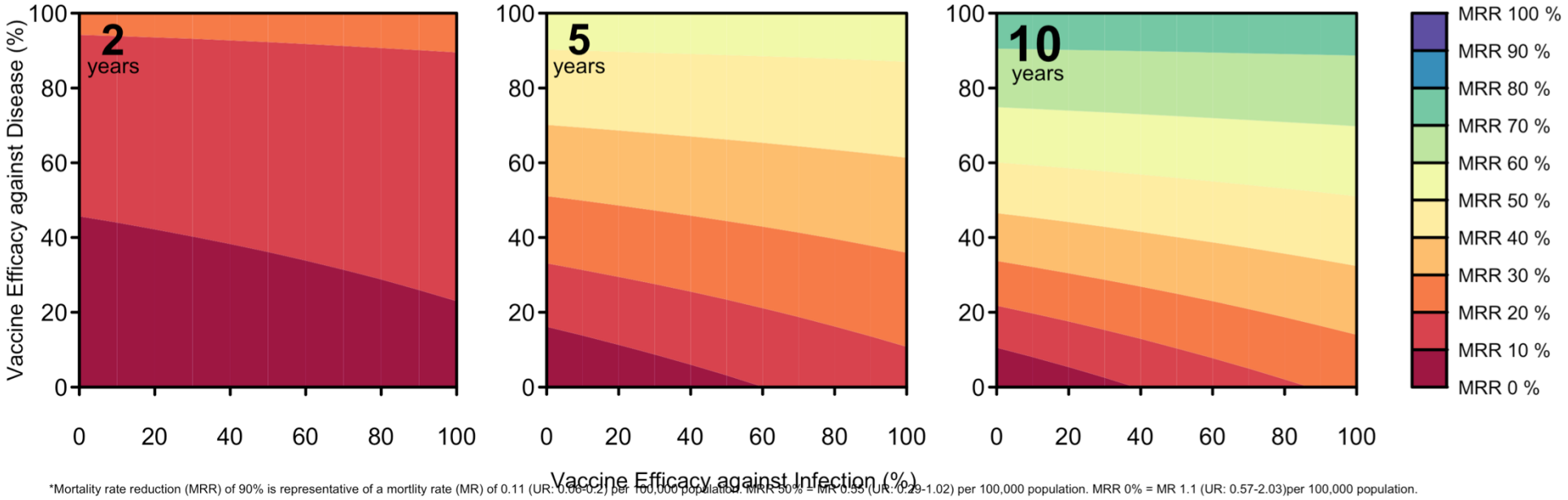


Figure 3.12: Median mortality rate reduction (%) for pre- and post-infection vaccine compared to no new vaccine baseline in 2050, by percentage efficacy against infection (x-axis), percentage efficacy against disease (y-axis), and duration of protection (as indicated in top left corner of panel). Although a variety of durations were explored, 2 years (left panel), 5 years (centre panel) and 10 years (right panel) duration of protection, all with 10-yearly mass vaccination campaigns, are presented here.

3.4.2.4 Cumulative TB cases and TB deaths averted 2025-2050

Heat maps demonstrating the cumulative number of cases and deaths averted 2025-2050 for the spectrum of vaccine efficacies and durations of protection of 2, 5 and 10 years for P&PI vaccines can be found in Figure 3.13 and Figure 3.14. A few key examples are given below.

A P&PI vaccine with 100% vaccine efficacy against both infection and disease could avert 3.6 million (UR: 3.1-3.9m) cases and 82,000 (UR: 45,000-147,000) deaths with two years duration of protection (Figure 3.13 and Figure 3.14, left panel), 7.7 million (UR: 6.7-8.4m) and 173,000 (UR: 94,000-308,000) deaths with five years protection (Figure 3.13 and Figure 3.14, centre panel), and 11.6 million (UR: 10.2-12.6m) cases and 270,000 (UR: 145,000-483,000) deaths by 2050 with 10 years protection(Figure 3.13 and Figure 3.14, right panel). At the lower end of the efficacy spectrum, a vaccine with 20% vaccine efficacy against both infection and disease could avert 1.0 million (UR: 0.8-1.1m) cases and 21,000 (UR: 11,000-38,000) deaths with two years duration of protection, up to 3.4 million (UR: 3.0-3.9m) cases and 74,000 (UR: 40,000-135,000) deaths with 10 years duration of protection.

A vaccine efficacy of 60% was considered an achievable target with new vaccines, therefore if vaccine efficacy against both infection and disease were both 60%, two to ten years' protection would avoid 2.5 million (UR: 2.2-2.7m) to 8.2 million (UR: 7.2-9.1m) cases and 56,000 (UR: 30,000-100,000) to 185,000 (UR: 101,000-333,000) deaths during 2025-2050. For a 10 year protection vaccine, if VE-POI were then increased to 80%, an additional 0.2 million cases would be averted; whereas if instead VE-POD were increased to 80%, an additional 1.7 million cases would be averted.

Additional years of protection would also avoid additional cases and deaths. For example, for a 100% VE-POD vaccine the average annual additional number of cases averted when increasing duration from two to three years is 1.5 million. With increasing duration, the absolute number averted continues to increase, but the

additional gains per year added are smaller, for example, increasing from seven to 10 years protection gains on average 0.7 million additional averted cases per year.

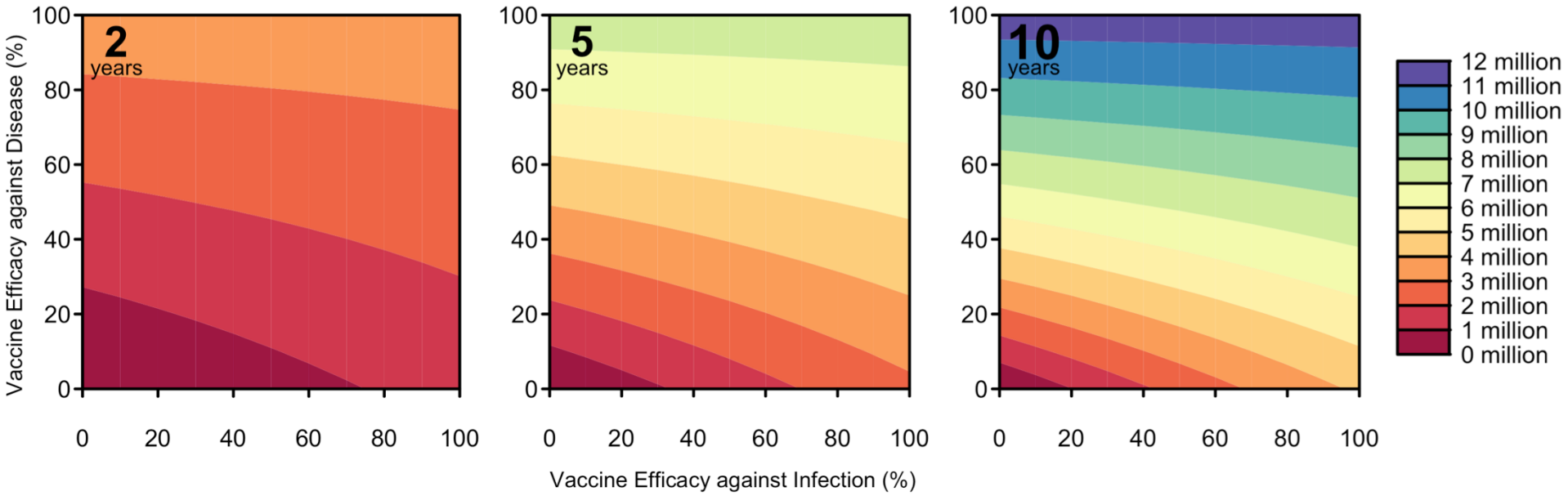


Figure 3.13: Median cumulative number of cases averted for the period 2025-2050 for pre- and post-infection vaccine compared to no new vaccine baseline, by percentage efficacy against infection (x-axis), percentage efficacy against disease (y-axis), and duration of protection (as indicated in top left corner of panel). Although a variety of durations were explored, 2 years (left panel), 5 years (centre panel) and 10 years (right panel) duration of protection, all with 10-yearly mass vaccination campaigns, are presented here.

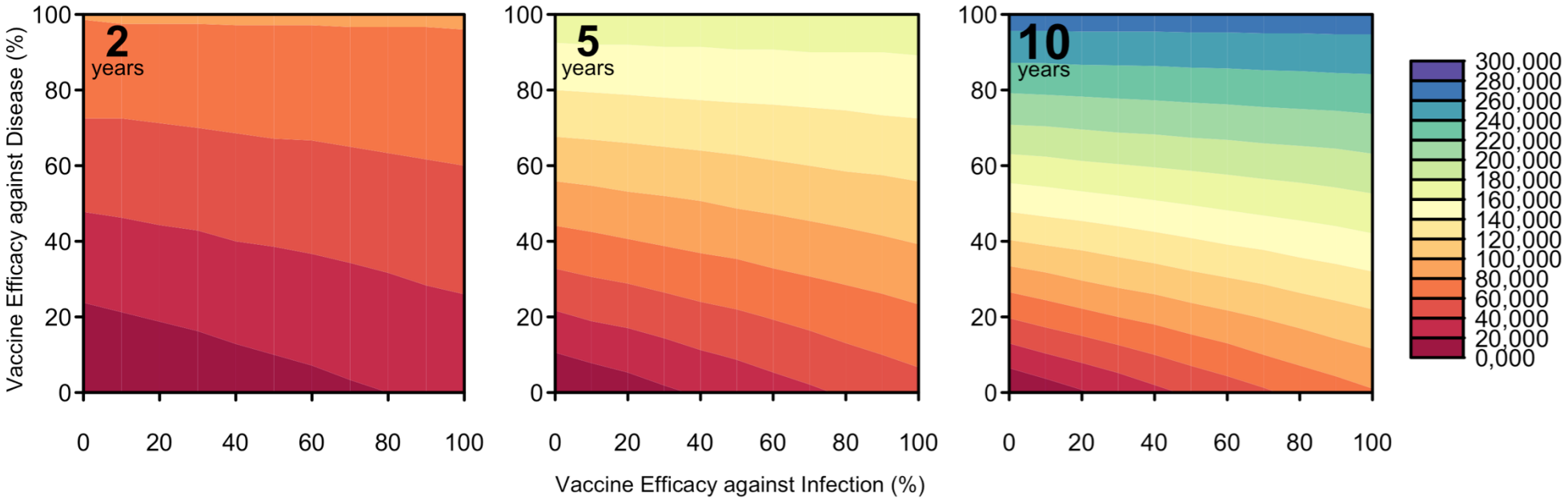


Figure 3.14: Median cumulative number of deaths averted for the period 2025-2050 for pre- and post-infection vaccine compared to no new vaccine baseline, by percentage efficacy against infection (x-axis), percentage efficacy against disease (y-axis), and duration of protection (as indicated in top left corner of panel). Although a variety of durations were explored, 2 years (left panel), 5 years (centre panel) and 10 years (right panel) duration of protection, all with 10-yearly mass vaccination campaigns, are presented here.

3.4.2.5 Incidence rate reduction and cases averted by 2035

Incidence rate reductions were estimated in 2035 (Figure 3.15), in addition to the estimated cases averted for 2025-2035 (Figure 3.16). Overall, similar trends were observed in terms of impact achieved by prevention of infection and disease efficacy characteristics in both 2035 and 2050. However, a difference was seen in the level of reduction achieved and patterns by duration of protection between these two time points. When comparing the percentage incidence rate reduction to the no new vaccine baseline in 2035, a marginally lower incidence rate reduction was estimated for durations of five years and above, whereas a higher percentage reduction in incidence rate was estimated in 2035 than 2050 for the two and three year durations of protection. This is due to the timing of vaccine waning relative to the time point at which impact is measured. In 2035, all vaccines had recently been boosted and therefore populations were protected in all duration scenarios; whereas in 2050, the shorter vaccine durations had already waned following the most recent (2045) mass campaign.

The numbers of cases averted by vaccination 2025-2035 are shown in Figure 3.16. In 2035, a 100% vaccine efficacy two-year duration vaccine could avert up to 1.7 million (UR: 1.5-1.8m) cases (left panel), a 5-year duration vaccine up to 3.5 million (UR: 3.1-3.7m) cases (centre panel), and a 10-year duration vaccine up to 5.6 million (UR: 5.0-6.0m) cases (right panel).

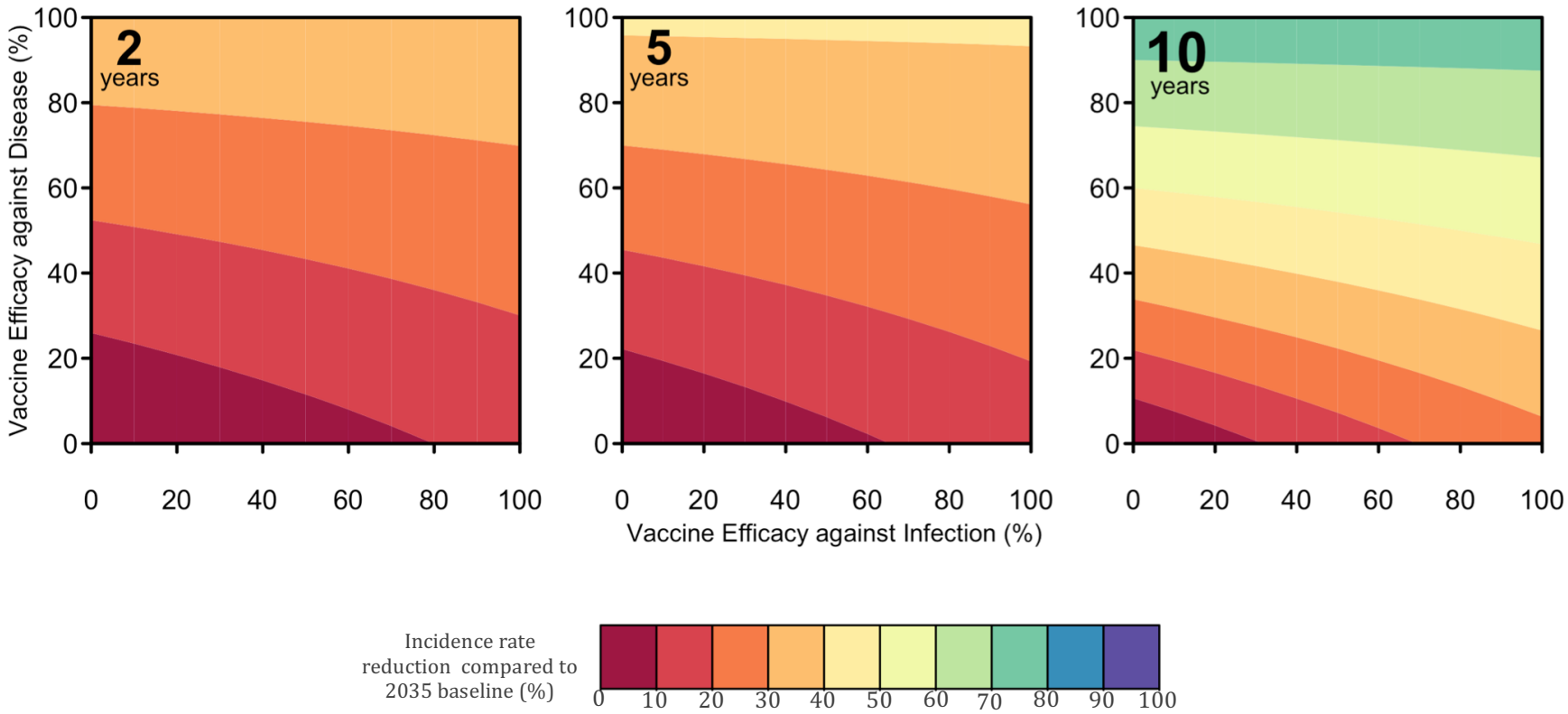


Figure 3.15: Median incidence rate reduction (%) for pre- and post-infection vaccine compared to no new vaccine baseline in 2035, by percentage efficacy against infection (x-axis), percentage efficacy against disease (y-axis), and duration of protection (as indicated in top left corner of panel). Although a variety of durations were explored, 2 years (left panel), 5 years (centre panel) and 10 years (right panel) duration of protection, all with 10-yearly mass vaccination campaigns, are presented here.

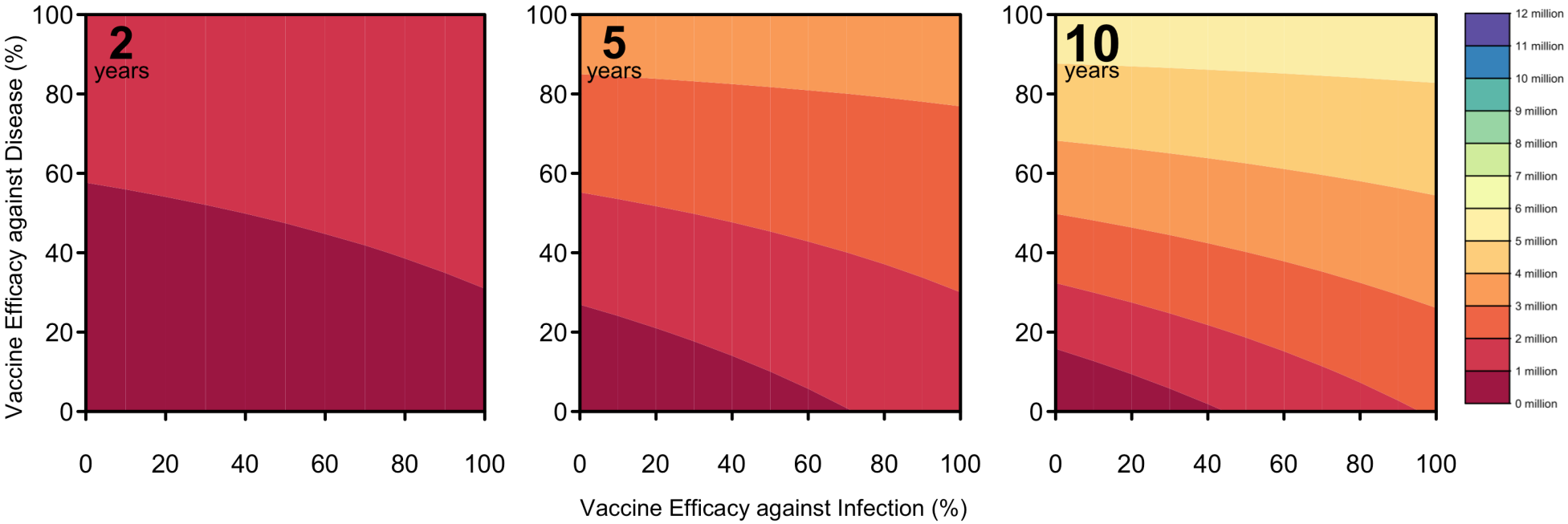


Figure 3.16: Median cumulative number of cases averted for the period 2025-2035 for pre- and post-infection vaccine compared to no new vaccine baseline, by percentage efficacy against infection (x-axis), percentage efficacy against disease (y-axis), and duration of protection (as indicated in top left corner of panel). Although a variety of durations were explored, 2 years (left panel), 5 years (centre panel) and 10 years (right panel) duration of protection, all with 10-yearly mass vaccination campaigns, are presented here.

3.4.2.6 Achievable impact with POI-only or POD-only vaccines

Incidence rate reductions achieved by POD-only (VE-POI=0%) (Figure 3.17, left panel) and POI-only (VE-POD=0%) (Figure 3.17, right panel), highlight again that for equivalent profiles the achievable impact is greater with a POD vaccine than POI.

Assuming 100% efficacy against disease (VE-POD=100%, VE-POI=0%), the 10 year duration of protection vaccine achieved 78.4% (UR: 76.6-80.4%) incidence rate reduction, averting 11.6 (UR: 10.8-12.4) million cases between 2025-2050. Whereas with 100% efficacy against infection (VE-POD=0%, VE-POI=100%) only 27.6% (UR: 23.0-34.7%) incidence rate reduction would be expected, averting only 3.9 million cases (UR: 3.6-4.1m). The number of cases averted by a 100% VE-POI vaccine were similar to a POD-only vaccine with a vaccine efficacy of approximately 20-30% (Figure 3.18).

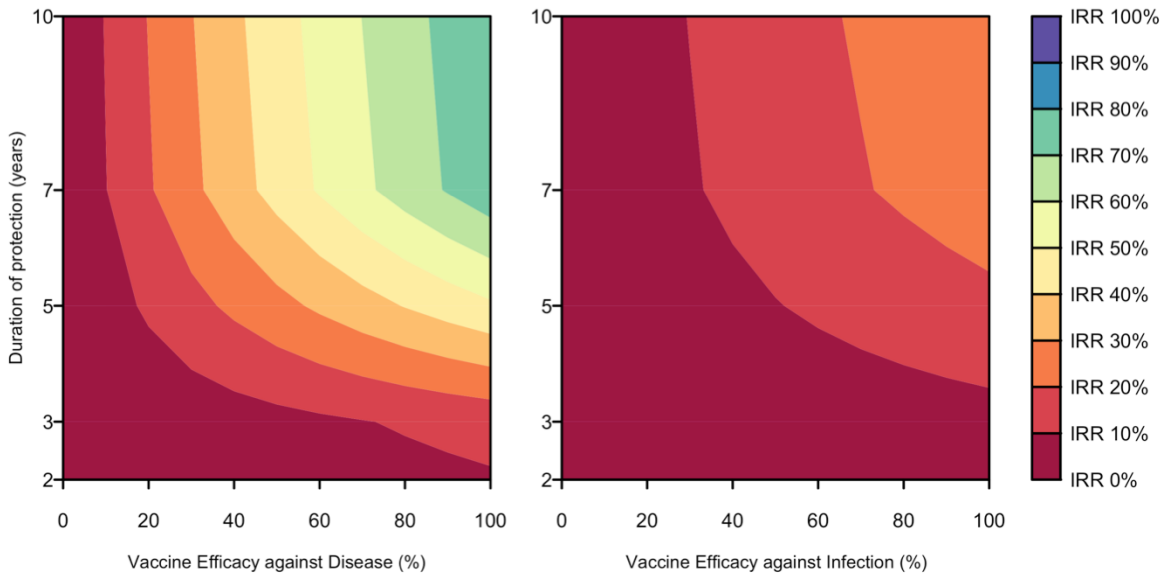


Figure 3.17: Median incidence rate reduction for pre- and post-infection vaccine compared to no new vaccine baseline in 2050, by percentage efficacy against disease or infection (x-axes), and duration of protection (y-axis). In the left hand panel, efficacy against infection was held at zero (i.e. POD-only), and in the right hand panel efficacy against disease was held at zero (i.e. POI-only).

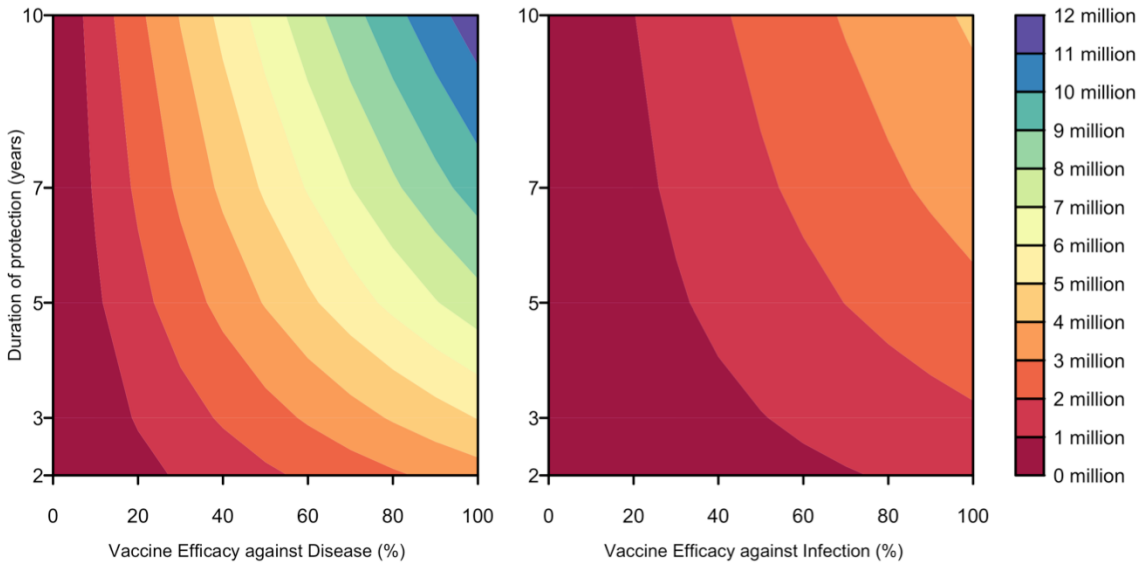


Figure 3.18: Median cumulative number of cases averted for the period 2025-2050 for pre- and post-infection vaccine compared to no new vaccine baseline, by percentage efficacy against disease or infection (x-axes), and duration of protection (y-axis). In the left hand panel, efficacy against infection was held at zero, and in the right hand panel efficacy against disease was held at zero.

Taking 60% vaccine efficacy as a more realistic efficacy scenario (Table 3.5), a POI-only vaccine with two to 10 years' protection was estimated to reduce 2050 incidence rates by 2.9% to 18.6% and avoid 0.8 to 2.5 million cases. The equivalent POD-only vaccine was estimated to provide a 2050 incidence rate reduction of 5.6% to 53.3% and avoid 2.2 to 7.5 million cases, up to three-times as many cases as the POI-only vaccine. A reduction in incidence rate of at least 20-29% could not be achieved with a 60% POI-only vaccine of any duration, whereas the POD-only vaccine would more than achieve this with a vaccine providing 5 years protection.

Table 3.5: Epidemiological impact of POI-only and POD-only vaccines with 60% vaccine efficacy pre- and post-infection

VE-POI (%)	VE-POD (%)	Duration of protection (years)	Percentage incidence rate reduction in 2050 compared to no new vaccine (uncertainty range)	Millions of cases averted 2025-2050 (uncertainty range)
60	0	2	2.9 (2.4-3.8)	0.8 (0.7-0.8)
60	0	3	4.3 (3.5-5.6)	1.1 (1.0-1.1)
60	0	5	11.3 (9.3-14.6)	1.6 (1.5-1.7)
60	0	7	17.0 (14.0-21.6)	2.2 (1.9-2.6)
60	0	10	18.6 (15.4-23.7)	2.5 (2.4-2.7)
0	60	2	5.6 (4.1-7.2)	2.2 (1.9-2.4)
0	60	3	8.3 (6.1-10.6)	3.1 (2.8-3.4)
0	60	5	31.7 (29.1-34.4)	4.8 (4.4-5.1)
0	60	7	51.0 (48.8-53.2)	6.0 (5.6-6.4)
0	60	10	53.3 (50.6-56.1)	7.5 (7.0-8.1)

3.4.2.7 Minimum vaccine efficacy for trial design

Assuming the 'worst case' scenario for duration (5 years) and vaccine efficacy against disease (0%) and with vaccine efficacy against infection as a trial outcome, the benchmark for stage-gating based upon VE-POI was estimated to be 100% (90-100%) if a 20-29% reduction in incidence rate in 2050 was required. Conversely, with vaccine efficacy against disease (VE-POD) as the trial outcome, only 50% (40-60%) vaccine efficacy would be required to be confident in achieving this level of epidemiological impact. With disease endpoints as the trial outcome, 70% (60-80%) VE-POD was estimated to provide confidence in achieving 30-39% incidence rate reduction in 2050, and up to 50-59% reduction could be achieved with a vaccine blocking progression to disease (VE-POD=100%). With vaccine efficacy against infection as the outcome, it would not be possible to be confident in achieving any higher IRR than 20-29%, unless the vaccine provided more than 5 years duration of protection or VE-POD was known.

3.4.2.8 Pre-infection and post-infection vaccines

The highest achievable incidence rate reduction with 100% efficacious vaccines against both infection and disease (POI&D) with 10 years duration of protection compared to no new vaccine baseline in 2050 was as little as 27% (UR: 23-34%) with a pre-infection vaccine, whereas with a post-infection vaccine achieved 71% (UR: 70-71%) (Figure 3.19). The equivalent P&PI vaccine reduced incidence rates by 79% (UR: 77-81%). These were associated with 3.9 million (UR: 3.5-4.7m), 9.3 million (UR: 8.0-10.4m) and 11.6 million (UR: 10.2-12.6m) cumulative cases averted between 2025 and 2050 (See Appendix B for graphs).

The sensitivity of the incidence rate reduction outcome to vaccine efficacy against disease and infection differ greatly between the three vaccine types. The post-infection (PSI) vaccine contours are almost horizontal, indicating that the vaccine impact comes almost exclusively from the vaccine efficacy for prevention of disease (Figure 3.19, right panel). The incidence rate reduction outcome is much more sensitive to both types of vaccine efficacy for a PRI vaccine (centre panel), with a slightly greater sensitivity to vaccine efficacy for prevention of infection. Due to the large contribution of reactivation disease in already infected populations to the overall burden, the P&PI vaccine heat map (left panel) most closely resembles the PSI results. The incidence rate reduction heat maps for PRI and PSI vaccines for the full range of durations explored are provided in Appendix B.

Uncertainty ranges are provided for results throughout the text. In addition, uncertainty ranges for P&PI, PRI and PSI vaccines for all combinations of 20% and 100% for both vaccine efficacies and durations of 5 and 10 years are presented in Figure 3.20.

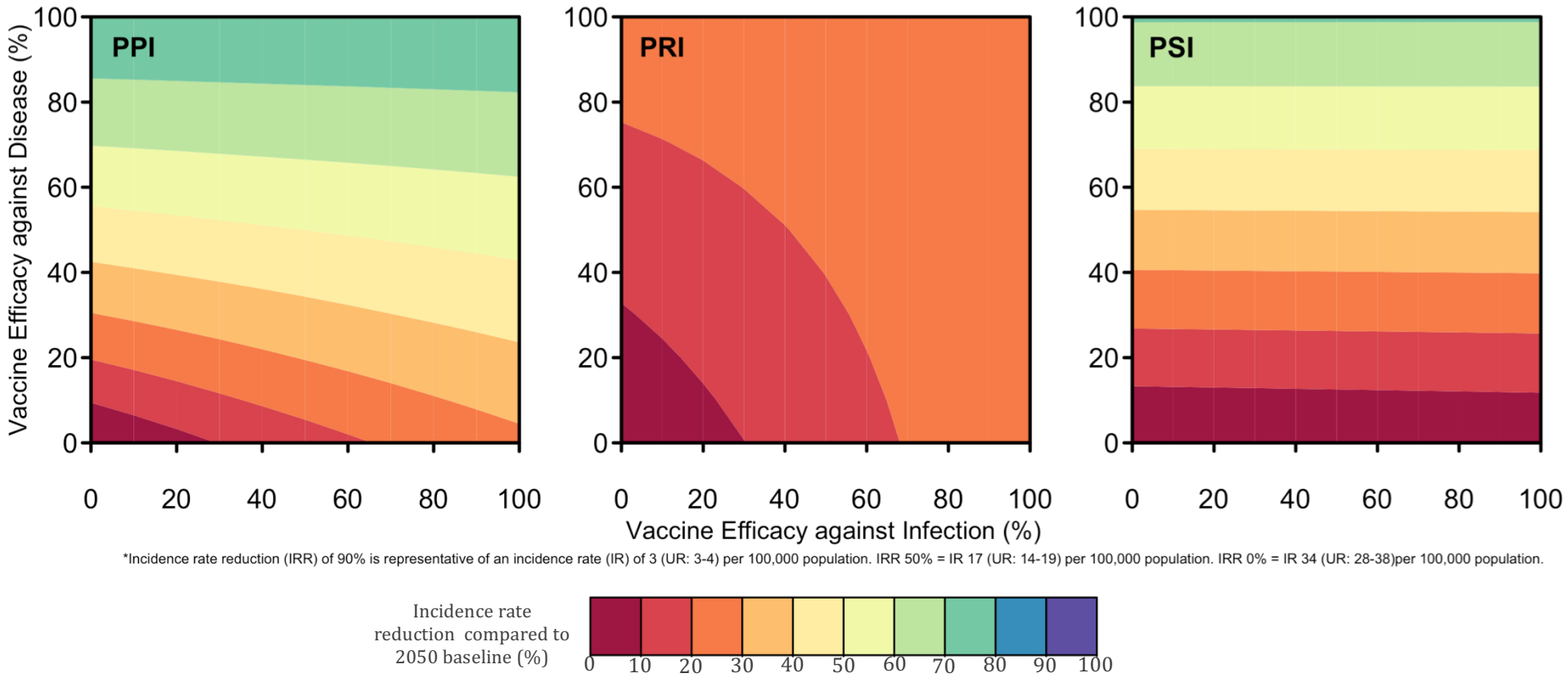


Figure 3.19: Median incidence rate reduction (%) for vaccines compared to no new vaccine baseline in 2050, by percentage efficacy against infection (x-axis), percentage efficacy against disease (y-axis), and infection status of protected population (as indicated in top left corner of panel). Pre- and Post-infection (left panel), pre-infection (centre panel) and post-infection (right panel) vaccines, all with 10 years of protection and 10-yearly mass campaigns, are presented here.

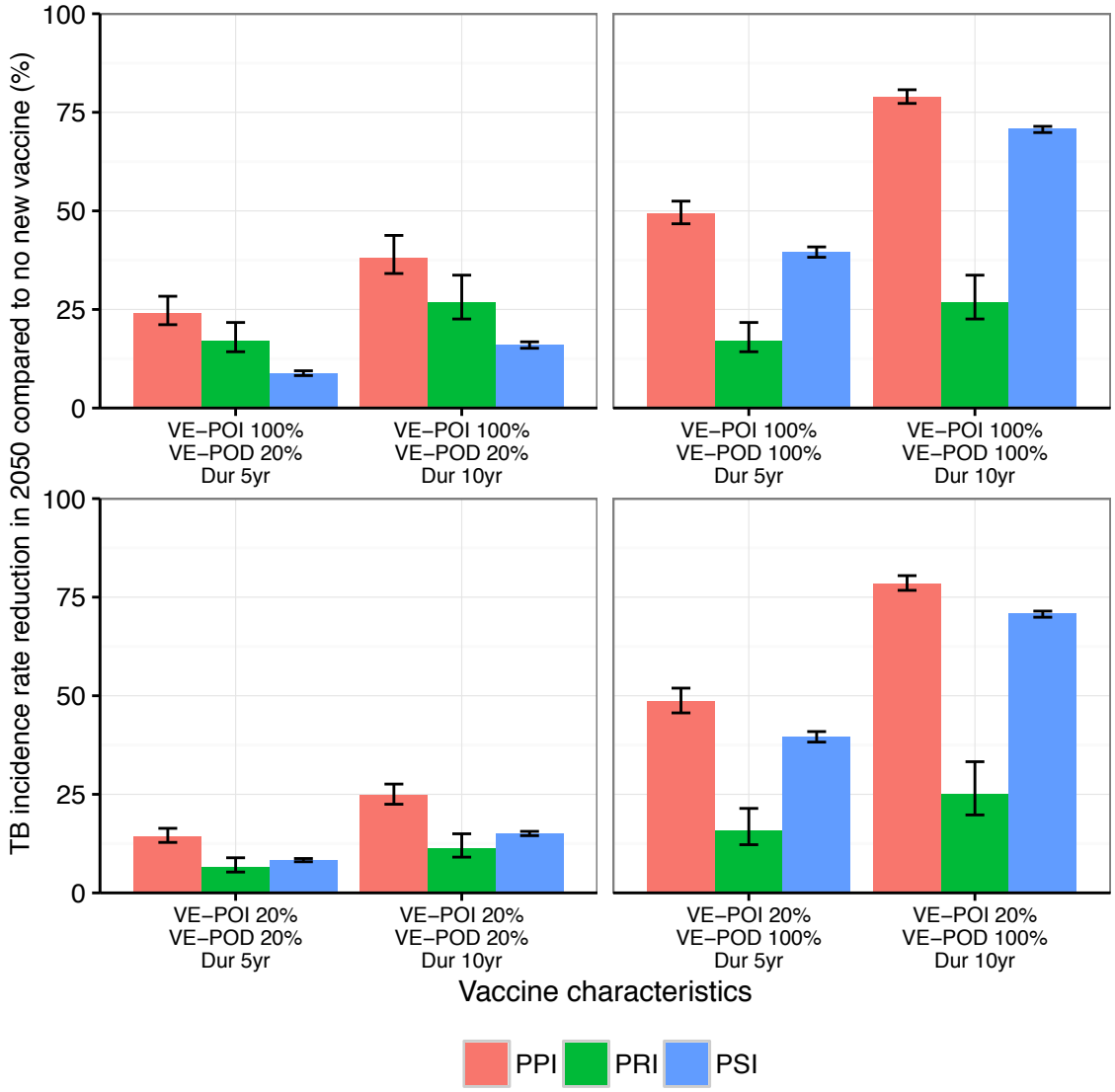


Figure 3.20: Medians and uncertainty ranges for estimates of incidence rate reduction in 2050 for combinations of durations of efficacy of 5 years and 10 years, vaccine efficacies for prevention of infection and prevention of disease of 20% and 100%, for P&PI (red), PRI (green) and PSI (blue) vaccines.

3.4.2.9 Incidence rate reduction in 2050 with 5-yearly mass campaigns

Increasing the frequency of mass campaigns to every 5 years improved the achievable impact in all duration scenarios below 10 years duration of protection (Figure 3.21). For the two-year duration of protection vaccines, the IRR more than doubled (104-120% increase) when mass vaccination was conducted 5-yearly as opposed to 10-yearly. For 3 to 7 year duration of protection vaccines, the incidence rate reduction achieved increased by approximately a quarter to nearly two thirds.

With 5-yearly boosting, even the shortest duration of protection explored (2 years) was able to achieve more than the absolute minimum of 20-29% incidence rate reduction. For 20-29% incidence rate reduction, at least 10% VE-POD is required (Table 3.6, bottom of VE-POD range), so a prevention of infection-only vaccine could not deliver this level of impact. The duration at which the achievable impact plateaus was 5 years, for which the results for VE-POI and VE-POD are almost identical to the 10-year duration vaccine with 10-yearly boosting (Table 3.2 and Table 3.6).

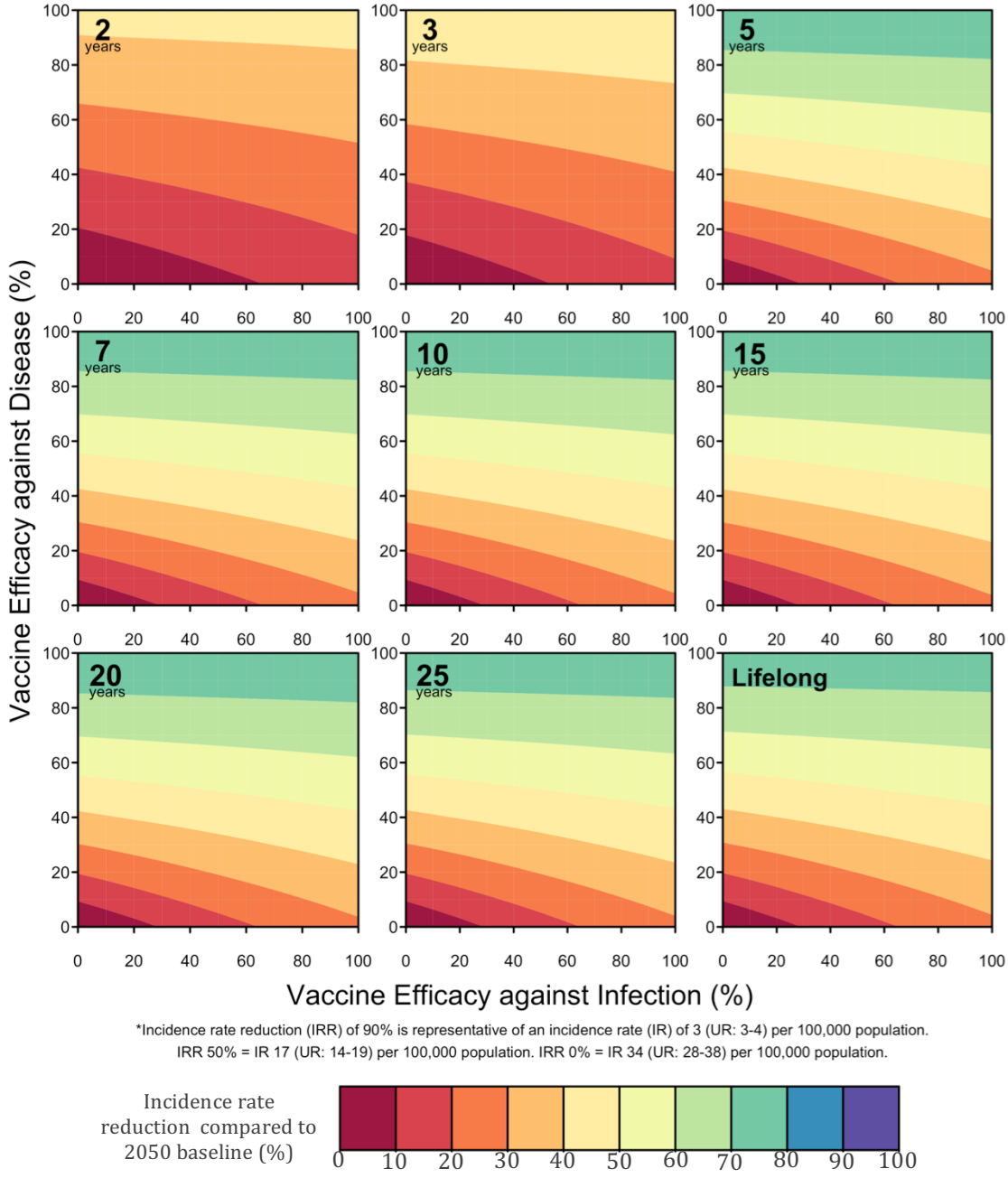


Figure 3.21: Median incidence rate reduction (IRR, %) for pre- and post-infection vaccine with 5-yearly mass vaccination campaigns compared to no new vaccine baseline in 2050, by percentage efficacy against infection (x-axis), percentage efficacy against disease (y-axis), and duration of protection (as indicated in top left corner of panel).

Table 3.6: Median and range of vaccine efficacies required to achieve 20-80% incidence rate reduction (in 10% steps) in 2050 compared to no new vaccine baseline for 2 and 5 year durations of protection and 5-yearly mass vaccination campaigns

Vaccine efficacy	Median VE to achieve 20-29% TB IRR 2050 (range)	Median VE to achieve 30-39% TB IRR 2050 (range)	Median VE to achieve 40-49% TB IRR 2050 (range)	Median VE to achieve 50-59% TB IRR 2050 (range)	Median VE to achieve 60-69% TB IRR 2050 (range)	Median VE to achieve 70-79% TB IRR 2050 (range)
2 year duration vaccine (with 5 year mass campaigns)						
VE-POD	50% (10-70%)	70% (50-90%)	100% (80-100%)	Not achievable	Not achievable	Not achievable
VE-POI	60% (0-100%)	50% (0-100%)	60% (0-100%)	Not achievable	Not achievable	Not achievable
5 year duration vaccine (with 5 year mass campaigns)						
VE-POD	10% (0-30%)	30% (0-40%)	40% (20-50%)	60% (40-70%)	90% (80-100%)	100% (100-100%)
VE-POI	50% (0-100%)	60% (0-100%)	50% (0-100%)	60% (0-100%)	50% (0-100%)	70% (0-100%)

3.5 Discussion

3.5.1 Discussion

A strategic, data-driven approach to the development of new TB vaccines will be essential to ensure evidence-based maximisation of the future impact of TB vaccines, by developing vaccines with the most appropriate characteristics and for the populations most in need. In this research, I explored vaccine intrinsic characteristics in the target product profile (duration of protection, vaccine efficacy against disease, and vaccine efficacy against infection) to estimate the potential population-level impact of new TB vaccines in China and to inform vaccine development directed towards maximisation of impact in this and similar settings.

In China, vaccine efficacy for prevention of disease was key to maximising impact on disease burden over the 2025-2050 time horizon, and provided most impact when effective post-infection or pre- and post-infection. For a vaccine effective pre- and post-infection with 10-yearly mass campaigns, at least 5 years duration of protection were required to achieve the absolute minimum incidence rate reduction of 20-29% in 2050, and impact was maximised when mass campaign frequency equalled duration of protection.

Required vaccine efficacies depend upon the minimum incidence rate reduction considered acceptable, but medians ranged from 40% to 100% for VE-POD and were mostly 50% to 60% for VE-POI, with wide ranges around both due to the different combinations capable of achieving such impact, given these two characteristics were correlated. Impact was highly sensitive to VE-POD and relatively insensitive to VE-POI due to the reactivation-driven nature of the epidemic in China.

The finding that duration of protection of 5 years would be required with 10-yearly mass campaigns to ensure at least 20-29% incidence rate reduction has important implications for clinical trials. Most efficacy studies plan for 2-3 years follow up, whereas these results suggest that follow up to demonstrate efficacy of at least 5

years will be required. Such a long phase IIB or phase III study could be prohibitively expensive, therefore an alternative could be to plan for a 2-3 year efficacy study with clinical or immunological follow up in a subset of patients continued up to 5 years. Flexibility would be required in the trial end point, with selection of immunological versus clinical follow up contingent upon identification of a correlate of protection.

Alternatively, an implementation approach could be taken if longer trials are not feasible or a new vaccine is found to provide limited duration of protection. By increasing the frequency of mass vaccination, the impact of vaccines with shorter durations of protection could be improved. Assuming mass campaigns every 5 years indicated that a duration of protection as low as 2 years could achieve up to 44% (UR: 42-46) incidence rate reduction in 2050, assuming 100% efficacy against infection and disease. The same impact was achieved with 5 years duration of protection and 5-yearly mass vaccination as with 10 years duration with 10-yearly mass vaccination, therefore shortfalls in duration of protection could be compensated for with increased frequency of mass campaigns, as long as coverage is maintained. It is clear from these results that planning for longer follow up in clinical trials or for shorter intervals between mass vaccination campaigns will likely be necessary to maximise vaccine impact, and should be planned for prospectively.

Impact was maximised when mass campaign frequency equalled duration of protection. As expected, durations of protection longer than the frequency of mass vaccination (10 years) did not provide much additional epidemiological gain. In these scenarios, revaccination occurred at the duration of protection, thus the proportion of the population protected was similar. From an epidemiological perspective, investment in longer duration vaccines would not necessarily be needed; however, cost-effectiveness analyses would be required, as less frequent mass campaigns could lead to large cost savings.

During the 2025-2050 vaccination period, most new TB cases were attributable to reactivation disease, many of which were likely infected before introduction of

vaccine. Hence, the incidence rate reduction achieved by P&PI vaccines was most sensitive to vaccine efficacy for prevention of disease, with limited gains from increases in efficacy against infection. Although there was value in efficacy for prevention of infection, in this setting greater population-level gains were achieved by improving vaccine efficacy for preventing development of disease. Therefore, development strategies for China and similar epidemics should aim to explore disease outcomes in trials, and accelerate development of candidate vaccines with at least some anticipated protection against disease.

The greatest achievable impact with the vaccines explored reduced incidence rates in 2050 by 79% (77-81%) compared to the no new vaccine baseline, reaching an incidence rate of 7 (UR: 6-8) per 100,000 population and averting 11.6 million (UR: 10.2-12.6m) cases and 270,000 (UR: 145,000-483,000) deaths by 2050. This is a sizeable burden of disease averted, and would be a significant contribution towards the WHO elimination goal (incidence of 1 case per million population by 2050),[37] as part of a package of TB control interventions. In contrast, two and three year durations of protection with 10-yearly mass campaigns would deliver less than 20% incidence rate reduction, and if coupled with low vaccine efficacies (20%), two-year duration vaccines would avert a more modest 1.0 million (UR: 0.8-1.1m) cases and 21,000 (UR: 11,000-38,000) deaths over the 26 year period. Maximising the incidence rate reduction and cases averted should therefore be the strategic goal, but in a challenging development environment, identifying the minimum acceptable incidence rate reduction can inform vaccine characteristics in the TPP minimum profile, and hopefully be a step towards more efficacious vaccines in the future.

Once the minimum acceptable incidence rate reduction in 2050 has been chosen by stakeholders, the appropriate section of Table 3.4 could be used within a target product profile or preferred product characteristic table. The durations to achieve the absolute minimum impact and the duration above which minimal additional epidemiological gains could be achieved with 10-yearly mass campaigns were identified as five years and 10 years protection, respectively, and are therefore the two durations reported. For these two durations and for each level of incidence rate

reduction, the median and range of vaccine efficacy estimates for prevention of infection and for prevention of disease were estimated. The median values provide the set of characteristics that together would likely provide the required level of protection (e.g. A P&PI vaccine with 40% VE-POD, 60% VE-POI *and* 5 years duration of protection was estimated to be likely to provide 20-29% IRR).

Estimated ranges around the medians reflect the characteristics potentially capable of delivering the required impact. These are wide and must be interpreted with caution because the vaccine efficacy for prevention of infection and disease to provide a given level of impact are correlated (Table 3.4). To explore an example, if an IRR of 20-29% was required and a clinical trial conducted with a vaccine of five years duration of protection with only infection or disease as an endpoint, either type of vaccine could potentially deliver this level of impact and a single-endpoint study could be appropriate. Though it should be noted that a POI-only vaccine would need to have a vaccine efficacy of 90%-100%, whereas a POD-only vaccine could achieve this impact with 40-60% vaccine efficacy. If conducting a trial only measuring VE-POI, there is a risk that a potentially suitable candidate could be discarded if VE-POD remained unknown and VE-POI were to fall below this high stage-gate threshold. Therefore, in China trials should either focus on VE-POD endpoints, or include both VE-POI and VE-POD endpoints. If a longer duration of protection were expected with a given vaccine, or both endpoints were measured and protection in both were anticipated, lower vaccine efficacies could provide sufficient impact, and therefore studies should be powered to measure these lower efficacies.

Target Product Profiles may be used for informing both clinical trial design and pipeline stage gating. Within the bounds of feasibility, clinical trials should be powered to detect the lowest vaccine efficacy that would provide sufficient population level impact. Therefore, TPP vaccine efficacies can inform trial endpoints and sample size. If a POI- or POD-alone vaccine would be insufficient to reach the target impact, both endpoints should be included in clinical trials.

In terms of stage gating after a study, a similar approach can be taken. Once clinical trial data are available to provide information on the characteristics of a new vaccine, the detailed heat maps (e.g. Figure 3.11) can allow estimation of the likely impact of a given vaccine. The estimated impact can be benchmarked against a minimum acceptable impact agreed prospectively by relevant stakeholders. If only one trial endpoint is available (i.e. only one of infection or disease), it may be preferable to stage gate assuming the stringent 'worst case' of zero vaccine efficacy against the unmeasured endpoint to give confidence to developers that at least the minimum incidence rate reduction could be achieved with that vaccine, potentially higher if there were efficacy against the unmeasured outcome.

Although the aim of this research was not to rank characteristics against one another, the incremental benefit from a percentage increase in VE-POD is much greater than the same improvement in VE-POI. Therefore, in this epidemiological setting, prioritisation of P&PI candidates with potential for prevention of disease activity would be rational.

A similar number of cases were averted with 100% vaccine efficacy against both infection and disease with two years duration of protection, as was estimated with 20% vaccine efficacy against each and 10 years duration of protection. Therefore, it is possible to compensate for lower vaccine efficacies with higher durations of protection and vice versa, but would require potentially large compensatory changes. The additional gains from an additional year of protection from a vaccine tend to be large with higher efficacy vaccines, so exploring more in-depth what the relative benefits of longer durations compared to increasing vaccine efficacies may also be informative if such a technical trade-off were to arise.

Although it is hoped that a vaccine effective both pre- and post-infection (P&PI) will be developed to maximise impact, some pipeline candidates may only be effective in one of these populations. It would make only a small difference to potential incidence rate reduction by 2050 in China if a 10 year duration post-infection (PSI) vaccine were

developed as opposed to a P&PI vaccine (IRR 71% (UR: 70-71%) versus 79% (UR: 77-81%), respectively). However, a pre-infection (PRI) vaccine would provide much less impact, with the highest achievable incidence rate reduction in 2050 as little as 27% (UR: 23-34%). Although the P&PI vaccine was a combination of the mechanisms of protection from PSI and PRI vaccines, it is clear that most impact from the P&PI vaccine came from protection against disease in those already infected, which was consistent with baseline epidemiological model outcomes demonstrating that the majority of cases in China were reactivation disease. A post-infection vaccine should be developed following a similar prevention of disease-focussed strategy as P&PI vaccines. The pre-infection vaccine was almost equally sensitive to efficacy against infection or disease, therefore clinical trials could benefit from including both endpoints to gauge the full potential of such vaccines.

Incidence rate reductions in 2050 were the primary outcome of interest in this study. However, several secondary outcomes helped provide a broader picture of the impact of the vaccine profiles. Estimated mortality rate reductions were similar to the incidence rate reductions, with the exception of durations of five years and below, in which mortality rate reductions were larger. This is a function of timing of the last mass campaign (2045) before outcome measurement in 2050. Shorter vaccines ($\leq$5 years) wane before the end of 2050, causing an incidence rate rebound; whereas the mortality rebound occurs after a lag, given that deaths occur some time after the development of disease. For seven years protection and above, a slightly higher incidence rate reduction than mortality rate reduction is observed, likely due to the fact that an individual can be protected from recurring incident cases, but death can only be avoided once. A similar pattern was observed when comparing incidence rate reductions in 2035 to 2050. Mass vaccination occurred in 2035, so unlike 2050 none of the vaccines had waned in 2035 (not shown).

Given the incidence rate reduction is sensitive to the relative timing of the last vaccination and the outcome measurement, the number of cases averted by the different profiles may also be of interest for decision making. The number of cases

and deaths averted for all of the profiles are demonstrated in heat maps in Figure 3.13 and Figure 3.14 and range from around 1.0 million (UR: 0.8-1.1m) cases and 21,000 (UR: 11,000-38,000) deaths up to 11.6 million (UR: 10.2-12.6m) cases and 270,000 (UR: 145,000-483,000) deaths during the 2025-2050 vaccination period.

Discussion with stakeholders suggested that shorter five-year spacing between mass campaigns could be feasible in some settings.[28] Reducing the frequency of mass campaigns from 10-yearly to 5-yearly more than doubled the incidence rate reduction achieved by 2 year protection vaccines, and also improved the incidence rate reductions achieved by the 3-7 year protection vaccines. Conducting mass campaigns this frequently over the long term could potentially prove incredibly costly, and a challenge to sustain momentum in the campaigns, especially as burden of disease declines. However, annual mass campaigns of 1-29 year olds with the Menafrivac vaccine achieved very high coverages year on year.[33] Therefore, if a short duration vaccine were developed and cost-effectiveness analyses were to suggest that 5-yearly mass campaigns were cost effective, there may be value in considering more frequent mass campaigns.

3.5.2 Research in context of existing literature

As discussed in Chapter 2, this research need was identified as no existing models had comprehensively explored the influence of vaccine characteristics on epidemiological impact. Such research is needed to help maximise vaccine impact by informing Target Product Profile 'ideal' characteristics, and ensuring at least a minimum required level of impact by helping inform 'minimum' characteristics. In addition, only two mathematical models had explored vaccine impact with very limited vaccine profiles in China, yet this is a high priority country given the high burden of disease and the existence of a candidate in phase III trials.

The results of this research support the existing literature,[8] suggesting that efficacy for prevention of disease is likely to provide greatest impact and more rapidly than prevention of infection vaccines. The literature to-date was divided as to whether

vaccines effective pre-infection or post-infection would provide greatest impact.[8] Results of the China sub-model in Dye (2013) indicated that post-infection vaccines delivered to latent individuals would provide greatest protection, but as discussed in Chapter 2, were interpreted with caution given few model or vaccine details were available. In the China model presented in this chapter, results supported the conclusions of the Dye study, as vaccines effective post-infection very clearly provided greatest epidemiological impact. This was due to the large and increasing proportion of cases arising from reactivation of latent infection.

Research preceding this work by Knight et al. explored the impact of adolescent/adult vaccination.[1] Results were reported at the global and income level, but was calibrated to overall TB incidence and mortality by country in 2009 and the population size in 2009 and 2050. The mass vaccination coverage assumption for China (73%) was similar to the assumption in our model, but routine coverage was assumed much higher, at 99% routine coverage of neonates or 10 year olds. The upper-middle income country (UMIC) group, which included China, experienced a 42.4% (UR: 30.8-54.9%) reduction in number of cases 2024-50 with introduction of a 60% efficacious P&PI, all-or-nothing prevention of disease vaccine delivered to adolescents/adults. In my study, I report percentage reduction in incidence rates, as opposed to number of cases, but results for the most similar vaccine are within the same region of impact (53.3%; UR: 50.6-56.1). In the Knight study, approximately 8 million cases were averted across all UMICs with this vaccine, and in this research an estimated 7.5 million were averted in China alone.

Two new studies have been published since this research was conducted. The first explored the impact of new TB vaccines in China during 2018-2035,[38] and the second was a global model exploring the importance of age-specificities in TB vaccine modelling.[39]

The Liu study compared neonatal vaccination to neonatal plus periodic mass all-age vaccination ('mixed vaccination') in China.[38] The mass revaccination frequencies

explored were mostly shorter (3-6 years) than were considered feasible in my study (5-10 years). Mass campaign coverages were lower (10-40%) in the Liu study, but were coupled with very high routine neonatal coverage (70-95%), and optimistic vaccine characteristics (100% lifetime efficacy against infection) and timing of introduction (2018). Although the neonatal campaigns delivered minimal impact, the mixed vaccination campaigns initiated in 2018 were able to achieve the 2035 WHO goal of an incidence rate of less than 10 cases per million population. The 100% efficacious POI vaccine in this chapter achieved only a 27% (UR: 23-33%) reduction compared to the no vaccine scenario by 2035, but the comparatively greater incidence reduction in the Liu study was due to the very optimistic implementation, efficacy, and duration of protection assumptions. In addition, the lack of relapse after recovery from disease in the Liu model likely underestimates future burden of disease (see section 4.3.6.9).

The Arregui model compared neonatal to adolescent vaccination (15 year olds).[39] The vaccine modelled was a POI-only vaccine and allowed development of disease, so behaved as a 'leaky' vaccine. As will be discussed in more depth in Chapter 4, the results of the Arregui model demonstrate the importance of age-structured models,[39] including demographic structure and heterogeneous contact patterns by age, therefore supporting the decision taken to create such a highly age structured model in this research. This age structure is particularly important for investigation of vaccine age targeting to adolescents compared to older adults in Chapter 4 of this thesis.

3.5.3 Study Strengths

This is the first study to provide an in-depth exploration of the impact of such a wide range of vaccine efficacies, durations of protection, and vaccination populations (pre- and post-infection). It is also the first to explore TB vaccine efficacies against infection and disease in combination, and the first China model to explore a more realistic array of vaccine characteristics. Such a comprehensive exploration of characteristics provides a unique data set to identify the characteristics and vaccination populations

required to maximise vaccine impact, and also to identify the characteristics required to achieve a minimum acceptable level of impact to inform TB vaccine TPPs, help inform clinical trial design (e.g. sample size and endpoints) and selection of recruitment populations. As more clinical trial data become available, model outputs can be used to predict future impact of pipeline vaccines.

Age-specificities of the model are discussed in more depth in Chapter 4, but it is worth mentioning that the model is highly reflective of the age structure of the population and of the TB epidemic in China. Particularly important in developing such age structured projections were the calibration to age-stratified demographic and epidemiological data, through age structure in the modelled parameters, including age-wise contact patterns. At time of development, this was the first TB vaccine model to incorporate heterogeneous mixing patterns by age. A recent publication has also incorporated mixing patterns, but used European data for global mixing patterns,[39] whereas in this model I improve on this by using contact pattern data specific to China.[15]

3.5.4 Study Limitations

There were several limitations to this study, most of which relate to model assumptions discussed elsewhere, but are mentioned again below.

The model is calibrated to a series of age-stratified country-level data. The mortality rate estimates for China are relatively low given the disease rates, therefore to achieve a model fit, wide mortality rate confidence interval targets for calibration and a relatively low case fatality rate were assumed. This should have minimal impact on the percentage mortality rate reduction, but means estimates of the number of deaths averted is likely an underestimate.

The model assumes geographically homogenous TB epidemiology in China, yet as highlighted by Gao *et al.* this is known not to be the reality. At this stage in vaccine development, the regional level of granularity is not required for vaccine

development decision-making and was therefore not required from this model. However, once a vaccine is available, regional tailoring of this model could be valuable for evidence-informed decision making for implementation at the regional level.

In terms of the model parameters, it was assumed that the rates of progression to active TB remained constant throughout the study period. However, given socioeconomic improvements over time, this approach might underestimate the progression rates in the earlier years and overestimate them in to the future.

In the current model structure, when recovered populations are reinfected they either develop primary disease or return to the latent state. Although this is an acceptable assumption, remaining in the recovered state instead of returning to the latent state could be a possible alternative assumption. Given the reactivation rates in recovered populations are generally assumed higher, this would likely lead to higher projected TB reactivation and burden estimates, therefore increasing the importance of vaccines for preventing infected populations against development of disease.

Case detection and successful treatment was assumed to occur in one time step, and lead to entering the recovered state. Given the majority of cases on appropriate treatment are thought to become non-infectious by two weeks on treatment, compared to the six-month time step of the model, this was considered an appropriate assumption. However, those not detected and successfully treated enter the prevalent state, and unless transitioning from non-infectious to infectious TB, do not experience another opportunity for treatment. Such details of the care cascade are beyond the scope of this research question, but could be approached in further work, especially if MDR TB were to become a focus of the research questions.

There are many unknowns in the implementation of these vaccines, therefore standardised assumptions with regards to delivery parameters such as coverage, age

of vaccination and frequency of mass vaccination have been made. It is unlikely that all adults would be vaccinated in mass campaigns, as some tailoring by local epidemiology would be likely. However, the all-age mass campaigns were selected to ensure coverage of the epidemiological peak(s), therefore if adolescent and/or adult targeting were conducted based upon robust evidence of the epidemiological peak, it would be expected that the overarching conclusions presented here would still hold. A second frequency of mass vaccination is considered in the scenario analysis, but others may be possible once more information about a new candidate's duration of protection and efficacy is known. As demonstrated in the scenario analysis, frequency of mass vaccination can have a substantial impact on the achievable population-level impact of shorter duration vaccines.

Modelled estimates of the median vaccine efficacy assumed 10-yearly mass campaigns, and were made for the duration allowing the minimum acceptable impact (5 years protection) and the duration above which minimal epidemiological gains were made (10 years). However, if more frequent mass campaigns were considered beneficial and affordable, lower vaccine efficacies could be acceptable, as was demonstrated in the scenario analysis. Also, if a higher minimum level of impact were required, a higher duration of protection may be required for calculation of vaccine efficacy medians. This could be explored in an additional analysis if needed.

The time horizon of vaccine implementation may be important for model outcomes. As discussed in the systematic review of TB vaccine models (Chapter 2),[8] longer time horizons can potentially influence the relative impact of pre- and post-infection vaccines, as even though post-infection vaccines may have greater initial impact, impact with pre-infection vaccines increases with longer time horizons.

The results presented here are representative of the Chinese epidemic, which is an epidemic driven by reactivation disease in an ageing population that has experienced high historical exposure. These results may be generalisable to other similar settings, such as other high burden ageing epidemics (e.g. Thailand, Russia), but it should be

noted that the relative impact of different characteristics is likely to differ in other epidemiological scenarios, and will be investigated in the planned continuation of this work exploring vaccine impact in India and South Africa.

3.5.5 Implications for vaccine development and policy

This research highlights the need for careful consideration of the criteria for key vaccine characteristics such as duration of protection (keeping feasible implementation in mind), and vaccine efficacy against infection or disease when developing TB vaccine Target Product Profile and Clinical Development Plans.

For development strategists, such as BMGF, WHO, cross-product bodies and developers, a key implication of this research is the importance of efficacy for prevention of disease for maximising impact on disease burden in China during 2025-2050. Although there was some value in prevention of infection, vaccine development strategies for China and similar epidemics should aim to accelerate development of candidate vaccines with at least some anticipated protection against disease.

There are several implications of these results for clinical trial, clinical development plan and stage gate design, relevant to development strategists and clinical triallists. Some higher levels of incidence rate reduction were only achievable with at least some vaccine efficacy against disease. Therefore, an infection outcome would not be sufficient to guarantee these higher levels of impact. If high levels of incidence rate reduction are required for stage gating, disease outcomes should be measured in clinical trials, or preferably both infection and disease outcomes. The vaccine efficacy required for a "go" decision when stage gating will most likely be higher if only one outcome is measured in a clinical trial (i.e. infection or disease), as if both outcomes are measured a non-zero efficacy in the second outcome can reduce the required efficacy in the first.

To achieve an absolute minimum incidence rate reduction of 20-29%, at least 5 years protection will be required if mass campaigns are every 10 years. Therefore, clinical trials should be designed to provide data on duration of protection beyond the usual 2-3 years of a phase IIB trial. For example, phase II clinical trials could include an extended immunological or infection/disease outcome follow-up in a subset of participants. Alternatively, if longer-term follow up data are unavailable at licensure, or the duration of protection of a product is known to be short, the results of this model have implications for decisions by country-level decision makers, as planning for more frequent revaccination campaigns will be critical to ensure the desired population level impact is achieved.

3.6 Supplementary materials: Chapter 3

3.6.1 Additional methods: Model equations

The series of difference equations comprising the vaccinated and unvaccinated model strata are provided. The first set of equations is valid for all time steps *except* that at the start of the year. The second section provides equations for the first time step of the year, and includes ageing, vaccination and vaccine waning.

The equations for the five *M.tb* sub-populations (uninfected (susceptible), latently infected, infectious active disease, non-infectious active disease, and recovered) in year *k*, time step *i* and age *j* are provided. The size of the time step is dt, which in this research was selected as 0.5 years as a balance between sufficient granularity and computing time/capacity. The time step in the model is given by $\mathrm{i} = \frac{\mathrm{k}-(\mathrm{start\ year})}{\mathrm{dt}} + 1$ for the first time step of the year and $\mathrm{i} = \frac{\mathrm{k}-(\mathrm{start\ year})}{\mathrm{dt}} + 2$ for the second time step of the year.

New-borns (births) in year *k* (*B[k]*) entered the population as uninfected at the start of the year. Life span was limited to no more than 100 years. Following the methodology of Schenzle (1984),[40] ageing was implemented in the model on an annual basis by transitioning the population in a given sub-population of age *j*, to the same sub-population of age *j*+1 at the very end of each year.

For the baseline scenario of no vaccination, $\theta_S, \theta_L, \theta_R$ and *d* were set to zero. In the vaccine scenarios, vaccination was delivered by setting the relevant theta to a non-zero value equal to vaccine coverage for the appropriate age groups during the first time step of each vaccination year.

Transmission

$$\lambda[i,j] = \eta cal \sum_{y=1}^{y=nygrp} \eta[m,y] z \left(\frac{I[i,y] + I_V[i,j]}{T[i,y]} \right)$$

Where $T[i,y] = \sum_{j=jmin}^{j=jmax} \left[S[i,j] + L[i,j] + I[i,j] + NI[i,j] + R[i,j] + S_V[i,j] + L_V[i,j] + I_V[i,j] + NI_V[i,j] + R_V[i,j] \right]$.

nygrp was the number of contact age groups, *m* was the age group of the individual exposed to infection (including age *j*), *y* was age group of contacts, $\eta[m,y]$ was number of respiratory contacts of age group *m* with contacts of age group y, ηcal was the calibration factor for model fitting, z was the probability of transmission per respiratory contact between an infectious active case and a susceptible person (which is later scaled for protection afforded by latent infection), and *jmin and jmax* were the lower and upper bounds of age classes within a contact age group (*y*).

Ageing was implemented using the method of Schenzle (1984).[40] In the first time step (*i*) of any given year, the updated values for populations of age *j* were functions of those aged one year younger (*j*-1) in the previous time step (*i*-1).

CDR[k,j] was the case detection rate (proportion) for year *k* in age *j*, and *CoT[k]* was the proportion of detected cases that were treated successfully in a given year.

Vaccination (θ) and waning of protection (*d*) were assumed to occur in the first time step of a given year. In the middle of the year, the only transitions for vaccinated populations were between states within the vaccinated stratum. As the vaccine was modelled as 'leaky', all those receiving vaccine were moved to the vaccinated stratum, thus $\theta_a[j,k]$ was equal to percentage coverage for age group *j* of infection state *a* in year *k.* Here, $d[i,j]$ is the risk of ending vaccine protection at time step *i* and age *j*. The vaccinated terms are not multiplied by *dt* as they only occur at set time steps in the year.

The parameters affected by vaccination depended upon the vaccine type. Prevention of infection vaccines with vaccine efficacy *effI* reduced infection rates (λ) by 1-*effI,* and prevention of disease efficacy (*effD*) reduced the parameters for reactivation/relapse from latency and recovered (*v* and *r*) and proportion developing primary active disease (*p*) by 1-*effD*. For vaccines only efficacious when delivered pre-infection (PRI), θ_L and θ_R were zero. For vaccines only efficacious when delivered post-infection (PSI), θ_S was zero, therefore S_V remained at zero. For those vaccines efficacious both pre- and post-infection, the vaccine was delivered to all groups that did not have active disease (i.e. non-zero and equal numbers for θ_S, θ_L and θ_R). Transitions between the vaccinated and unvaccinated active disease classes were unidirectional, only occurring upon waning of protection.

Time steps NOT the first time step of the year

Unvaccinated (non-first time step)

Susceptibles

$$S[i,j] = S[i-1,j] - (u[j] + \lambda[i-1,j])S[i-1,j]dt$$

Latent

$$L[i,j] = L[i-1,j] + \lambda[i-1,j](1-p[j])(S[i-1,j] + xR[i-1,j])dt - (v[j] + \lambda[i-1,j]p[j]x + u[j])L[i-1,j]dt$$

New infectious active TB cases

$$new\_I[i,j] = \lambda[i-1,j]p[j]f[j](S[i-1,j] + xL[i-1,j] + xR[i-1,j])dt + v[j]f[j]L[i-1,j]dt + r[j]f[j]R[i-1,j]dt + wNI[i-1,j]dt$$

New non-infectious active TB cases

$$new\_NI[i,j] = \lambda[i-1,j]p[j](1 - f[j])(S[i-1,j] + xL[i-1,j] + xR[i-1,j])dt + v[j](1 - f[j])L[i-1,j]dt + r[j](1 - f[j])R[i-1,j]dt$$

Infectious active TB cases

$$I[i,j] = I[i-1,j] + (1 - CDR[k,j]\times CoT[k])new\_I[i,j] - (n[j] + u[j] + ui[j])I[i-1,j]dt$$

Non-infectious active TB cases

$$NI[i,j] = NI[i-1,j] + (1 - CDR[k,j]\times CoT[k]\times e)new\_NI[i,j] - (n[j] + u[j] + uni[j] + w)NI[i-1,j]dt$$

Recovered

$$R[i,j] = R[i-1,j] + n[j](I[i-1,j] + NI[i-1,j])dt + (CDR[k,j]\times CoT[k])(new\_I[i,j] + e\times new\_NI[i,j]) - (r[j] + \lambda[i-1,j]x + u[j])R[i-1,j]dt$$

Vaccinated (non-first time step)

Vaccinated susceptibles

$$S_V[i,j] = S_V[i-1,j] - (u[j] + (1 - effI)\lambda[i-1,j])S_V[i-1,j]dt$$

Vaccinated latent

$$\begin{aligned} L_V[i,j] \;=\; & L_V[i-1,j] \\ & + (1-effI)\lambda[i-1,j](1-(1-effD)p[j])(S_V[i-1,j] \\ & + xR_V[i-1,j])dt \\ & - \big((1-effD)v[j] + (1-effI)\lambda[i-1,j]\times(1-effD)p[j]x \\ & + u[j]\big)L_V[i-1,j]dt \end{aligned}$$

New vaccinated infectious active TB cases

$$\begin{aligned} new\_I_V[i,j] = \; & (1-effI)\lambda[i-1,j](1-effD)p[j]f[j](S_V[i-1,j] \\ & + xL_V[i-1,j] + xR_V[i-1,j])dt \\ & + \; (1-effD)v[j]f[j]L_V[i-1,j]dt \\ & + (1-effD)r[j]f[j]R_V[i-1,j]dt + wNI_V[i-1,j]dt \end{aligned}$$

New vaccinated non-infectious active TB cases

$$\begin{aligned} new\_NI_V[i,j] = \; & (1-effI)\lambda[i-1,j](1-effD)p[j](1 \\ & - f[j])(S_V[i-1,j] + xL_V[i-1,j] + xR_V[i-1,j])dt \\ & + \; (1-effD)v[j](1-f[j])L_V[i-1,j]dt + (1-effD)r[j](1 \\ & - f[j])R_V[i-1,j]dt \end{aligned}$$

Vaccinated infectious active *TB cases*

$$\begin{aligned} I_V[i,j] = I_V[i-1,j] + \; & (1-CDR[k,j]\times CoT[k])new\_I_V[i,j] \\ & - \; (n[j]+u[j]+ui[j])I_V[i-1,j]dt \end{aligned}$$

Vaccinated non-infectious active TB cases

$$\begin{aligned} NI_V[i,j] = NI_V[i-1,j] + \; & (1-CDR[k,j]\times CoT[k]\times e)new\_NI_V[i,j] \\ & - \; (n[j]+u[j]+uni[j]+w)NI_V[i-1,j]dt \end{aligned}$$

Vaccinated Recovered

$$R_V[i,j] = R_V[i-1,j] + n[j](I_V[i-1,j] + NI_V[i-1,j])dt + (CDR[k,j]\times CoT[k])(new\_I_V[i,j] + e\times new\_NI_V[i,j]) - \big((1-effD)r[j] + (1-effI)\lambda[i-1,j]x + u[j]\big)R_V[i-1,j]dt$$

<u>First time step of the year (ageing and vaccine delivery/waning)</u>

Unvaccinated (first time step of the year)

Susceptibles

If *j*=1: $S[i,1] = B$

If *j*≠1:

$$S[i,j] = S[i-1,j-1] - (u[j-1] + \lambda[i-1,j-1])S[i-1,j]dt - \theta_S[k,j]S[i-1,j-1] + d[k,j](1-\theta_S[k,j])S_V[i-1,j-1]$$

Latent

$$L[i,j] = L[i-1,j-1] + \lambda[i-1,j-1](1-p[j-1])(S[i-1,j-1] + xR[i-1,j])dt - (v[j-1] + \lambda[i-1,j-1]p[j-1]x + u[j-1])L[i-1,j-1]dt - \theta_L[k,j]L[i-1,j-1] + d[k,j](1-\theta_L[k,j])L_V[i-1,j-1]$$

New infectious active TB cases

$$new\_I[i,j] = \lambda[i-1,j-1]p[j-1]f[j-1](S[i-1,j-1] + xL[i-1,j-1] + xR[i-1,j-1])dt + v[j-1]f[j-1]L[i-1,j-1]dt + r[j-1]f[j-1]R[i-1,j-1]dt + wNI[i-1,j-1]dt$$

New non-infectious active TB cases

$$new\_NI[i,j] = \lambda[i-1,j-1]p[j-1](1 - f[j-1])(S[i-1,j-1] + xL[i-1,j-1] + xR[i-1,j-1])dt + v[j-1](1-f[j-1])L[i-1,j-1]dt + r[j-1](1-f[j-1])R[i-1,j-1]dt$$

Infectious active TB cases

$$I[i,j] = I[i-1,j-1] + (1 - CDR[k,j]\times CoT[k])new\_I[i,j] - (n[j-1] + u[j-1] + ui[j-1])I[i-1,j-1]dt + d[k,j]I_V[i-1,j-1]$$

Non-infectious active TB cases

$$NI[i,j] = NI[i-1,j-1] + (1 - CDR[k,j]\times CoT[k]\times e)new\_NI[i,j] - (n[j-1] + u[j-1] + uni[j-1] + w)NI[i-1,j-1]dt + d[k,j]NI_V[i-1,j-1]$$

Recovered

$$R[i,j] = R[i-1,j-1] + n[j-1](I[i-1,j-1] + NI[i-1,j-1])dt + (CDR[k,j]\times CoT[k])(new\_I[i,j] + e\times new\_NI[i,j]) - (r[j-1] + \lambda[i-1,j-1]x + u[j-1])R[i-1,j-1]dt - \theta_R[k,j]R[i-1,j-1] + d[k,j](1-\theta_R[k,j])R_V[i-1,j-1]$$

Vaccinated (first time step of the year)

Vaccinated susceptibles

$$S_V[i,j] = S_V[i-1,j-1] - (u[j-1] + (1-effI)\lambda[i-1,j-1])S_V[i-1,j-1]dt + \theta_S[k,j]S[i-1,j-1] - d[k,j](1-\theta_S[k,j])S_V[i-1,j-1]$$

Vaccinated latent

$$L_V[i,j] = L_V[i-1,j-1] + (1-effI)\lambda[i-1,j-1](1-(1-effD)p[j-1])(S_V[i-1,j-1]+xR_V[i-1,j-1])dt - \big((1-effD)v[j-1] + (1-effI)\lambda[i-1,j-1]\big)\times(1-effD)p[j-1]x+u[j-1]\big)L_V[i-1,j-1]dt + \theta_L[k,j]L[i-1,j-1] - d[k,j](1-\theta_L[k,j])L_V[i-1,j-1]$$

New vaccinated infectious active TB cases

$$new\_I_V[i,j] = (1-effI)\lambda[i-1,j-1](1-effD)p[j-1]f[j-1](S_V[i-1,j-1]+xL_V[i-1,j-1]+xR_V[i-1,j-1])dt + (1-effD)v[j-1]f[j-1]L_V[i-1,j-1]dt + (1-effD)r[j-1]f[j-1]R_V[i-1,j-1]dt + wNI_V[i-1,j-1]dt$$

New vaccinated non-infectious active TB cases

$$new\_NI_V[i,j] = (1-effI)\lambda[i-1,j-1](1-effD)p[j-1](1-f[j-1])(S_V[i-1,j-1]+xL_V[i-1,j-1] + xR_V[i-1,j-1])dt + (1-effD)v[j-1](1-f[j-1])L_V[i-1,j-1]dt + (1-effD)r[j-1](1-f[j-1])R_V[i-1,j-1]dt$$

Vaccinated infectious active TB cases

$$I_V[i,j] = I_V[i-1,j-1] + (1-CDR[k,j]\times CoT[k])new\_I_V[i,j] - (n[j-1]+u[j-1]+ui[j-1])I_V[i-1,j-1]dt - d[k,j]I_V[i-1,j-1]$$

Vaccinated non-infectious active TB cases

$$NI_V[i,j] = NI_V[i-1,j-1] + (1 - CDR[k,j]\times CoT[k]\times e)new\_NI_V[i,j] - (n[j-1] + u[j-1] + uni[j-1] + w)NI_V[i-1,j-1]dt - d[k,j]NI_V[i-1,j-1]$$

Vaccinated recovered

$$R_V[i,j] = R_V[i-1,j-1] + n[j-1](I_V[i-1,j-1] + NI_V[i-1,j-1])dt + (CDR[k,j]\times CoT[k])(new\_I_V[i,j] + e\times new\_NI_V[i,j]) - \big((1-effD)r[j-1] + (1-effI)\lambda[i-1,j-1]x + u[j-1]\big)R_V[i-1,j-1]dt + \theta_R[k,j]R[i-1,j-1] - d[k,j](1-\theta_R[k,j])R_V[i-1,j-1]$$

3.6.2 Additional Methods: Justification of parameters and calibration data

Justification of selection and sources for natural history, demographic, social and control parameters are provided in this section.

3.6.2.1 Natural history (biological) parameters

Natural history parameter prior ranges sampled from in the calibration process and the calibrated posterior ranges are summarised in Table 3.7. Natural history parameters were assessed for age variance based upon the literature and biological plausibility, as described below.

Parameters considered invariant by age or with insufficient data demonstrating age variability were probability of transmission per infectious contact (*z*), protection from active disease due to latent infection (*x*) and conversion from non-infectious to infectious active case (*w*). These were modelled with the same value for all ages based on estimates from the literature. The ranges sampled from during calibration were in-line with values applied in Knight et al. and historical literature.[1,13,41]

However, age-related differences in immunity, such as immune immaturity in children and immunosenescence and co-morbidities in the elderly, are believed to influence the course of infection and disease.[42,43] Therefore, to improve the modelled epidemiology, parameters with an epidemiological or biological basis for age-based differences were calibrated separately for each age group. These included the proportion progressing directly to active disease (*p)*, proportion developing infectious disease (*f)*, risk of reactivation of latent infection or relapse of recovered disease (*v* and *r*), risk of natural cure from active disease (*n*) and an age-wise calibration factor for TB mortality rate (*uiscale*). The age groups with calibrated parameters were children (0-14 years), adolescents and adults (15-64 years) and the elderly (≥65 years).

Where available, parameter ranges for calibration were based upon available literature (see references in Table 3.7). However, little literature exists to inform parameters for older adults and the elderly,[43] and no previous models were identified with age-specific parameters for these age groups. Based upon published literature and expert opinion comparing the impact on immune responses of HIV versus ageing, it was assumed that HIV-positive parameter ranges were an appropriate proxy for the upper bound of immunosenescence in old age (Richard Aspinall, oral communication, 5th June 2015).[44] When HIV-positive ranges did not overlap with HIV-negative adult parameters, the lower bound of HIV-positive ranges were taken (*v* and *r*), and where overlap occurred the upper range of HIV-positive parameters were used (*p*, *f* and *n*). Therefore, the ranges for calibration of elderly parameters encompassed the HIV-negative adult parameters up to either the least extreme or all values of the HIV-positive parameters (see Table 3.7). Immune senescence is a gradual process, therefore to smooth the transition between adult and elderly parameters for *v, r, and n,* 55-64 year olds were programmed to take the mean of the 15-54 years and elderly calibrated values.

In model calibration, the age-specific values for a given parameter were sampled independently. Where data or biological plausibility were suggestive of an age-wise relationship between parameters, constraints were set to ensure only parameter sets adhering to those relationships were retained. For example, given TB mortality is higher at the age extremes,[43] parameter sets were only retained if the child and elderly TB mortality calibration factors (*uiscaleC* and *uiscaleE*) were greater than or equal to the adult parameter (*uiscaleA)*.

The model parameter determining the distribution between infectious and non-infectious TB (*f*) was constrained so that the proportion of infectious disease in the elderly was equal to or less than the adult parameter. This is based upon literature suggesting that although approximately 75% of elderly cases manifest as pulmonary disease,[45] many present atypically with minimal pulmonary symptoms.[45,46] In addition, an increasing proportion of extrapulmonary disease is observed with increasing age in this population.[43,47]

The proportion of the population directly developing primary disease (*p*) and reactivation/relapse disease (*v* and *r*) may be higher in the elderly due to immunosenescence weakening the immune system's ability to control infection.[46,47] Reactivation has been demonstrated to occur decades after infection, therefore even historical latent infection can reactivate.[48] An epidemiological model exploring the incubation period between infection and disease demonstrated that incubation periods became shorter with age at infection,[49,50] supporting the assumption that the risk of primary progression to disease is likely higher in older age groups. Therefore a constraint was included to ensure the elderly parameter was at least equal to the adult parameter. Conversely, natural cure (*n*) was considered potentially less likely in elderly populations, so parameter sets with equal or lower natural cure in the elderly were retained. Details of these constraints can be found in Table 3.7.

Two fitting factors (*rmortTB* and *uiscale*) were employed for calibration of TB mortality in infectious and non-infectious disease (u_i *and* u_{ni}). First TB mortality was scaled by calibration factor *rmortTB* (increasing mortality if *rmortTB* greater than zero, and reduced if less than zero), then scaled by age with age-specific calibration factors for children, adults and the elderly (*uiscaleC, uiscaleA,* and *uiscaleE*).

3.6.2.2 Demographic parameters

The population size at model initiation, and birth and death rates over time form the basis for the demographic structure of the modelled population. The model was initiated in 1900, with initiation population size (p0) set as part of the calibration process (see 3.6.3.1). The 'burn in' period (1900-1999) ahead of the period of interest (2000-2050) ensured the populations in the calibration (2000-2010) and outcome (2025-2050) periods had appropriate levels of historical exposure to produce relevant age-wise rates of infection and disease.

Demographic data were obtained from the UN population division 2012 revision.[4] The initial model birth rate was calculated by dividing UN average annual birth estimates for 1975-1980 by the UN reported 1980 population of China.[4] The initial

birth rate was implemented from initiation to 1989. To account for the impact of China's one-child policy, a lower birth rate (based upon 2010 estimates) was implemented from 1990-2009. Although the one-child policy was introduced in 1979, data indicate policy implementation was not instantaneous, as birth rates did not begin to decline until 1989/1990, therefore the lower birth rates were implemented in 1990 as opposed to 1979.[51] From 2010 onwards, modelled birth rates were estimated annually using UN fertility and population estimates/predictions.[4] In the model, number of births each year was estimated by applying birth rates as described above to modelled population size.

The all-cause mortality was calculated in the model by year of age and calendar year. This was modelled by using UN life expectancy data for China to calculate annual risk of death by age [1/(life expectancy at a given age)], for 1950, 2000, and annually from 2010-2050, during which the rate is constant in 5-year steps, but varies by age.[4] To fit mortality by age and over time, it was necessary to have steps in mortality rates to reflect historical changes in mortality in China. Background mortality was higher historically due to social and healthcare determinants, so in the years preceding 2000, mortality estimates calculated from UN population data for 1950 were used, as those who were elderly in the calibration period would have been born around this time, so reflected their historical life expectancy. From 2000-2009, the mortality estimates for 2000 were used, then from 2010 to 2050 the modelled mortality tracked the UN reported data and predictions.[4]

Table 3.7: Natural history and demographic parameters (tables adapted from Knight et al.[1])

Parameter	Symbol	Description	Prior proposal range and constraints	References	Parameter range observed in 1000 model fits
Births and all-cause death	*B[k]*	Number of births in year *k*	Birth rate calculated from UN fertility and population estimates and applied to modelled population. 1980s birth rate pre-1990, 2010 value 1990-2009, and tracks annual UN data from 2010 onwards.	UN Population Division (revision 2012)[4]	n/a
	u[j]	Background (all-cause) death risk at age *j*	*u[j] = 1/LE[j]* Death rate calculated from UN life expectancy estimates by age for China and applied to modelled population. 1950s birth rate pre-2000, 2000 value 2000-2009, and tracks annual UN data and predictions from 2010 onwards.	UN Population Division data (2012 revision) [4]	n/a

	*rmort**	Calibration factor for all-cause mortality	Calibrated at single value to match population size by age group in 2010 and 2050.
			Range: -1 to 1
			If (*rmort* < 0): *u[j]*=(1+*rmort*) *u[j]*
			If (*rmort* ≥0): *u[j]*=(1-*u[j]*)rmort + *u[j]*
					-0.6

Transmission	*λ[i,j]*	*M.tb* transmission risk (force of infection) in time step *i* for age *j*	Calculated in the model: $\lambda[i,j] = \eta cal \sum_{y=1}^{y=nygrp} \eta[m,y] z \left(\frac{I[i,y] + I_V[i,y]}{T[i,y]} \right)$		n/a

	ηcal	Scaling and calibration factor for daily number of respiratory contacts	Scales respiratory contacts to annual number of contacts and calibrates to TB incidence Calibration range: 0-5 (range based upon preliminary modelling)		1.85-2.82
	η[m,y]	Daily number of respiratory contacts by age group *m* and contacts in age group *y*	Calibrated by *ηcal* to match TB incidence Initial values and mixing patterns taken from Read *et al.* 2014,[15] made symmetrical to account for reporting bias.[15]		n/a
	z	Probability of transmission per respiratory contact between an infectious and uninfected individual	Fixed value from literature: 0.1	Dye et al 2008[52] Abu Raddad et al 2009 [41] Knight et al 2014[1]	0.1
Progression to active disease	*p[j]*	Proportion of (re-) infected Uninfected, Latents or Recovereds directly developing active TB (fast/ primary disease), by year of age	*p[j < 15]* = 0.01-0.06 *p[j ≥ 15, <65]* = 0.08 – 0.2 *p[j ≥ 65]* = 0.08-0.36 (adult and HIV-positive range) Constraint: Parameter set only retained if elderly parameter selected was greater than or equal to adult parameter.	A Dye et al 1998[13,*] Knight et al 2014[1] Abu Raddad et al 2009 [41,**] Dye et al 2008[52]	*p[j < 15]* = 0.014- 0.030 *p[j ≥15, <65]* = 0.14-0.2 *p[j ≥65]* = 0.21- 0.36
	x	Protection from re-infection or developing active TB due to being latently infected or	*(1-x)* is the value for the level of protection afforded Range: 0.25 – 0.41	Vynnycky and Fine 1997 [53] Abu Raddad et al 2009 [41] Dye et al 2008[52] Gomes et al 2007[54]	0.25-0.36

		recovered from infection			
	v[j]	Risk of reactivation in latently infected population	*v [j<15]* = 0.0001 - 0.0003 *v [j≥15, j<65]* = 0.0001 - 0.0003 *v[j≥65]* = 0.0001 - 0.04 (adult, and up to lower bound of HIV-positives) Constraint: Parameter set only retained if elderly parameter selected was greater than or equal to adult parameter.	Dye et al 1998[13] Gomes et al 2007[54] Schulzer et al 1992[55] Knight et al 2014[1] Schaaf et al 2010[43]	*v[j <15]* = 0.00010-0.00019 *v [j≥15 & j<65]* = 0.00018 - 0.00028 *v[j* 55-64*]* = 0.00020-0.00194 *v[j ≥ 65]* = 0.00020 - 0.00369
Infectious TB	*f[j]*	Proportion of new active cases directly becoming infectious, by age	*f[j < 15]* = 0-0.15 *f[j ≥ 15, <65]* = 0.25-0.75 *f[j ≥ 65]* = 0.19-0.75 (point estimate from elderly study, range from adult and HIV-positive ranges) Constraint: Parameter set only retained if adult parameter selected was greater than or equal to elderly parameter.	Abu Raddad et al 2009[41] Dye et al 2008[52] Yokishawa et al 1992[45] Marion et al 2009[46] Rajagopalan and Yoshikawa 2000[47] Schaaf et al 2010[43]	*f[j<15]= 0.10-0.15* *f[j ≥ 15, <65] = 0.66-0.75* *f[j ≥ 65]* = 0.61-0.75
	w	Risk of converting from non-infectious to infectious active case	0.007 – 0.02	Dye et al 1998[13] Ferebee et al 1970[56]	0.009-0.018
TB mortality	u_i;	Death risk for infectious untreated TB, varies by age	0.6 Calibrated to TB mortality by *rmortTB* and *uiscale* (see below)	Tiemersma et al 2011[57]	/
	u_{ni}	Death risk for non-infectious untreated TB, varies by age	0.21 Calibrated to TB mortality by *rmortTB* and *uiscale* (see below)	Tiemersma et al 2011[57]	/

	*rmortTB**	Calibration factor for u_i and u_{ni}	Sampled to calibrate u_i and u_{ni} to TB mortality.
			Range sampled: -1 to 1
			
			If (*rmortTB* < 0): $u_i^a=(1+rmortTB)u_i$, $u_{ni}^a=(1+rmortTB)u_{ni}$
			If (*rmortTB* ≥0): $u_i^a=(1-u_i)rmortTB + u_i$, $u_{ni}^a=(1-u_{ni})rmortTB + u_{ni}$
					-0.96 to -0.75

	uiscale	Calibration factor to vary u_i and u_{ni} by age	*uiscale* [*j* < 15] = 0-2 *uiscale [j ≥ 15, <65]* = 0-2 *uiscale [* j ≥ 65] = 0-2 *Where* u_i^b= uiscale*u_i^a Constraint: Parameter set only retained if child and elderly parameters selected were greater than or equal to adult parameter.[43]		*uiscale [j < 15] = 0.98-2.00* *uiscale [j ≥ 15, <65] = 0.10-1.0* *uiscale [j ≥ 65] = 0.18-1.56*
Natural cure and relapse	*n*	Annual risk of natural cure for TB cases, varies by age	*n[j<55]* = 0.1 – 0.25 *n[j 55-64]* = (*n[j<55] + n[j ≥ 65])/2* *n[j ≥ 65]* = 0.1 – 0.25 Constraint: Parameter set only retained if adult parameter selected was greater than or equal to elderly parameter.	Abu Raddad et al 2009[41] Dye et al 2008[52]	*n[j<55]* = 0.16-0.24 *n[j 55-64]* = 0.13-0.18 *n[j ≥ 65]* = 0.10-0.15
	r	Annual risk of relapse from recovered to active TB, varies by age	*r[j* <15] = 0.005-0.015 *r[j ≥ 15, <55]* = 0.005-0.015 *r[j ≥ 55, <65] =* (*r[j<55] + r[j ≥ 65])/2* *r[j ≥ 65]* = 0.005-0.2 (adult, and up to lower bound of HIV-positives)	Gomes et al 2004[58] Knight et al 2014[1] Schaaf et al 2010[43]	*r[j* <15] = 0.005-0.010 *r[j ≥ 15, <55]* = 0.005-0.007 *r[j ≥ 55, <65]* = 0.005-0.015 *r[j ≥ 65]* = 0.005-0.025

3.6.2.3 Social mixing parameters

Heterogeneous age-specific social mixing patterns were incorporated into the model. Previous TB vaccine models have assumed homogenous random mixing patterns within modelled populations,[1,41,52,59] but such homogenous mixing does not appropriately represent the reality of age-assortativity in contact patterns. One recently published model included age-assortativity in contacts, applying European contact data to global populations.[39] However, age structure and cultural factors in the surveyed population affect generalisability of contact patterns. In this model, I applied China-specific age-wise contact patterns from the Read *et al.* study to the modelled population in China to appropriately represent the likelihood of a contact event occurring.[15]

The Read study estimated the total and average daily number of contacts by age strata (0-5, 6-19, 20-64 and ≥ 65years) from contact diaries, in a study enrolling 1821 participants from southern China.[15] Contacts were defined as face-to-face conversation or skin-on-skin touch, so are broadly representative of the type of contact relevant to *M.tb* transmission.[15] The study demonstrated heterogeneity in the average number of contacts between age groups, in particular that the largest number of contacts tended to be within the participant's age group or with the 20-64 years age group (Figure 3.22, left).[15] Importantly, the total number of daily contacts was lowest for the elderly population. The clear differences in absolute number of contacts made by each age group and the age assortativity in the contacts made provided a clear rationale for incorporating age-specific contact patterns in this model.

The heterogeneous mixing pattern from this study was incorporated in the model in the term $\eta[m, y]$. This is a matrix of the average number of contacts per person per day between an individual in the participant group (*m*) and individuals in a contact age group (*y*) (c_{ym}) (Figure 3.22). In the calibration process, this matrix is multiplied by a scaling factor (*ncal*). Reporting or participation biases in the study produced asymmetry in the number of contacts between pairs of age groups in the reported

contact matrix (Figure 3.22, left).[15] Symmetry was achieved using the methods of Baguelin et al.,[60] by estimating the total number of contacts from all subjects in participant age group *m* with contact age group *y*, and for the age groups reversed, taking an average of the two, and converting back to a contact rate by dividing the number of contacts by the size of the participant age group (equation below and Figure 3.22, right):

$$c_{ym} = \left(\frac{1}{2}\left((d_{ym} * T_m) + (d_{my} * T_y)\right)\right)/T_m$$

where d_{ym} is average number of contacts of participants in group m with people in group y, d_{my} is average number of contacts of participants in group *y* with people in group m, and T_m and T_y are the number of participants in age group m and y, respectively.[60]

In the model, this contact matrix is then multiplied by the probability of transmission occurring upon a respiratory contact between an infectious and uninfected (*z*) and by the proportion of the contact population that are infected $\left(\frac{\mathrm{I[i,y]}}{\mathrm{T[i,y]}}\right)$. These are summed over the contact age groups to give *Mtb* transmission risk (force of infection) in time step i for age j (*λ[i,j]*). As described in section 3.3.3, the matrix is also multiplied by a scaling factor (*ncal*) in the fitting process to calibrate the total number of contacts to the epidemiological data.

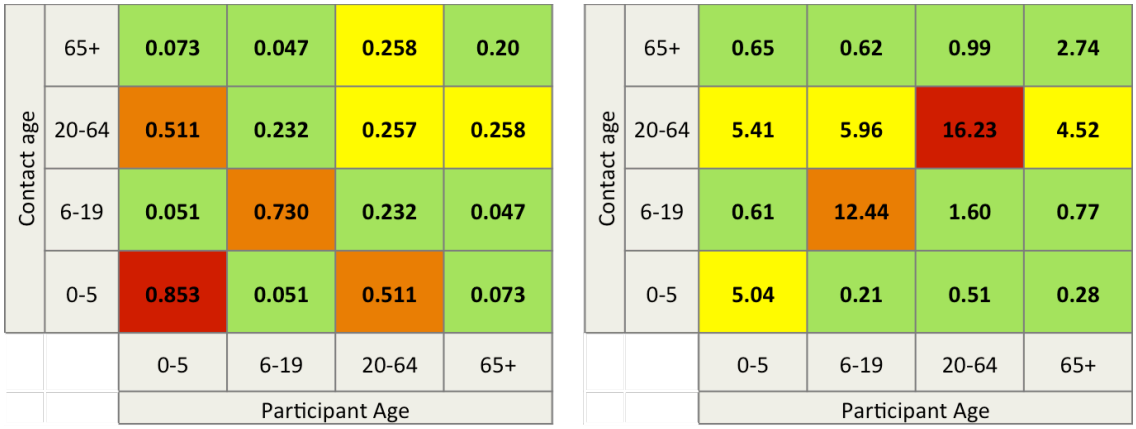


Figure 3.22: Average number of reported contacts per day (left, Read et al. 2014.[15]) and average number of contacts per person in the participant age group per day following matrix averaging to ensure symmetric total numbers of contacts (right)

3.6.2.4 Existing TB Control measure parameters

The existing control measure parameters are summarised in Table 3.8.

BCG coverage was assumed to remain at current levels as existing coverage is high and stable, therefore was not explicitly modelled, as the impact of BCG is intrinsic to the calibration data.

Data from China and some other countries have indicated that the case detection rate (CDR) may be lower in elderly populations than adults, due to diagnostic difficulties and reduced access to care.[61] Age-wise differences were reflected in the model by exploring age-specific CDR calibration factors for <55 years and the elderly, with parameter sets retained only if adult CDR was at least equal to the elderly parameter. An older adult calibration factor was calculated as the mean of the adult and elderly factors. These older adult parameters were to allow a more gradual parameter transition in to old age.

Although case detection estimates fluctuated over time, the data suggest a large, rapid increase from 35% to 74% between 2002 and 2005.[17] WHO 'case detection' requires diagnosis and reporting, whereas in the model it only represents diagnosis;

therefore, the low early estimates of case detection and the sudden rapid rise were likely an artificially rapid jump reflecting a more steady (2000-2010), but still large, increase in diagnosis ratios due to scale up of TB control efforts, coupled with a very rapid improvement in reporting. Therefore, the WHO China CDR data for 1990-2012 were used as the basis for developing a generalised logistic function to reflect a rapid but steady increase in CDR over this period, and a much slower but continued improvement (up to a limit of 100%) in detection beyond this. As described in the calibration section (3.6.3), this CDR was then calibrated to epidemiological data by age-specific scaling factors (CDRscale) for 'children, adolescents and adults', and the elderly, with 55-64 year olds as the average of younger adults and the elderly value. The value of the scaled CDR was capped to ensure scaling could not increase CDR beyond 100%. A scaling factor (*e*) for the CDR was included to account for the relative case detection of bacteriologically negative and extrapulmonary TB disease compared to bacteriologically positive disease.

WHO reported treatment success rates were employed between 1994 and 2011, and the 1994 and 2011 rates held constant before and after this period, respectively.[16] Although the lack of future improvement could be considered somewhat pessimistic, given the high levels of treatment success (95%) reported in recent years,[16,17] and that only 3.6% of the 197 countries studied in Knight et al. report a higher 2011 treatment success rate than China,[1,16] there remains minimal opportunity for improving existing control measures, therefore this was considered a likely scenario.

Table 3.8: Control measure parameters

Parameter	Symbol	Description	Prior proposal range and constraints	References	Parameter range observed in 1000 fits
Case detection	*CDR[k]*	Case Detection Ratio (proportion of new active TB cases detected and started on treatment) year k	2012 baseline: 89% CDR data for 1990-2012 were used as the basis for developing a generalised logistic function to reflect a rapid but steady increase in CDR over this period.	WHO TB burden estimates [17]	/
	CDRscale	Calibration factor for CDR, varies by age	Calibrated to match TB prevalence, mortality and incidence. Range sampled: -0.5 to 2 independently for <55year olds and ≥65 year olds. 55-64 year olds took the mean of the older and younger sampled values. Constraint: Parameter set only retained if adult calibration factor selected was greater than or equal to elderly parameter.[43] *CDR[k]=CDRscale*CDR[k]*		*CDRscale[j<55]* = 0.33-0.51 *CDRscale[j ≥ 55, <65]* = 0.09 – 0.28 *CDRscale[j ≥ 65]* = -0.19 - 0.10
	e	Relative case detection rate of non-infectious cases	0.4 – 0.8	Assumed [14]	0.58-0.80
Treatment	*CoT[k]*	Treatment success proportion (cured or complete treatment) in year *k*	2011 baseline: 95% WHO data 1994-2011, then constant CoT from 2011 onwards	WHO treatment outcomes [16]	n/a

3.6.3 Additional Methods: Calibration methodology

A two-stage calibration process was employed, with the first stage a manual calibration to age-stratified UN population projections for 2010 and 2050,[18] and the second stage an adaptive rejection ABC MCMC-based method using a modified *easyABC* package to calibrate to 18 country-specific epidemiological data ranges.

3.6.3.1 Demographic calibration

Demographic data, including age-stratified population size estimates for 2010 and 2050, were obtained from the UN population division 2012 revision (Table 3.9).[4] The model was adjusted to align with the overall population size and four age stratifications (0-14 years, 15-54 years, 55-64 years and ≥65 years) at these two time points. Uncertainty ranges were only available for 0-14 years and 15-54 years, as UN population estimates explore uncertainty in fertility assumptions, which do not impact older age groups over this time frame.[5]

Table 3.9: Demographic calibration targets for China

Calibration Factor	Year	Age (years)	Estimate	Lower bound	Upper bound
Population size (thousands)[4]	2010	All	1,359,822	n/a	n/a
		0-14	246,707	n/a	n/a
		15-54	863,710	n/a	n/a
		55-64	135,859	n/a	n/a
		65+	113,546	n/a	n/a
	2050	All	1,384,976	1,208,829	1,579,558
		0-14	204,187	124,968	300,552
		15-54	623,982	527,054	722,200
		55-64	225,492	n/a	n/a
		65+	331,315	n/a	n/a

Calibration of modelled demography to the above targets was achieved by manual adjustment of the population size at model initiation (p0) and a fitting factor for mortality rates (*rmort*). The fitted values, which gave the best overall fit to the selected calibration data were 1,267,142 million for p0 and -0.6 for *rmort*. Model demographic fits are shown in Figure 3.7.

3.6.3.2 Epidemiological calibration targets

Estimates and confidence intervals around age-stratified prevalence of bacteriologically-positive pulmonary TB disease were obtained from national prevalence surveys conducted in 2000 and 2010, reported in Wang *et al.* (2014).[19] Data from the 1990 survey were not used, as the diagnostic protocol was less sensitive, and age-stratified data were not reported. Although all-TB prevalence rates were available, clinically diagnosed TB included in these estimates was likely an overestimate of true prevalence due to over diagnosis based on clinical signs, and did not align well with the disease states in the model. Therefore, calibration was to the biologically-positive prevalence data, as these most closely represented the infectious disease state in the model (I).

Disease Surveillance Point data, considered nationally representative, were used to estimate mortality rates by age.[21] Published adjusted mortality estimates did not align with the age bands required for calibration, therefore age groups were brought together using published crude data to estimate unadjusted rates. In the published analyses the adjusted estimates were adjusted twice, first using methods for complex sampling surveys, and second for under-reporting of mortality. These adjustments were not reproducible with the available data and methodological information, but uncertainty ranges for calibration were based upon estimates of under reporting from other data sources. Causes of misclassification of TB deaths include HIV co-infection and malnourishment, and occur more frequently in deaths at home or in rural areas. Wang et al. and Yang et al. estimated TB mortality mis-classification (underestimation) in China, particularly due to miscoding as pneumonia,[62] at

approximately 50%.[63,64] Therefore the upper range of the calibration ranges was set at 50% higher than the unadjusted point estimates, which encompassed the published adjusted estimates. Although misclassification of non-cases as cases is less common, it does occur, so symmetric uncertainty ranges around the point estimate were assumed.

Country-level TB incidence stratified by age is not currently reported by the WHO, therefore age-specific notification rates of all TB (including smear positive, smear negative, and extra-pulmonary TB) were employed for model fitting.[20] Up to 20% over-reporting of notifications may occur due to the use of clinical algorithms as opposed to biological testing for diagnosis. Up to 20% under-reporting was estimated, as it was assumed that the reported data only captured public (CDC) notifications, which constitute 80% of cases, and also cases that are asymptomatic or do not exhibit standard symptoms may be missed.[65,66] Therefore, confidence intervals for calibration were +/-20% of the WHO-reported notification point estimates.[17]

To ensure an appropriate case detection rate in the fitting process, the model was also calibrated to the all-age TB incidence rate in 2010 obtained from the WHO database.[17] WHO incidence rates and ranges are retrospectively updated each year, therefore ranges fitted to are those for 2010, but accessed in July 2016.

Prevalence of infection surveys using TST and Quantiferon (QFN) tests in four rural sites in China provide age-stratified estimates of latent tuberculosis infection (LTBI) prevalence.[22] However these data were only used for comparison, not for calibration, as there was some uncertainty in generalisability of the four study sites to the whole of China. In addition, without a gold standard test for comparison, the sensitivity and specificity of TST and QFN tests are not fully understood.[67] Given the modelled latent infection captures all infected individuals, it is unknown the extent to which the study data would need to be adjusted to align with the modelled population.

3.6.3.3 Epidemiological calibration

Calibration to the 18 China-specific epidemiological data points was achieved using an adaptive rejection ABC MCMC-based method using a modified *easyABC* package.[23,24]

The parameter space in this model was highly multi-dimensional (24 natural history and case management parameters), therefore un-seeded MCMC would likely have required infeasible numbers or lengths of MCMC chains to identify suitable regions of the parameter space. Therefore, the adapted package was employed to allow for provision of seed parameter estimates to initiate the MCMC chains.[23]

To identify a series of suitable seed estimates, a maximum likelihood-based approach was implemented. One million parameter sets were randomly selected from literature-informed prior ranges and run through the 'no new vaccine' baseline model. Such a large number of parameter sets ensured thorough exploration of the parameter space. Variances were estimated for the calibration data, assuming the data were normally distributed. For each of the million model outputs, the joint likelihood function for observing the calibration data given the model outputs was calculated to assess goodness of fit of each. The twenty highest likelihood runs were selected as seeds for the adaptive rejection ABC MCMC-based calibration.

In the adaptive rejection ABC MCMC-based calibration, the aim was to fit within the calibration ranges of all 18 epidemiological data points (Table 3.10). Acceptance/rejection criteria in the MCMC chains were based upon a sufficient number (ranging 13-18, see adaptive fitting below) of epidemiological model outputs falling within the calibration ranges. Adaptive fitting was employed, increasing the acceptance criterion during a series of ABC MCMC chain sets from 13 to 18 data point fits, to ultimately identify parameter space fitting all 18 data points ('fully fitted').[68] In each set of chains, between five and twenty of the highest fits of the previous set were used as seeds for the subsequent set of chains, and the acceptance criterion

incrementally increased as the fit improved. Once parameter sets providing a fully fitted model had been identified using this method, five were selected randomly using the *runif* command as seeds for five 200,000-run ABC-MCMC chains. 2152 full fits were obtained from these final million ABC MCMC runs, of which one thousand were randomly selected using *runif*. As an adaptive ABC-MCMC was implemented, MCMC chains were not contiguous, therefore removal of initial burn-in period from the final chains was not required. This was checked by comparing outputs from the first and second half of the chains, which confirmed similarity throughout the chains. The selected 1000 parameter sets were used to estimate the median and uncertainty ranges of the model.

Table 3.10: Epidemiological calibration targets for China

Calibration Factor	Year	Age (years)	Estimate	Lower bound	Upper bound
All TB notification rate (/100,000/yr)[20]	2010	All	63.9	51.1	76.7
		0-14	2.7	2.2	3.2
		15-54	64.6	51.7	77.5
		55-64	104.4	83.5	125.2
		65+	143.1	114.5	171.7
TB Mortality Rate (/100,000/yr)[21]	2010	All	3.37	1.69	5.06
		0-14	0.29	0.15	0.44
		15-59	1.91	0.96	2.87
		60+	15.69	7.85	23.54
Microbiologically-positive pulmonary TB prevalence rate (/100,000/yr)[19]	2000	≥15	178	163	195
		15-29	92	72	116
		30-59	155	126	189
		60+	596	510	698
	2010	≥15	116	101	132
		15-29	59	40	86
		30-59	99	77	129
		60+	346	294	407
All age TB incidence rate (/100,000/yr)[17]	2010	All age	78	72	83

3.6.4 Additional Methods: Vaccine Parameters

Vaccine introduction was implemented in 2025, as this was considered the earliest likely date of implementation of a new vaccine. Modelled characteristics of novel vaccines have previously been categorised within four dimensions: the host infection status required for efficacy, the effect type, vaccine efficacy, and duration of protection.[8] An additional characteristic is whether the modelled vaccine efficacy is 'all-or-nothing' or 'leaky'. Implementation parameters included timing of vaccination, ages vaccinated and coverage.

3.6.4.1 Host infection status

Previous exposure of the immune system may impact the immune response to the vaccine antigen. Therefore, the efficacy of a given candidate may be reliant upon whether the vaccinated individual is *M.tb*-naïve, currently latently infected or ever-infected. Therefore, four infection status combinations were modelled (Table 3.11). Pre-infection vaccines were assumed to prevent development of active disease only in never-infected individuals ($\theta_S > 0, \theta_L = 0, \theta_R = 0$). Post-infection (PSI) vaccines were assumed efficacious when delivered to either latently infected populations or those who had recovered from active disease ($\theta_S = 0, \theta_L > 0, \theta_R > 0$). A vaccine effective pre- and post infection (P&PI), producing immunity in all except those with active disease ($\theta_S > 0, \theta_L > 0, \theta_R > 0$), was modelled to estimate the greatest preventative vaccine impact given the efficacy and other vaccine parameters. Therapeutic vaccines were not explored in this study, so none were considered effective in the populations with active disease.

Table 3.11: Host infection statuses in which each vaccine type is effective

Vaccine	Never infected (S)	Latently infected (L)	Active Disease (I or NI)	Recovered from disease (R)
PRI	Yes	No	No	No
PSI	No	Yes	No	Yes
P&PI	Yes	Yes	No	Yes

3.6.4.2 Effect type and vaccine efficacy

As described in the main text and model equations, prevention of infection vaccines reduced the infection parameter, and prevention of disease vaccine reduced the parameters for development of disease (proportion developing primary active disease, and risk of reactivation from latent and recovered classes). Modelled vaccines did not affect infectiousness of disease, severity of disease, case fatality ratio, likelihood of natural cure, or case detection.

As the research question aimed to explore the impact of a comprehensive set of vaccine efficacies, the full range of vaccine efficacies from 0% to 100%, in 10% steps, were explored for both prevention of infection and prevention of disease and in combination with each other.

3.6.4.3 Vaccine Duration and Waning of Protection

Modelled durations were 2, 3, 5, 7, 10, 15, 20, 25 years and lifelong. The minimum expected duration of protection anticipated from clinical trials is two years. Although it is unlikely that data on immune responses of any longer than five years would be known at registration, given BCG is known to protect for at least 10 years, potentially longer,[69] it is possible that these longer durations of protection would be possible with new vaccines. Vaccination is modelled as 'take', and waning assumed to be exact. No Immunosenescent waning is assumed.

3.6.4.4 Leaky versus all-or-nothing vaccine

The model that the basic structure was developed from modelled all-or-nothing prevention of disease vaccines,[1] meaning that the proportion moved to the vaccinated stratum was vaccine efficacy scaled by coverage ('effective coverage'), with the vaccinated stratum 100% protected against disease. The different assumption modeled here was of 'leaky' vaccine efficacy, in which the proportion moved to the vaccinated stratum equals coverage, and the vaccinated stratum receive partial protection against infection and/or disease.

Prevention of infection vaccine efficacy (*effI*) was applied to the infection parameter (λ) by multiplying it by (1-*effI*), and prevention of disease efficacy (*effD*) reduced the proportion developing primary active disease (*p*), and risk of reactivation from the latent (*v*) and recovered classes (*r*). In this leaky protection model, the vaccinated population can develop infection and disease, just with a lower risk than the unvaccinated population.

There were two main reasons for choosing leaky vaccine efficacy. Firstly, Ragonnet (2015) has demonstrated that models vaccinating with all-or-nothing vaccines tend to have greater impact than leaky vaccines.[70] Therefore, to ensure conservative estimates in this model, which is important when estimating the potential achievable impact of a vaccine, a leaky vaccine was more appropriate.[70] Secondly, from a technical perspective, all-or-nothing vaccination could not be used to model a vaccine with simultaneous efficacy against infection and disease, as is the research question here. To model these activities concurrently would require either 1) separate vaccine strata for POI and POD efficacy, which was considered an unrealistic scenario as it requires populations to experience either protection against infection *or* disease, or 2) the two efficacy types combined in to one vaccine stratum, but this would not be representative of both types of efficacy, which is integral to the research question and therefore required.

3.6.4.5 Vaccine implementation and coverage

Timing and ages for routine and mass vaccination are explained in detail in the main text. In brief, annual routine vaccination at 9 years was based upon using the HPV vaccination platform, mass vaccination of adolescents and adults was to ensure coverage of the epidemiological peak, and 10-yearly mass campaigns were considered the most feasible scenario by stakeholders.

Coverage of routine vaccination of 9 year olds was a conservative assumption (80%). This was based upon a combination of gross secondary school enrolment ratio from China (94.3%) taking in to account vaccination coverage of those attending school would not be 100%, and the HPV coverage achieved for 9 year olds in South Africa (87%).[30,31] Coverage of mass vaccination was modelled at 70%, based upon Menafrivac data and influenza vaccination in China.[34,35] Experience from serial vaccination suggests approximately a 10% drop in coverage between doses,[32] and Menafrivac campaigns in sub Saharan Africa achieved coverages of 70-98% of 1-29 year olds.[33] However, given the broader age range for vaccination in this model, and that routine elderly influenza vaccination in China only reaches coverages of 36-49%,[34,35] the lower end of the Menafrivac coverage was considered an appropriate assumption (70%).

3.6.5 Additional Methods: Selection of outcome measures

In recent modelling to inform TB antimicrobial TPPs, given there was an effective standard of care, the research aimed to focus on improving one or two priority characteristics that would provide the most impact.[71] Without an existing vaccine for adolescent/adult tuberculosis, such prioritization is of less relevance. Instead, based upon stakeholder consultation, my research aimed to more broadly inform targets for the overall efficacy and duration profile of the vaccine required to achieve given levels of population-level impact.

The primary outcome measure of impact selected was incidence rate reduction in 2050 compared to the no new vaccine baseline. To inform the TPP characteristics, estimates of the values of the duration and vaccine efficacies for prevention of infection and disease characteristics required to achieve pre-defined levels of impact were required. Given the interdependence of these three characteristics, two durations of protection were identified (see 3.3.5), and median values and ranges for the prevention of infection and disease characteristics required to achieve a given population-level reduction in TB incidence rate in 2050 were estimated. More importantly, the heatmaps demonstrating population level reduction in TB incidence rate compared to no new vaccine baseline in 2050 for all evaluated values of VE-POI, VE-POD and duration of protection outputs should be referred to in TPPs.

3.7 Chapter 3 References

1. Knight GM, Griffiths UK, Sumner T, et al. Impact and cost-effectiveness of new tuberculosis vaccines in low- and middle-income countries. *Proc Natl Acad Sci U S A* 2014; **111**(43): 15520-5.
2. World Health Organization. Global Tuberculosis Report 2016. 2016. http://apps.who.int/iris/bitstream/10665/250441/1/9789241565394-eng.pdf?ua=1 (accessed 10th May 2017).
3. Department of Economic and Social Affairs Population Division. World Population Ageing: 1950-2050 (ST/ESA/SER.A/207). 2001. http://www.un.org/esa/population/publications/worldageing19502050/ (accessed 22nd January 2017).
4. United Nations Population Division. World Population Prospects: The 2012 Revision, Highlights and Advance Tables. ESA/P/WP.228. 2012. https://esa.un.org/unpd/wpp/publications/Files/WPP2012_HIGHLIGHTS.pdf (accessed 23rd June 2014).
5. United Nations Population Division. World Population Prospects: The 2015 Revision, Methodology of the United Nations Population Estimates and Projections, Working Paper No. ESA/P/WP.242. 2015. https://esa.un.org/unpd/wpp/Publications/Files/WPP2015_Methodology.pdf2015).
6. Huynh GH, Klein DJ, Chin DP, et al. Tuberculosis control strategies to reach the 2035 global targets in China: the role of changing demographics and reactivation disease. *BMC medicine* 2015; **13**: 88.
7. Anhui Zhifei Longcom Biologic Pharmacy Co. Phase III Clinical Study of Efficacy and Safety of Vaccae™ to Prevent Tuberculosis. 27th December 2016. https://clinicaltrials.gov/show/NCT01979900 (accessed 3rd January 2017).
8. Harris RC, Sumner T, Knight GM, White RG. Systematic review of mathematical models exploring the epidemiological impact of future TB vaccines. *Hum Vaccin Immunother* 2016; **12**(11): 2813-32.
9. Aeras. The Aeras Annual Report 2015. 2015. http://www.aeras.org/annualreport2015 (accessed 20th January 2017).
10. Evans TG, Schrager L, Thole J. Status of vaccine research and development of vaccines for tuberculosis. *Vaccine* 2016; **34**(26): 2911-4.
11. R Core Team. R: A language and environment for statistical computing. . R Foundation for Statistical Computing, Vienna, Austria; 2014.
12. Ahmad D, Morgan WKC. How long are TB patients infectious? *CMAJ: Canadian Medical Association Journal* 2000; **163**(2): 157-.
13. Dye C, Garnett GP, Sleeman K, Williams BG. Prospects for worldwide tuberculosis control under the WHO DOTS strategy. *Lancet (London, England)* 1998; **352**(9144): 1886-91.
14. Lombardi G, Di Gregori V, Girometti N, Tadolini M, Bisognin F, Dal Monte P. Diagnosis of smear-negative tuberculosis is greatly improved by Xpert MTB/RIF. *PloS one* 2017; **12**(4): e0176186.

15. Read JM, Lessler J, Riley S, et al. Social mixing patterns in rural and urban areas of southern China. *Proceedings of the Royal Society B: Biological Sciences* 2014; **281**(1785).
16. WHO. TB treatment outcomes. 2015. http://www.who.int/tb/country/data/download/en/ (accessed 24th August 2015).
17. WHO. WHO TB burden estimates. 2015. http://www.who.int/tb/country/data/download/en/ (accessed 3rd July 2016).
18. United Nations Department of Economic and Social Affairs Population Division. World Population Prospects: The 2015 Revision, custom data acquired via website. 2015. https://esa.un.org/unpd/wpp/Download/Standard/Population/.
19. Wang L, Zhang H, Ruan Y, et al. Tuberculosis prevalence in China, 1990-2010; a longitudinal analysis of national survey data. *The Lancet* 2014; **383**(9934): 2057-64.
20. WHO. Global Tuberculosis Report 2013. 2013. http://www.who.int/tb/publications/global_report/en/.
21. Zhang H, Huang F, Chen W, et al. Estimates of tuberculosis mortality rates in China using the disease surveillance point system, 2004-2010. *Biomed Environ Sci* 2012; **25**(4): 483-8.
22. Gao L, Lu W, Bai L, et al. Latent tuberculosis infection in rural China: baseline results of a population-based, multicentre, prospective cohort study. *Lancet Infect Dis* 2015; **15**(3): 310-9.
23. Jabot F, Faure T, Dumoulin N, Albert C, Adapted by Funk S and Knight G. EasyABC (R package, adapted version). 2014.
24. Marjoram P, Molitor J, Plagnol V, Tavare S. Markov chain Monte Carlo without likelihoods. *Proc Natl Acad Sci U S A* 2003; **100**(26): 15324-8.
25. WHO. Human papillomavirus vaccines: WHO position paper, October 2014. *WHO Weekly Epidemiological Report* 2014; **89**(43): 465-92.
26. Lambert P-H. In: Haneknom W, editor.; 2016.
27. Vekemans J. In: Harris R, editor.; 2016.
28. Stakeholders from Aeras; BMGF, TBVI, and WHO. TB modelling stakeholder teleconference. 2016.
29. HPV information centre. China: Human Papillomavirus and Related Cancers, Fact Sheet 2016 (2016-12-15). 2016. http://www.hpvcentre.net/statistics/reports/CHN_FS.pdf (accessed 17th December 2016).
30. UNESCO Institute for Statistics. Education: gross enrolment ratio by level of education. 2016. http://data.uis.unesco.org/?queryid=142 (accessed 3rd January 2017).
31. HPV information centre. South Africa: Human Papillomavirus and Related Cancers, Fact Sheet 2016 (2016-12-15). 2016.
32. Widmeyer G. In: Hannekom W, editor.; 2016.
33. Harouna Djingarey M. Roll out of the meningococcal A conjugate vaccine through mass vaccination campaigns in countries of the African meningitis belt. 2014. http://www.who.int/immunization/sage/meetings/2014/october/2.DJINGAREY_Session6_SAGE_Oct2014_FINAL_21Oct2014.pdf?ua=1 (accessed 28th November 2016 2016).

34. Wu S, Yang P, Li H, Ma C, Zhang Y, Wang Q. Influenza vaccination coverage rates among adults before and after the 2009 influenza pandemic and the reasons for non-vaccination in Beijing, China: a cross-sectional study. *BMC Public Health* 2013; **13**: 636.
35. Zheng Y, Yang P, Wu S, et al. A cross-sectional study of factors associated with uptake of vaccination against influenza among older residents in the postpandemic season in Beijing, China. *BMJ Open* 2013; **3**(11).
36. Willem Hanekom. Teleconference. In: Harris R, editor.; 2016.
37. WHO and ERS. Framework towards TB elimination in low-incidence countries. 2014. http://www.who.int/tb/publications/Towards_TB_Eliminationfactsheet.pdf?ua=1 (accessed 12th January 2017).
38. Liu S, Li Y, Bi Y, Huang Q. Mixed vaccination strategy for the control of tuberculosis: A case study in China. *Math Biosci Eng* 2017; **14**(3): 695-708.
39. Arregui S, Sanz J, Marinova D, et al. A data-driven model for the assessment of age-dependent patterns of Tuberculosis burden and impact evaluation of novel vaccines. *bioRxiv* 2017: Online first.
40. Schenzle D. An age-structured model of pre- and post-vaccination measles transmission. *IMA J Math Appl Med Biol* 1984; **1**(2): 169-91.
41. Abu-Raddad LJ, Sabatelli L, Achterberg JT, et al. Epidemiological benefits of more-effective tuberculosis vaccines, drugs, and diagnostics. *Proceedings of the National Academy of Sciences* 2009; **106**(33): 13980-5.
42. Rajagopalan S. Tuberculosis and aging: a global health problem. *Clin Infect Dis* 2001; **33**(7): 1034-9.
43. Schaaf HS, Collins A, Bekker A, Davies PDO. Tuberculosis at extremes of age. *Respirology* 2010; **15**(5): 747-63.
44. Pathai S, Bajillan H, Landay AL, High KP. Is HIV a Model of Accelerated or Accentuated Aging? *The Journals of Gerontology: Series A* 2014; **69**(7): 833-42.
45. Yoshikawa TT. Tuberculosis in Aging Adults. *Journal of the American Geriatrics Society* 1992; **40**(2): 178-87.
46. Marion CR, High KP. Tuberculosis in Older Adults. In: Norman D, Yoshikawa T, eds. Infectious Disease in the Aging: A Clinical Handbook. Totowa, NJ: Humana Press; 2009: 97-110.
47. Rajagopalan S, Yoshikawa TT. Tuberculosis in the elderly. *Z Gerontol Geriatr* 2000; **33**(5): 374-80.
48. Lillebaek T, Dirksen A, Baess I, Strunge B, Thomsen VO, Andersen AB. Molecular evidence of endogenous reactivation of Mycobacterium tuberculosis after 33 years of latent infection. *The Journal of infectious diseases* 2002; **185**(3): 401-4.
49. Vynnycky E, Fine PE. Lifetime risks, incubation period, and serial interval of tuberculosis. *American journal of epidemiology* 2000; **152**(3): 247-63.
50. Vynnycky E, Nagelkerke N, Borgdorff MW, Van Soolingen D, Van Embden JDA, Fine PEM. The effect of age and study duration on the relationship between 'clustering' of DNA fingerprint patterns and the proportion of tuberculosis disease attributable to recent transmission. *Epidemiology and infection* 2001; **126**(1): 43-62.
51. National Bureau of Statistics of China. China Statistical Yearbook 2014, Population Birth rate, Death rate and Natural Growth rate of population. 2014. http://www.stats.gov.cn/tjsj/ndsj/2014/indexeh.htm (accessed 3rd July 2015).

52. Dye C, Williams BG. Eliminating human tuberculosis in the twenty-first century. *Journal of the Royal Society, Interface / the Royal Society* 2008; **5**(23): 653-62.
53. Vynnycky E, Fine PE. The natural history of tuberculosis: the implications of age-dependent risks of disease and the role of reinfection. *Epidemiology and infection* 1997; **119**(2): 183-201.
54. Gabriela MGM, Rodrigues P, Hilker FM, et al. Implications of partial immunity on the prospects for tuberculosis control by post-exposure interventions. *Journal of theoretical biology* 2007; **248**(4): 608-17.
55. 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. *Tubercle and Lung Disease* 1992; **73**(1): 52-8.
56. Ferebee SH. Controlled chemoprophylaxis trials in tuberculosis. A general review. *Bibliotheca tuberculosea* 1970; **26**: 28-106.
57. 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.
58. Gomes MG, Franco AO, Gomes MC, Medley GF. The reinfection threshold promotes variability in tuberculosis epidemiology and vaccine efficacy. *Proceedings Biological sciences / The Royal Society* 2004; **271**(1539): 617-23.
59. Dye C, Glaziou P, Floyd K, Raviglione M. Prospects for tuberculosis elimination. *Annu Rev Public Health* 2013; **34**: 271-86.
60. Baguelin M, Flasche S, Camacho A, Demiris N, Miller E, Edmunds WJ. Assessing Optimal Target Populations for Influenza Vaccination Programmes: An Evidence Synthesis and Modelling Study. *PLoS medicine* 2013; **10**(10): e1001527.
61. Cheng SM, Liu EY, Du X. [The impact of geriatric tuberculosis patients on the tuberculosis control strategy in China]. *Zhonghua liu xing bing xue za zhi = Zhonghua liuxingbingxue zazhi* 2004; **25**(8): 655-7.
62. Rao C, Yang G, Hu J, Ma J, Xia W, Lopez AD. Validation of cause-of-death statistics in urban China. *International journal of epidemiology* 2007; **36**(3): 642-51.
63. Wang WB, Zhao Q, Yuan ZA, Jiang WL, Liu ML, Xu B. Deaths of tuberculosis patients in urban China: a retrospective cohort study. *The International Journal of Tuberculosis and Lung Disease* 2013; **17**(4): 493-8.
64. Yang G, Rao C, Ma J, et al. Validation of verbal autopsy procedures for adult deaths in China. *International journal of epidemiology* 2006; **35**(3): 741-8.
65. Cheng J, Wang L, Zhang H, Xia Y. Diagnostic Value of Symptom Screening for Pulmonary Tuberculosis in China. *PloS one* 2015; **10**(5): e0127725.
66. Lin H, Wang L, Zhang H, Ruan Y, Chin DP, Dye C. Tuberculosis control in China: use of modelling to develop targets and policies. *Bulletin of the World Health Organization* 2015; **93**: 790-8.
67. Pai M, Denkinger CM, Kik SV, et al. Gamma interferon release assays for detection of Mycobacterium tuberculosis infection. *Clinical microbiology reviews* 2014; **27**(1): 3-20.
68. Gelman A, Shirley K. Handbook of Markov Chain Monte Carlo. In: Brooks S, Gelman A, Jones GL, Meng X, eds. Handbook of Markov Chain Monte Carlo: CRC Press; 2011.

69. Abubakar I, Pimpin L, Ariti C, et al. Systematic review and meta-analysis of the current evidence on the duration of protection by bacillus Calmette-Guerin vaccination against tuberculosis. *Health technology assessment (Winchester, England)* 2013; **17**(37): 1-372, v-vi.
70. Ragonnet R, Trauer JM, Denholm JT, Geard NL, Hellard M, McBryde ES. Vaccination Programs for Endemic Infections: Modelling Real versus Apparent Impacts of Vaccine and Infection Characteristics. *Scientific reports* 2015; **5**: 15468.
71. Dowdy DW. Driving Change to End TB, From new drugs to new regimens. Union world conference on lung health; 2016 25th October 2016; Liverpool, UK; 2016.

CHAPTER 4: Age targeting of new TB vaccines in China

Summary of Chapter 4

The objective of chapter 4 (*objective 3*) was to investigate the impact of age-targeted vaccination programmes to adolescents versus older adults on population-level TB epidemiology.

In this chapter, I further adapted the calibrated China model developed in chapter 3 to explore age targeting of new TB vaccines. This chapter reports the results of exploring the impact of vaccination of adolescents (15-19 years) versus older adults (60-64 years) with new TB vaccines. A subset of vaccines, selected based upon the results of the modelling in chapter 3, were explored in this chapter. The contributions of the authors to this work are described in the paper cover sheet. This work is reported as ***research paper 2***, currently under review at the Lancet Global Health.

Citation (paper 2): Harris RC, Sumner T, Knight GM, Evans T, Cardenas V, Chen C, and White RG. New tuberculosis vaccines – the impact of age targeted vaccination in China and implications for vaccine development. *Under review*

Earlier iterations of this work were submitted to Aeras as a report, and were presented at the TB Vaccine 4th Global Forum (21st April 2015) and the Union Conference for Lung health (6th December 2015).

CHAPTER 4 Age targeting of new TB vaccines in China

4.1 Introduction

4.1.1 Background

Modelling studies have demonstrated that routine vaccination of adolescents and/or periodic mass vaccination of adults provides much greater impact more rapidly than routine vaccination of neonates.[2-5] Adolescent/adult vaccination has also been demonstrated to be more cost effective, as vaccines with as low as 20% efficacy and 5 years duration of protection demonstrated a cost effective price, whereas neonatal vaccines were only found to have a cost effective price at much higher efficacies and durations.[2]

Accordingly, vaccine implementation in Chapter 3 was modelled as routine vaccination of 9 year olds alongside HPV vaccination, with periodic mass vaccination of adolescents/adults. This takes advantage of a likely future vaccination platform, and helps maximise absolute impact through mass vaccinations. However, the resources required to conduct such mass campaigns, especially the more frequent scenarios, would be substantial. Therefore, age-targeting of routine vaccination could help ensure a balance between maximisation of vaccine impact and efficient use of resources.

In Chapter 3, the baseline epidemiological model demonstrated a very clear shift in the age distribution of TB in China between 2000 and 2050. In 2000, most cases occurred in the 15-64 year old age group (75.8%; UR: 71.3-81.0); whereas by 2050, 74.5% (UR: 70.2-78.6) were estimated to occur in the elderly (≥65 years old). This was concurrent with a transition from a new-infection-driven to a reactivation-driven epidemic. This is indicative of the potential importance of older adult age groups in efforts to maximise future impact of new TB vaccines.

Given these results, a clear research need was identified to investigate whether the current focus on targeting routine vaccination to adolescents or an possible

alternative targeting strategy focused on older adults could potentially help maximise impact. No existing research explores the impact of vaccinating older adults or the elderly against TB, therefore this is the first of its kind for TB vaccines.[6]

Older adults are a population almost entirely excluded from current clinical trials. The M72-AS01E vaccine candidate in phase IIB trials recruits only 18-50 year olds,[7] and the DAR-901 candidate is currently in trials in 13-15 year olds.[8] The only candidate recruiting up to 64 year olds is the *M.vaccae* candidate in phase III trials in China,[9] and there are no candidates recruiting the elderly. Older age groups tend to be excluded due to uncertainty with regards to safety and immunogenicity. However, the robust efficacy achieved zoster vaccines in older adults and the elderly provides precedent for considering TB vaccine development in older age groups.[10,11] Therefore, results of this study are of importance for informing recruitment populations for future trials and, if found to be efficacious in these age groups, implementation strategies once a candidate is registered.

In this chapter, I further developed the model reported in Chapter 3, to explore the relative impact of vaccinating adolescents versus older adults in China.

4.1.2 Aims and objectives

This Chapter contributes towards the first aim of the thesis and objective 3:

Aim 1: Generate mathematical models of age-stratified demography and TB epidemiology to explore the population-level epidemiological impact of vaccine characteristics and implementation strategies for potential new TB vaccines, using China as a case study.

Objective 3: Using the calibrated model of the TB epidemic in China and main vaccine characteristics (objective 2), investigate the impact of age-targeted vaccination programmes to adolescents versus older adults on population-level TB epidemiology.

4.1.3 Overview of modelled vaccine characteristics

Detailed methods are presented in the relevant sections of the manuscript and supplementary materials. A brief summary follows of the key vaccine characteristics chosen based upon the results of Chapter 3, and the main differences compared to Chapter 3 (Table 4.1).

To explore age targeting, in the following research adolescent vaccines were routinely delivered to 15 year olds, and older adult vaccines to 60 year olds, with a one-off 3-year mass campaign of ages 16-19 and 61-64 year olds. Selection of these vaccination ages took into consideration the high burden age groups observed in Chapter 3, logistical feasibility of delivery, the adolescent age group of current strategic focus, and safety considerations for the older age group. To explore the impact of implementation success, both 30% and 70% vaccination coverage was explored.

In Chapter 3 it was demonstrated that prevention of disease vaccines will be needed to maximise the future impact of new TB vaccines in China. Therefore, in this chapter the model was adapted to focus on prevention of disease vaccination as 'all-or-nothing' protection.

Vaccines effective post-infection were shown to deliver greatest impact in China over the time horizon explored, but given that the prevalence of infection was estimated to be substantially higher in older age groups, in this chapter I continue to model pre-infection, post infection and pre- and post-infection vaccines, so as not to bias the results towards older adult vaccination. In Chapter 3, post-infection vaccines were assumed to provide efficacy both in latency and in populations recovered from disease. However, it is possible that post-infection vaccines would be registered for use only in latency, so the relative impact of these two types of post-infection vaccine was explored in this study.

Based upon results from Chapter 3, I decided to explore a subset of vaccine efficacies of 40%, 60% and 80%. These were anticipated to be technically feasible, and in Chapter 3 were estimated to achieve 40-49% up to 60-69% incidence rate reduction in 2050 if delivered as a P&PI POD vaccine with 10 years protection and 10-yearly mass vaccination campaigns.

In this chapter, duration of protection was limited to 10 and 20 years. Based upon expert opinion, these were considered a likely and an optimistic scenario for the duration of protection, and allowed exploration of the impact of targeted vaccination. In Chapter 3, waning occurred exactly at the duration of protection for all individuals and was not altered by the age of the vaccinated individual. This was improved to more closely reflect biological variation in this chapter, by modelling waning at the population level as normally distributed around the mean duration of protection, with a standard deviation of 10% of the mean. Secondly, in older age groups it is biologically plausible that immunosenescence could lead to more rapid waning of protection. To avoid bias towards older adult vaccination, immunosenescent waning was represented by accelerated waning in the population aged ≥65 years. In the main analysis this was assumed at 2% per year, but was explored for 0-5% in sensitivity analyses.

A summary of the comparisons between the vaccines modelled in chapters 3 and 4 is provided in (Table 4.1), and further detail is provided in the manuscript in section 4.2.

Table 4.1: Summary of differences in vaccine characteristics between the model developed in Chapter 3 compared to Chapter 4.

Vaccine characteristic	Vaccines modelled in Chapter 3	Vaccines modelled in this chapter (Chapter 4)
Routine vaccination age	9 year olds	15 or 60 year olds
Mass vaccination age	≥10 years old	16-19 or 61-64 year olds
Mass frequency	Every 10 years (or duraton of protection if longer), starting in 2025	One-off, during 2025-2027
Coverage routine/mass	80% / 70%	30% or 70% / 30% or 70%
Effect type	POI&D	POD-only
Host infection status required for efficacy	P&PI, PRI, PSI-L&R	P&PI, PRI, PSI-L&R, PSI-L
Vaccine efficacy	0-100%	40-80%
Take or degree efficacy	Leaky (‘degree’)	All-or-nothing (‘take’)
Duration of protection	2 years to lifelong	10 or 20 years
Waning	All wane at exactly duration of protection.	Normally distributed around mean (s.d. 1 or 2 years). Elderly immunoscenescent waning 2%.

4.2 Research Paper 2: New tuberculosis vaccines – the impact of age targeted vaccination in China and implications for vaccine development

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	Rebecca Claire Harris
Principal Supervisor	Richard White
Thesis Title	New tuberculosis vaccines – the impact of age targeted vaccination in China and implications for vaccine development

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?	Submitted to the Lancet Global Health
Please list the paper's authors in the intended authorship order:	Rebecca C. Harris, Tom Sumner, Gwenan M. Knight, Tom Evans, Vicky Cardenas, Chen Chen, and Richard G White.
Stage of publication	**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)	See next page

Student Signature: [redacted] **Date:** 29/9/17

Supervisor Signature: [redacted] **Date:** 29/9/17

Paper 2: New tuberculosis vaccines – the impact of age targeted vaccination in China and implications for vaccine development

Authors: Rebecca C. Harris, Tom Sumner, Gwenan M. Knight, Tom Evans, Vicky Cardenas, Chen Chen, and Richard G White.

Author contributions: All authors contributed to development of the research question. The research methodology was developed by myself with input from Prof. Richard White, Dr Tom Sumner and Dr Gwenan Knight. I collected the data for parameterisation and calibration of the model from secondary sources. The basic structure of the model was based upon a model developed by Dr Gwenan Knight, but I substantially modified the model with guidance from Prof. Richard White, Dr Tom Sumner and Dr Gwenan Knight, including development of age structure and parameterisation, calibration to age stratified epidemiological (prevalence, notifications, mortality and incidence) and demographic calibration targets, age targeted vaccination and immunosenescence, incorporation of social mixing patterns, and development of model outcomes (e.g. incidence rate reduction, NNV) amongst other developments. I carried out the model calibration, with guidance from Dr Tom Sumner. Interpretation of results were my own and I wrote the first draft of the manuscript and incorporated co-author feedback manuscript. All authors reviewed manuscript drafts and approved the final manuscript.

New tuberculosis vaccines – the impact of age targeted vaccination in China and implications for vaccine development

Authors: Rebecca C. Harris,[a,*] Tom Sumner,[a] Gwenan M. Knight,[b] Tom Evans,[c,d] Vicky Cardenas,[c,e] Chen Chen,[f,g] and Richard G White.[a]

[a] TB Modelling Group, TB Centre and CMMID, London School of Hygiene & Tropical Medicine, Keppel Street, London, WC1E 7HT, UK

[b] Imperial College London, UK

[c] Aeras, 1405 Research Blvd, Rockville, MD 20850, USA

[d] Current affiliation: Vaccitech Limited, Park End Street, Oxford, OX1 1JD

[e] Current affiliation: The Aurum Institute, 29 Queens Road, Johannesburg, South Africa

[f] Aeras Asia, Suite 1426, Tower A, Pacific Century Place, Chaoyang District, Beijing 100027, P.R. China

[g] Current affiliation: Division of Clinical Epidemiology and Aging Research, DKFZ, Heidelberg, Germany

Keywords: tuberculosis, vaccine, infectious disease epidemiology, mathematical model, theoretical models, adolescent, older adult, elderly, China

Abstract

Background: Tuberculosis is the leading single-pathogen cause of death worldwide, with China the third largest contributor to global tuberculosis incidence. New tools, such as new vaccines, are needed to meet WHO tuberculosis goals. Current tuberculosis vaccine development strategies mostly target infants or adolescents, but given China's ageing epidemic, vaccinating older ages may be important. We explored the potential impact of new tuberculosis vaccines in China, targeted to adolescents (15-19 years) or older adults (60-64 years) and varying vaccine characteristics, to inform strategic vaccine development.

Methods: A *Mycobacterium tuberculosis* transmission model was calibrated to age-stratified demographic and epidemiological data from China. Vaccine implementation (age targeting and coverage) and characteristics (efficacy, duration of protection, and host infection status required for efficacy) were explored. Outcomes included incidence and mortality rate reduction in 2050 compared to no new vaccine, and number needed to vaccinate per case averted.

Findings: By 2050, results suggest that 74·5% (uncertainty range (UR): 70·2-78·6) of incident tuberculosis cases in China would occur in the elderly (≥65 years), and 75·1% (UR: 66·8-80·7) would be due to reactivation, rather than recent transmission. The vaccine profiles explored averted up to 3·0million cases over 2025-2050. All vaccine profiles delivered to older adults had higher population-level impact, and lower number needed to vaccinate per case or death averted, than if delivered to adolescents. Vaccinating older adults with post-infection vaccines provided substantially greater impact than pre-infection vaccines. Results were robust to sensitivity analysis.

Interpretation: Adolescent-targeted tuberculosis vaccines, the focus of many current development plans, would have low impact in ageing, reactivation-driven epidemics like China. Instead, an efficacious post-infection vaccine delivered to older adults will be critical to maximise population-level impact in this setting, and would provide a

crucial contribution towards achieving WHO goals. Older adults should be included in tuberculosis vaccine clinical development and implementation planning.

Funding: Aeras and UK MRC

Research in context panel

Evidence before this study: A recently published systematic review of the tuberculosis (TB) vaccine epidemiological modelling literature (Harris et al, 2016) was updated for articles published up to 19th July 2017. The published literature suggests that targeting new vaccines to adolescents would have greater impact than infant vaccination, however no previous research has explored the potential impact of new TB vaccines targeted to older adults or the elderly. Two studies exploring the impact of new TB vaccines in China compared vaccination at birth to mass vaccination, and suggested that post-infection mass vaccination may be important for elimination in China. However, critically, age specificity in demographic and epidemic dynamics in China has not been addressed in vaccine models.

Added value of this study: Our research is the first to explore the potential impact of new TB vaccines targeted to older adults, and to compare this approach to the current strategic focus on adolescent vaccination. This data-driven model is calibrated to age-stratified data from the WHO, UN population division, and empirical studies for demographics, prevalence, mortality, notification and incidence rates, including capturing temporal trends, and incorporates empirical data on non-random age mixing. This study is the first to show that older adult vaccination with new TB vaccines is likely to be critical to maximising impact in settings such as China, and strengthens the conclusions from previous studies suggesting that vaccines effective post-infection will be needed in China.

Implications of all the available evidence: Although adolescent vaccination may be a suitable global-level strategy, important differences in country-level epidemiology and demography may cause adolescent-targeted TB vaccines to have low impact in ageing, reactivation-driven TB epidemics such as in China. In these settings, an efficacious post-infection vaccine delivered to older adults will be critical to maximise population-level impact. Older adults should be included in TB vaccine clinical development and implementation planning.

Background

Despite progress in tuberculosis (TB) prevention and care during the last two decades,[12] China remains the third largest contributor to global TB incidence, accounting for 918,000 (788,000-1,060,000) new cases in 2015.[13] Increasing public sector investment in TB care has contributed to case detection rates reaching 87% and treatment success of 95%.[12,13] Given the substantial achievements in scale-up of existing TB care and prevention options, mathematical modelling suggests that gains possible from further investment in existing tools alone are unlikely to reach the World Health Organization (WHO) 2025, 2035 and 2050 TB goals in China.[14-18] Innovative tools, such as new TB vaccines, have been shown to be essential to achieving these targets.[14-17]

Recent years have seen increasingly rapid advances in the development of new TB vaccines, with thirteen candidates in clinical trials, including four in phase IIB/III.[19] TB vaccine target product profiles (TPPs) outline the desired vaccine characteristics and recipient populations for new vaccines, and clinical development plans provide the pathway to achieving these criteria. Mathematical modelling provides a systematic framework to explore the potential future impact of vaccine characteristics and implementation strategies to inform TPP development. Impact achieved varies by epidemiological and demographic setting,[5,6] therefore country-specific models are needed for appropriate vaccine selection and implementation planning such as age targeting. A prophylactic TB vaccine candidate is in Phase III clinical trials in China at time of writing,[9] therefore modelling to elucidate the potential impact of new TB vaccines and implementation strategies in China is urgently needed.

Traditional vaccine development pathways focussed on infant vaccination, though recently the field has experienced a paradigm shift towards adolescents and adults,[20] supported by insights from mathematical modelling.[2,5] However, older adult (60-64 years) and elderly (≥65 years) populations remain neglected in TB vaccine development pathways, despite several developed and developing countries experiencing high TB disease burden in older age groups.[21] Given the dramatic and

on-going population ageing in China,[22] and elevated prevalence of latent infection in older populations,[1] a shift in burden to older age groups could be expected.

The modelling literature for TB vaccines has recently been systematically reviewed.[6] Although studies have explored vaccination of infants compared to adolescents or all ages, no model explores the potential impact of new TB vaccines for older adults or the elderly.[6] Dynamics of TB disease and demographics are intimately linked, and influence the impact of age-targeted interventions.[5] Only two studies have modelled new TB vaccines in China,[4,23] and neither accounted for age specificities in epidemiology or demographics. Further, in these studies, vaccination was either at birth or all-ages, and vaccine characteristics modelled were unclear or unrealistic (e.g. 100% vaccine efficacy).[4,23] Although not exploring vaccines, a modelling study in China showed that existing TB interventions alone would not reach WHO 2035 targets, whereas development of preventative therapy for the elderly could be "transformational".[16]

In this paper, we aimed to estimate future trends in TB epidemiology, and the impact of new TB vaccines in China by varying implementation strategies (age-targeting and coverage) and vaccine characteristics (efficacy, duration of protection, and host infection status required for efficacy).

Methods

Model structure and calibration

An age-stratified population-level compartmental deterministic transmission model calibrated to China's TB epidemic was developed in R,[24] based upon the model developed by Knight et al. (2014).[2] A summary follows. See supplementary materials for full details.

The model represents five infection states: uninfected, latent infection, bacteriologically-positive active TB disease, bacteriologically-negative active TB disease, and recovered from active TB disease populations. Transitions between states represent acquisition of infection, development of primary, reactivation or relapse disease, effective detection and treatment, or natural cure (Supplementary Figure S1). Births enter the uninfected state, all-cause mortality occurs in all states, and TB mortality occurs in active disease states. Uninfected, latent and recovered states comprised both unvaccinated and vaccinated strata. Age was modelled in single years from 0-100. Main age categories for parameters and outcomes were children (0-14 years), adolescents and adults (15-64 years), and the elderly (≥65 years). Age categories for vaccine implementation were adolescents (15-19 years) or older adults (60-64 years).

Age-stratified natural history parameter prior ranges, summarised in Supplementary Table S1, were identified based upon available data. HIV has been identified as a proxy for elderly immunosenescence,[25] therefore for parameters without available elderly data, priors were based upon data from HIV populations. Temporal evolution of the age distribution of infection and disease was captured historically and prospectively by calibration to age-stratified epidemiological and demographic data using age-specific natural history parameters and contact patterns. To provide a more gradual older-age transition in reactivation, relapse, natural cure and case detection parameters, 55-64 year old parameters were estimated as the mean of the adult and elderly calibrated values. Heterogeneity in age-wise contact patterns were based upon data from a study in Southern China.[26] Treatment success was based upon historical data, and assumed to plateau beyond 2011 due to high success rates

achieved (95%).[27] Case detection followed a generalised logistic function based upon WHO 1990-2010 data.[28] BCG delivery was assumed to remain constant, so was not explicitly modelled.

A two-stage calibration process was employed, with the first stage a manual calibration to UN 2010 and 2050 age-stratified demographic data and projections,[29] and the second stage an adaptive rejection Approximate Bayesian Computation-Markov Chain Monte Carlo (ABC-MCMC)-based method using a modified *easyABC* package to calibrate to historical epidemiological data (see Supplementary Section 4.3.3 for details).[30,31] To provide the temporal and age-stratified granularity required, the model was calibrated to 18 epidemiological data ranges: China's bacteriologically-positive prevalence rates (≥15, 15-29, 30-59 and ≥60 years) in 2000 and 2010,[12] 2010 notification rates (all-age, 0-14, 15-54, 55-64, and ≥65 years),[32] 2010 mortality rates (all-age, 0-14, 15-59 and ≥60 years),[33] and 2010 all-age incidence rate.[28] After ranking by likelihood, the 20 highest from one million parameter sets randomly selected from uniform priors were used as seeds for the adaptive rejection ABC-MCMC. Sequential parallel ABC-MCMC chains were employed, seeding with acceptances from the previous chains and adapting the rejection criterion until a full model fit to all 18 data ranges was achieved. Results and uncertainty ranges were based on one thousand parameter sets randomly selected from the acceptances in the final set of ABC-MCMC runs.

Epidemiological outcomes

Epidemiological outcomes were calculated annually for 2000-2050 in the baseline ('no new vaccine') scenario. Contribution by age group to annual incident cases was calculated as the proportion of total incident cases arising from a given age group. The proportion of incident disease due to new transmission versus reactivation was estimated annually. The population attributable fraction (PAF) of each age group to annual *M.tb* infections was estimated using methods described previously.[34] Prevalence of latent infection was estimated by age, for comparison to empirical data.[1]

Vaccine characteristics and implementation

Two implementation characteristics (age targeting and coverage) and three vaccine characteristics (vaccine efficacy, host infection status required for efficacy, and duration of protection) were varied, exploring 96 vaccination scenarios in total.

Two age targeting scenarios were explored: vaccination of adolescents (15-19 year olds) and of older adults (60-64 year olds). Vaccination was implemented 2025-2050. Routine vaccination was provided annually throughout this period to 15 or 60 year olds. Initial catch-up campaigns were performed in 2025-2027 for 16-19 year olds or 61-64 year olds.

Routine vaccine coverage of 30% and 70% of the target populations was explored. Only those with active disease were excluded from vaccination. Annual coverage for catch-up campaigns was assumed to be a third of the routine coverage. We assumed no latent infection screening prior to vaccination, but each vaccine was only effective in the specified host infection status groups. Therefore, the proportion effectively immunised in a given year was the product of coverage in hosts with the specified infection status, and vaccine efficacy.

Vaccine efficacy (VE) was modelled as 40%, 60% and 80% protective against development of TB disease ('prevention of disease'). Vaccination was modelled as 'take' (all-or-nothing) protection. Vaccine protection in four host infection statuses was explored: 1) uninfected individuals (pre-infection: PRI), 2) latently infected individuals (post-infection, latency-only: PSI-L), 3) individuals either latently infected or recovered from active disease (post-infection: PSI-L&R), or 4) individuals either uninfected, latently infected, or recovered from active disease (pre- and post-infection: P&PI).

The rate of vaccine waning was set so that the duration of vaccine protection was normally distributed with a mean of 10 or 20 years and a standard deviation of 1 or 2 years, respectively. Waning returned the population to the equivalent unvaccinated

state. To account for immunosenescence, an additional 2% annual waning was assumed for elderly populations, and was varied between 0-5% in sensitivity analyses.

All combinations of these characteristics and implementation strategies were explored. For reporting of some outcomes, an 'intermediate' vaccination scenario was defined as 60% VE, 10-year protection and 70% coverage.

Vaccine impact outcomes

Primary outcomes were the percentage incidence rate reduction and mortality rate reduction in each vaccination scenario compared to the no new vaccine baseline in 2050. Secondary outcomes were cumulative number of TB cases or deaths averted 2025-2050 compared to baseline, and the cumulative number needed to vaccinate (NNV) per case or death averted 2025-2050. These outcomes were compared by age targeting and other vaccine characteristics.

Role of the funding source

The study sponsor (Aeras), a non-profit Product Development Partnership, contributed to development of the research question, and commented on manuscript drafts.

Results

Baseline (no new vaccine) scenario

The model fitted overall and age-stratified data for demography and TB prevalence, notification, incidence and mortality rates (Figure 1A and Supplementary Figures S3-8).

Modelled all-age infection prevalence was 15·7% (UR: 12·8-18·7) in 2013, increasing almost linearly from 1·3% (UR: 1·0-1·6) in 5-9 year olds, to 37·0% (UR: 31·1-42·6) in ≥70 year olds (Supplementary Figure S10). Although not calibrated to these data, modelled age-stratified infection prevalence aligned closely with published Quantiferon data.[1] At vaccine introduction, 2·2% (UR: 1·7-2·8) of 15-19 year olds and 25·3% (UR: 20·5-30·2) of 60-64 year olds were latently infected (Supplementary Figure S11).

Between 2000 and 2050, a substantial shift was projected in the age distribution of incident TB cases, from 75·8% (UR: 71·3-81·0) in adolescents and adults (15-64 years) in 2000, to 74·5% (UR: 70·2-78·6) in the elderly (≥65 years) in 2050 (Figure 1B). A concurrent age shift was observed in the source of new infections (Figure 1C and Supplementary Section 4.3.5.1). Over the same period, the model predicted a transition from a recent-infection-driven, to a reactivation-driven epidemic. Between 2000 and 2025, the estimated proportion of incident disease due to reactivation rose from 28·6% (UR: 21·9-32·5) to 60·1% (UR: 51·1-64·9). By 2050, 75·1% (UR: 66·8-80·7) of all incident disease cases were estimated to derive from reactivation (Figure 1D), though the proportion due to reactivation was lower in 15-19 year olds (6·1%; 4·8-7·8%) than in 60-64 year olds (78·1%; UR: 71·8-81·8%).

In the baseline scenario, the all-age TB disease incidence rate was projected to decline from 56·2/100,000 population/year (UR: 51·4-60·0) in 2025 to 33·7/100,000 population/year (UR: 27·6-38·8) in 2050, and the mortality rate from 2·0/100,000 population/year (UR: 1·2-3·4) to 1·1/100,000 population/year (UR: 0·6-2·2) (Supplementary Figures S5-6).

Epidemiological impact of new TB vaccines

Vaccine impact, in terms of incidence rate reduction in 2050 compared to the no new vaccine baseline in 2050, varied by age targeting, host infection status required for efficacy, vaccine efficacy, coverage and duration of protection (Figure 2).

All vaccine profiles explored provided greater population-level impact when delivered to older adults than adolescents (Figure 2). The relative differences were demonstrated by the 'intermediate' vaccine profile (Table 1), where the median population level incidence rate reduction by older adult vaccination ranged from 1.9- to 157.5-times greater than observed with adolescent vaccination, depending on host infection status required for efficacy. For the same intermediate profiles, the estimated median cumulative number needed to vaccinate (NNV) per case averted 2025-2050 for older adult vaccination were 0.011 to 0.796 times the NNV estimates for adolescent vaccination.

The relative impact of pre- versus post-infection vaccines varied by vaccination age (Figure 2). For older adult vaccination, the predicted impact, from lowest to highest, was: vaccines effective pre-infection (PRI), post-infection in latency (PSI-L), post infection in individuals either latently infected or recovered from active disease (PSI-L&R), and pre- and post-infection (P&PI). Vaccination of older adults with post-infection vaccines provided greater impact and required vaccination of fewer individuals per case averted than pre-infection vaccines (Figure 2 and Supplementary Table S4). For adolescent vaccination, although greatest impact was also with P&PI vaccines, pre-infection vaccines were predicted to have greater impact than the two post-infection (PSI-L and PSI-L&R) vaccines.

The relative incidence rate reduction by older adult compared to adolescent vaccination was smallest with pre-infection vaccines ('intermediate' vaccine older adult:adolescent ratio 1·9; UR: 1·5-2·6), and greatest with post-infection (PSI-L&R) vaccines (ratio 157·5; UR: 119·3-225·6).

Inducing vaccine protection in both latent and recovered populations (PSI-L&R) provided markedly greater impact than protection only in latency (PSI-L) (Figure 2 and Supplementary Figure 14). This was especially important in older adult populations, where latency-only vaccines (PSI-L) provided 44% (UR: 20-89%) less impact than PSI-L&R vaccines (Figure 2).

As expected, increased vaccine efficacy, duration of protection or coverage led to greater incidence rate reductions (Figure 2). The cumulative NNV per case or death averted 2025-2050 declined with increasing vaccine efficacy and duration of protection, as additional benefit was gained without delivery of additional vaccines (Supplementary Tables S7-8). Although raising coverage substantially increased cases averted, many more vaccine doses were required, so cumulative NNV was found to increase slightly at higher coverages. The same patterns of vaccine impact were observed for mortality as for incidence.

Of the vaccine profiles and implementation strategies explored, the most effective was the vaccine effective P&PI with 80% VE, 20-year protection and 70% coverage. Delivered to older adults this vaccine averted 3·0 million (UR: 2·5-3·5m) cases 2025-2050, whereas when delivered to adolescents only 502,000 (UR: 431,000-591,000) cases were averted (Figure 2 and Supplementary Table S4). The lowest impact vaccines explored averted as few as 2,000 cases during 2025-2050.

Immunosenescent waning assumptions had minimal impact on the primary outcome. Increasing annual immunosenescent waning from 2% to 5% reduced the "intermediate" vaccine incidence rate reduction from 13·8% to 12·9%. Conversely, assuming no immunosenescence increased this estimate to 14·4%.

Discussion

During 2025-2050, all vaccine profiles explored in China provided higher population-level impact on cases or deaths, and lower cumulative number needed to vaccinate per case or death averted, when delivered to older adults (60-64 years) than adolescents (15-19 years). With the most effective vaccine profile and coverage explored, up to 3·0 million cases were averted between 2025-2050. By 2050, the no new vaccine baseline results suggest that around 75·1% of incident TB cases would be due to reactivation of existing infections rather than recent transmission, and 74·5% of incident cases would occur in the elderly (≥65 years). Vaccination of older adults with post-infection vaccines provided much greater impact than pre-infection vaccines. Adolescent vaccination was found to provide limited impact on China's TB epidemic 2025-2050.

In the no new vaccine baseline, transition from an adult, transmission-driven epidemic to an elderly, reactivation-driven epidemic was predicted to occur before a new vaccine is likely to be launched, as a result of population ageing concurrent with reduced transmission due to improved TB diagnosis, care and prevention.

Even though immunosenescence disadvantaged older adult vaccination, targeting older adults consistently provided greater epidemiological impact than targeting adolescents, regardless of the vaccine characteristics explored. Although older adult campaigns delivered more vaccines than adolescent campaigns for the same coverage due to the larger population size, number needed to vaccinate per case averted was lower, often considerably, than for adolescent vaccination, suggesting cost effectiveness calculations may also support older adult vaccination strategies. Even if lower coverage was assumed for older adult campaigns (30%) than school-based adolescent campaigns (70%), vaccines targeted to older adults would still be likely to provide at least equivalent, if not greater, impact than adolescent vaccination.

The relative impact of pre-infection versus post-infection vaccines varied in adolescent and older adult populations. Vaccination of older adults achieved higher population-level impact with post-infection vaccines, whereas higher impact was achieved in adolescents by pre-infection vaccines; explained by the substantially higher prevalence of latent infection in older adults. Pre-infection vaccines provided similarly low population-level impact whether vaccinating adolescents or older adults. Post-infection vaccines delivered to older adults will be central to achieving the highest possible impact. However, post-infection vaccines would have negligible impact if delivered to adolescents. Post-infection vaccines effective in both latent and recovered populations provided greater impact than latency-only, particularly when delivered to older adults, indicating vaccines also effective against relapse after recovery from active disease could have important additional value. Vaccines effective both pre-and post-infection provided greatest impact in both age groups, and would be the ideal candidate for development, but at a minimum a post-infection vaccine suitable for older adult vaccination should be developed for TB prevention in China if highest population level impact before 2050 is the primary goal.

Of the vaccine profiles and delivery strategies explored, an up to 32% reduction in TB incidence rate could be achieved in 2050, which would be a substantial contribution towards WHO 2050 goals. Even greater impact could be achieved with less conservative delivery strategies (e.g. one off mass vaccination to all adults), or if higher efficacy or duration of protection were achieved. Our model demonstrates that an appropriately designed and delivered new TB vaccine could provide a substantial contribution to TB prevention in China.

Limitations

There is a notable data gap with regards to elderly TB natural history parameters. We represented this uncertainty by sampling from elderly parameter priors spanning HIV-negative and HIV-positive adult ranges in the calibration process. Given population ageing, the assumption that the contact matrix remained constant throughout the study period may provide conservative impact estimates. The vaccine was assumed to provide “all-or-nothing” protection, but the alternative “leaky”

assumption could reduce impact estimates.[35] Due to high current treatment success and case detection estimates, these were assumed to plateau. If improvements in current care measures were achieved, or better diagnostics or treatments introduced, the impact achieved by vaccines could potentially be reduced. However, this would not affect conclusions with respect to age targeting unless such an intervention varied in effectiveness by age.

Impact was evaluated up to 2050 to align with the WHO 2050 elimination goal, so results are specific to 2025-2050. Full benefit of vaccines delivered pre-2050 would extend beyond 2050, therefore relative vaccine impact may change over longer horizons.[6] Better vaccines (e.g. lifelong protection) or coverage could improve impact,[23] but were not considered likely scenarios by experts. Infant vaccination could be explored, but impact would be very low over this time horizon given the declining force of infection in China.[2,5] To answer the research question, we deliberately restricted mass campaigns to narrow age ranges (15-19 and 60-64 years). If a new efficacious vaccine became available, broader campaigns could be implemented to increase overall impact, but would not affect our overall conclusions.

Value of this research

This research is the first to explore the potential impact of new TB vaccines targeted to older age groups, and to provide a comparison of this approach to the current strategic focus on adolescent vaccination.[20] This required age-specific model parameterization and calibration, and is the first China TB model to calibrate to age-stratified mortality and notification rates and include data-informed non-random mixing by age. Our results are consistent with the conclusions of the Dye study regarding the importance of post-infection vaccines in China,[23] but provide a critical clarification: that population-level impact with such vaccines is contingent on delivery to older age groups. These results build on the Huynh study demonstrating the importance of controlling elderly reactivation disease.[16] Our results extend previous work which suggested that adolescent/adult campaigns would have a greater impact than infant vaccination,[2,4,5] as we develop the argument further by demonstrating

that in ageing epidemics, most impact could be achieved by targeting older adults, potentially minimising the resources required per case averted.

Policy impact

Adolescent-targeted vaccines, the focus of current development plans, would have low impact in ageing, reactivation-driven TB epidemics like China. In these settings, an efficacious post-infection vaccine delivered to older adults will be critical to maximise population-level impact and would provide a crucial contribution towards achieving the WHO 2050 TB goals. Older adults should be included in TB vaccine clinical development and implementation planning.

Declarations

List of abbreviations

ABC-MCMC	Approximate Bayesian Computation – Markov Chain Monte Carlo
BCG	Bacillus Calmette-Guérin vaccine
NNV	Cumulative number needed to vaccinate
PAF	Population attributable fraction
P&PI	Pre- and post-infection
PRI	Pre-infection
PSI	Post-infection
PSI-L	Post-infection effective in latency only
PSI-L&R	Post-infection effective in individuals latently infected or recovered from active disease
TB	Tuberculosis
TPP	Target product profile
UR	Uncertainty range
VE	Vaccine efficacy
WHO	World Health Organization

Declaration of interests

RGW, TS, and GK report no potential competing interests. RCH provided consultancy to GSK vaccines, ending in 2015 and outside the scope of the submitted work. TE and VC were employed at the TB vaccine development organization Aeras (Rockville, USA), and CC was employed as Aeras Asia (Beijing, China), during the period when this work was completed.

References

1. Wang L, Zhang H, Ruan Y, et al. Tuberculosis prevalence in China, 1990-2010; a longitudinal analysis of national survey data. *The Lancet* 2014; **383**(9934): 2057-64.
2. World Health Organization. Global Tuberculosis Report 2016. 2016. http://www.who.int/tb/publications/global_report/en/ (accessed 21st December 2016).
3. Lin H, Wang L, Zhang H, Ruan Y, Chin DP, Dye C. Tuberculosis control in China: use of modelling to develop targets and policies. *Bulletin of the World Health Organization* 2015; **93**: 790-8.
4. Houben RM, Menzies NA, Sumner T, et al. Feasibility of achieving the 2025 WHO global tuberculosis targets in South Africa, China, and India: a combined analysis of 11 mathematical models. *The Lancet Global health* 2016; **4**(11): e806-e15.
5. Huynh GH, Klein DJ, Chin DP, et al. Tuberculosis control strategies to reach the 2035 global targets in China: the role of changing demographics and reactivation disease. *BMC medicine* 2015; **13**: 88.
6. Dye C, Williams BG. Eliminating human tuberculosis in the twenty-first century. *Journal of the Royal Society Interface* 2008; **5**(23): 653-62.
7. Xu K, Ding C, Mangan CJ, et al. Tuberculosis in China: A longitudinal predictive model of the general population and recommendations for achieving WHO goals. *Respirology* 2017: Online first.
8. Aeras. Aeras Annual Report. 2015. http://www.aeras.org/annualreport2015 (accessed 10th November 2016).
9. Harris RC, Sumner T, Knight GM, White RG. Systematic review of mathematical models exploring the epidemiological impact of future TB vaccines. *Hum Vaccin Immunother* 2016; **12**(11): 2813-32.
10. Arregui S, Sanz J, Marinova D, et al. A data-driven model for the assessment of age-dependent patterns of Tuberculosis burden and impact evaluation of novel vaccines. *bioRxiv* 2017: Online first.
11. Anhui Zhifei Longcom Biologic Pharmacy Co. Phase III Clinical Study of Efficacy and Safety of Vaccae™ to Prevent Tuberculosis. 27th December 2016. https://clinicaltrials.gov/show/NCT01979900 (accessed 3rd January 2017).
12. Aeras and TBVI. TB Vaccine Research and Development: A Business Case for Investment. http://www.aeras.org/pdf/TB_RD_Business_Case_Draft_3.pdf (accessed 14th January 2017 2017).
13. Knight GM, Griffiths UK, Sumner T, et al. Impact and cost-effectiveness of new tuberculosis vaccines in low- and middle-income countries. *Proc Natl Acad Sci U S A* 2014; **111**(43): 15520-5.
14. Negin J, Abimbola S, Marais BJ. Tuberculosis among older adults – time to take notice. *International Journal of Infectious Diseases* 2015; **32**: 135-7.
15. United Nations Population Division. World Population Prospects: The 2012 Revision, Highlights and Advance Tables. ESA/P/WP.228. https://esa.un.org/unpd/wpp/publications/Files/WPP2012_HIGHLIGHTS.pdf (accessed 23rd June 2014).
16. Gao L, Lu W, Bai L, et al. Latent tuberculosis infection in rural China: baseline results of a population-based, multicentre, prospective cohort study. *Lancet Infect Dis* 2015; **15**(3): 310-9.

17. Dye C, Glaziou P, Floyd K, Raviglione M. Prospects for tuberculosis elimination. *Annu Rev Public Health* 2013; **34**: 271-86.
18. Liu S, Li Y, Bi Y, Huang Q. Mixed vaccination strategy for the control of tuberculosis: A case study in China. *Math Biosci Eng* 2017; **14**(3): 695-708.
19. R Core Team. R: A language and environment for statistical computing. . R Foundation for Statistical Computing, Vienna, Austria; 2014.
20. Pathai S, Bajillan H, Landay AL, High KP. Is HIV a Model of Accelerated or Accentuated Aging? *The Journals of Gerontology: Series A* 2014; **69**(7): 833-42.
21. Read JM, Lessler J, Riley S, et al. Social mixing patterns in rural and urban areas of southern China. *Proceedings of the Royal Society B: Biological Sciences* 2014; **281**(1785).
22. WHO. TB treatment outcomes. 2015. http://www.who.int/tb/country/data/download/en/ (accessed 24th August 2015).
23. WHO. WHO TB burden estimates. 2015. http://www.who.int/tb/country/data/download/en/ (accessed 3rd July 2016).
24. United Nations Department of Economic and Social Affairs Population Division. World Population Prospects: The 2015 Revision, custom data acquired via website. 2015. https://esa.un.org/unpd/wpp/Download/Standard/Population/.
25. Jabot F, Faure T, Dumoulin N, Albert C, Adapted by Funk S and Knight G. EasyABC (R package, adapted version). 2014.
26. Marjoram P, Molitor J, Plagnol V, Tavare S. Markov chain Monte Carlo without likelihoods. *Proc Natl Acad Sci U S A* 2003; **100**(26): 15324-8.
27. WHO. Global Tuberculosis Report 2013. 2013. http://www.who.int/tb/publications/global_report/en/.
28. Zhang H, Huang F, Chen W, et al. Estimates of tuberculosis mortality rates in China using the disease surveillance point system, 2004-2010. *Biomed Environ Sci* 2012; **25**(4): 483-8.
29. Orroth KK, White RG, Korenromp EL, et al. Empirical observations underestimate the proportion of human immunodeficiency virus infections attributable to sexually transmitted diseases in the Mwanza and Rakai sexually transmitted disease treatment trials: Simulation results. *Sexually transmitted diseases* 2006; **33**(9): 536-44.
30. Ragonnet R, Trauer JM, Denholm JT, Geard NL, Hellard M, McBryde ES. Vaccination Programs for Endemic Infections: Modelling Real versus Apparent Impacts of Vaccine and Infection Characteristics. *Scientific reports* 2015; **5**: 15468.

Figures and Tables

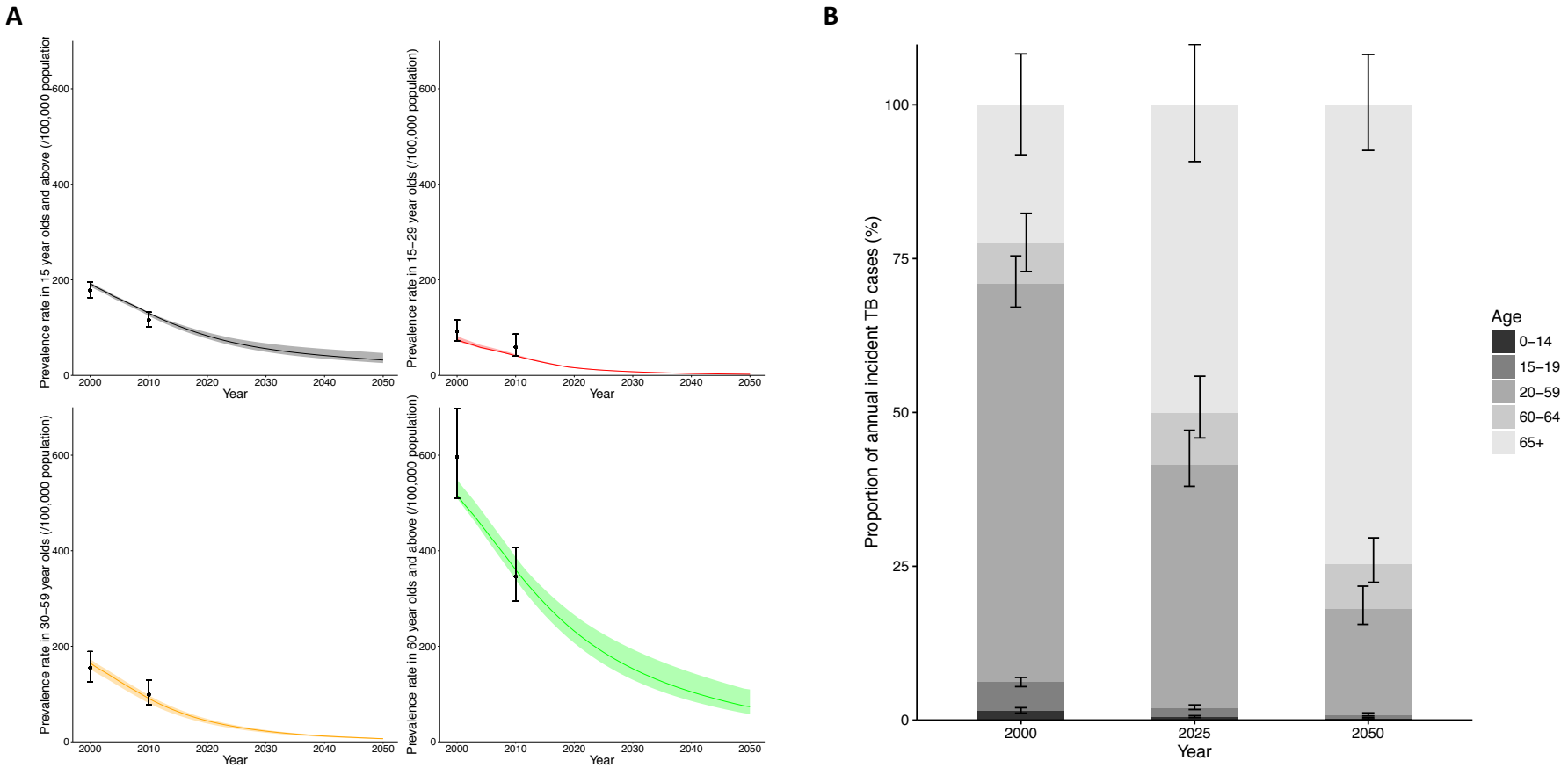


Figure 1: (A) Predicted (lines) and empirical data (points) for microbiologically-positive TB prevalence in ≥15 year olds (black), 15-29 year olds (red), 30-59 year olds (yellow) and ≥60 year olds (green), calibrated to survey-estimated 2000 and 2010 microbiologically-positive pulmonary TB prevalence rates.[12] Ranges due to uncertainty in natural history parameters identified in the calibration process are shown by shaded ribbons. For other calibration see Supplementary Section 4.3.3. **(B) Predicted percentage of annual incident TB cases by age and calendar year in the baseline (no new vaccine) scenario.** Ranges represent the uncertainty across the 1000 model fits around the proportion of incident cases in a given age group.

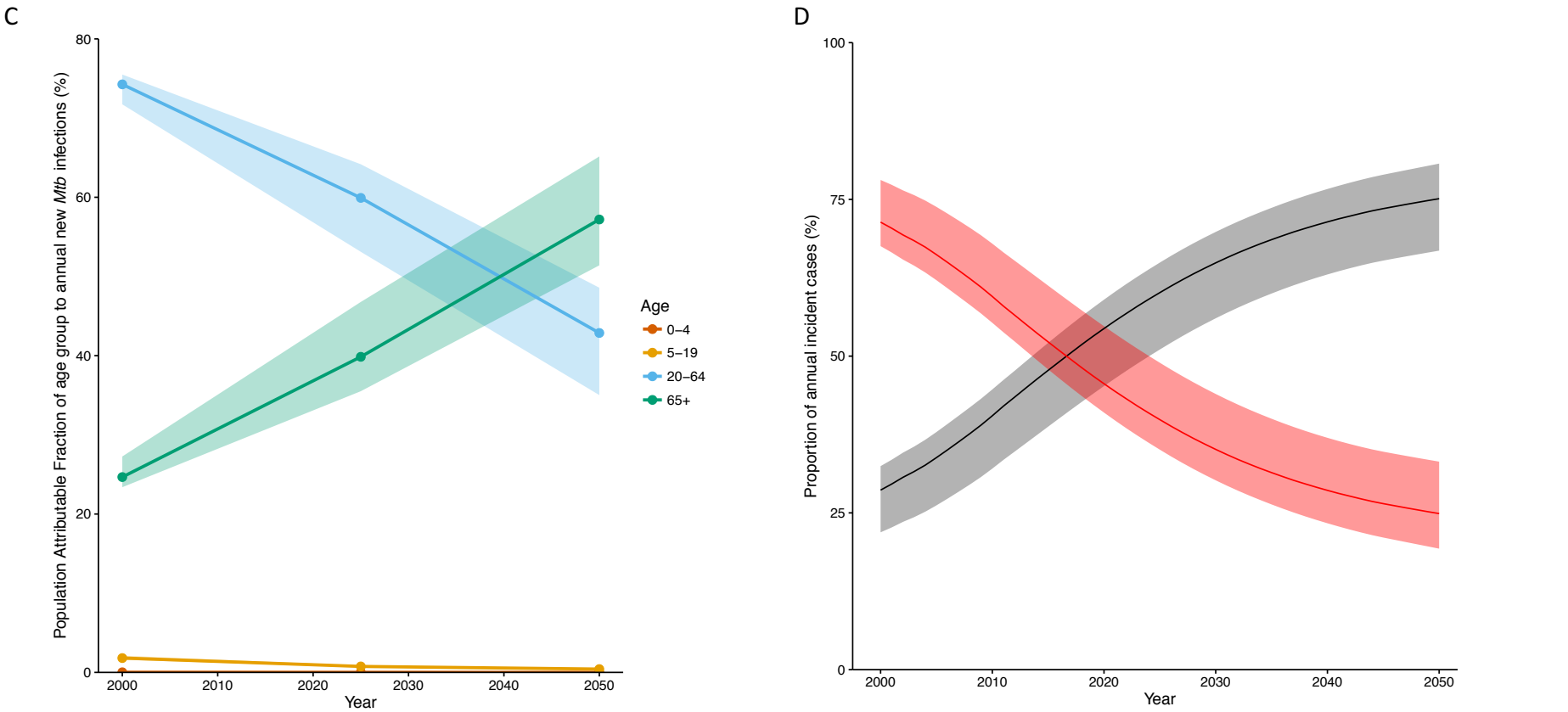


Figure 1: (C) Predicted population attributable fraction of each age group to new *M.tb* infections, by calendar year. Lines represent median estimates for 0-4 year olds (red), 5-19 year olds (yellow), 20-64 years (blue) and 65 years and above (green). Ranges due to uncertainty in natural history parameters are shown by shaded ribbons. *M.tb: Mycobacterium tuberculosis.* **(D) Predicted proportion of all annual incident disease cases due to recent infection (red) versus reactivation (black), by calendar year.** Ranges due to uncertainty in natural history parameters are shown by shaded ribbons.

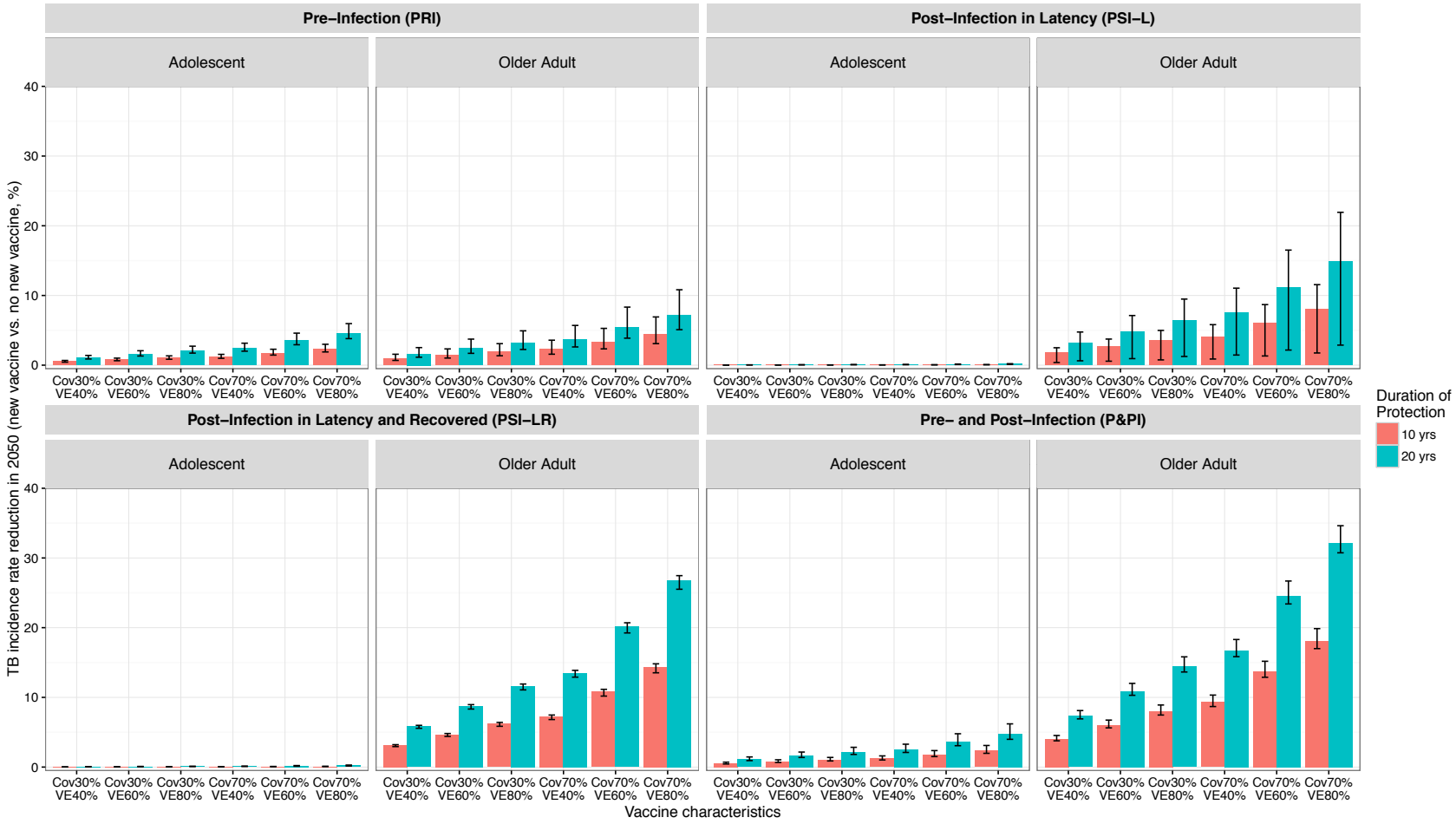


Figure 2: Predicted population level impact on TB incidence rate in 2050 compared to the no new vaccine scenario, by host infection status required for efficacy, age targeting, vaccine efficacy, coverage and duration of protection. Vaccination was implemented in 2025-2050, with adolescent vaccination delivered routinely to 15 year olds and a 3-year catch up campaign in 16-19 year olds 2025-2027, and older adult vaccination delivered routinely to 60 year olds with a 3-year catch up campaign in 61-64 year olds 2025-2027. Cov: vaccine coverage, VE: Vaccine Efficacy.

Table 1: Predicted population level impacts of the 'intermediate' vaccine profile (60% vaccine efficacy, 10 years protection and 70% coverage), by host infection status required for efficacy and age targeting. Ratios presented are of older adult:adolescent values. NNV = Number needed to vaccinate

Host Infection status required for efficacy	Age vaccinated	TB incidence rate reduction in 2050, % (range)	TB incidence rate reduction ratio in 2050 (range)	Mortality rate reduction in 2050, % (range)	TB mortality rate reduction ratio in 2050 (range)	Number of TB cases averted 2025-2050, thousands (range)	Ratio of cases averted (range)	Number of TB deaths averted 2025-2050, thousands (range)	Ratio of deaths averted (range)	NNV per case averted (range)	Ratio of NNV per case averted (range)	NNV per death averted (range)	Ratio of NNV per death averted (range)
Pre-infection (PRI)	**Adolescent**	**1.7** (1.4-2.3)	**1.9** (1.5-2.6)	**0.9** (0.5-1.4)	**4.0** (2.2-7.7)	**248** (214-292)	**1.4** (1.2-1.9)	**3** (2-5)	**3.3** (1.9-7.2)	**1,278** (1,087-1,481)	**0.796** (0.632-0.970)	**101,379** (77,813-144,867)	**0.360** (0.166-0.627)
	Older adult	**3.3** (2.3-5.3)		**3.5** (2.6-5.3)		**370** (287-504)		**9** (5-21)		**1,022** (752-1,318)		**44,613** (36,654-51,666)	
Post-infection latency only (PSI-L)	**Adolescent**	**0.05** (0.04-0.07)	**120.1** (26.4-215.5)	**0.02** (0.01-0.04)	**229.7** (42.8-593.3)	**8** (6-11)	**82.2** (18.0-155.2)	**0.09** (0.05-0.20)	**170.6** (30.1-497.0)	**40,065** (29,505-52,492)	**0.015** (0.008-0.066)	**2,623,571** (2,094,819-4,317,445)	**0.007** (0.002-0.040)
	Older adult	**6.1** (1.3-8.7)		**6.4** (1.3-9.4)		**658** (131-1,081)		**16** (2-45)		**574** (350-2,886)		**31,324** (25,609-55,197)	
Post-infection in latency & recovered from active disease (PSI-L&R)	**Adolescent**	**0.07** (0.05-0.09)	**157.5** (119.3-225.6)	**0.04** (0.02-0.06)	**312.6** (186.5-584.8)	**12** (9-16)	**109.7** (83.8-154.6)	**0.14** (0.07-0.29)	**253.8** (140.3-550.8)	**26,831** (20,437-34,840)	**0.011** (0.008-0.014)	**1,383,557** (1,285,122-2,556,696)	**0.005** (0.002-0.009)
	Older adult	**10.8** (10.2-11.2)		**11.7** (11.0-12.1)		**1,295** (1,037-1,469)		**33** (16-65)		**292** (257-365)		**13,164** (10,132-14,590)	
Pre- and post-infection (P&PI)	**Adolescent**	**1.8** (1.5-2.4)	**7.7** (6.4-8.8)	**0.9** (0.5-1.4)	**16.1** (10.4-0.3)	**259** (224-304)	**6.3** (5.1-7.3)	**3** (2-5)	**14.0** (9.0-31.9)	**1,223** (1,043-1,414)	**0.188** (0.164-0.233)	**94,346** (73,535-137,392)	**0.085** (0.038-0.133)
	Older adult	**13.8** (12.9-15.2)		**14.9** (14.2-16.0)		**1,643** (1,403-1,893)		**42** (22-81)		**230** (199-269)		**10,292** (8,048-11,534)	

4.3 Paper 2 supplementary appendix

The supplementary appendix to research paper 2 is included in this section. As this model was a further development of the model described in Chapter 3, to avoid excessive repetition of methods and baseline model fitting results, sections from the paper supplement that repeat methodology or results described previously have been replaced with signposting to the relevant sections in Chapter 3.

4.3.1 Model structure

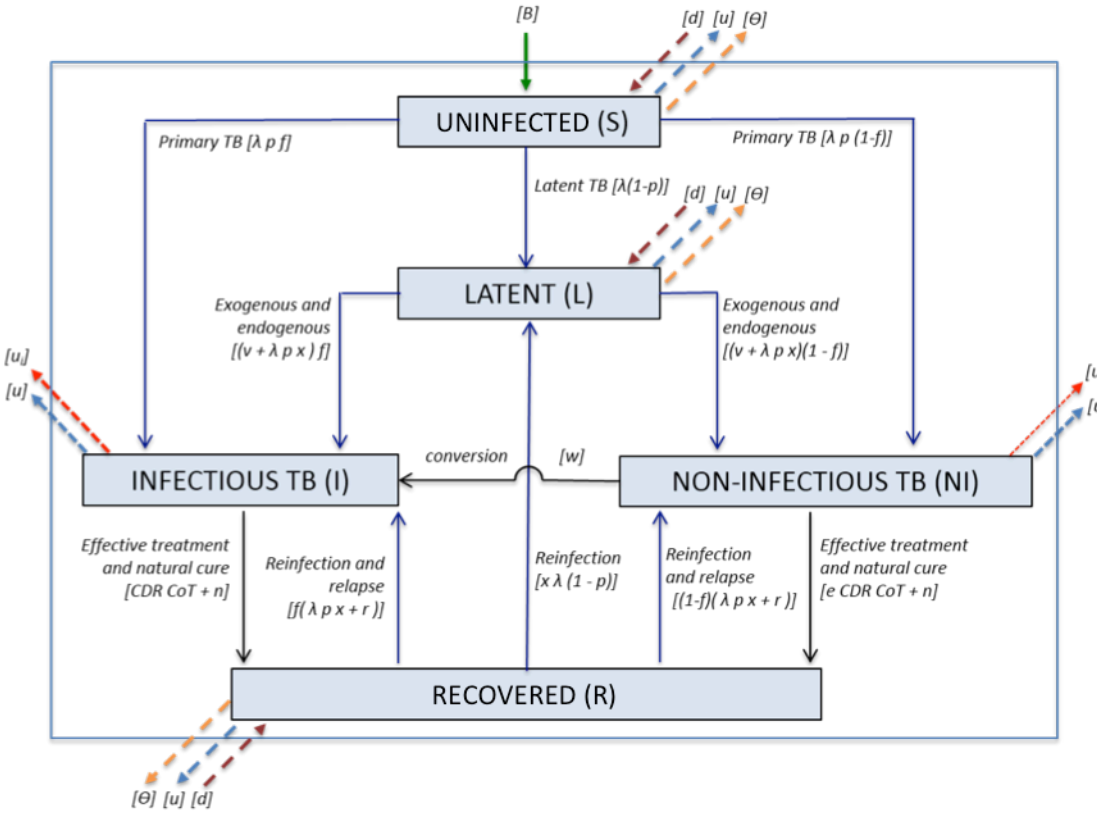
An age-structured, compartmental, deterministic transmission model (Figure S1) described by a series of difference equations (section 4.3.2) was developed based upon the epidemiological model described in Chapter 3. The narrative description of TB natural history represented by the unvaccinated stratum can therefore be found in Chapter 3 section 3.3.

In this study, the vaccine was modelled as all-or-nothing protection, as opposed to leaky/degree protection, thus did not permit development of disease in the vaccinated population. As a consequence, the transitions between the unvaccinated and vaccinated strata differ from Chapter 3. Therefore, the full set of equations for both strata are provided in section 4.3.2 below. The new model vaccine structure in this study is described below.

Vaccination was incorporated into the model through a distinct vaccinated stratum, consisting of uninfected (S), latently infected (L) and recovered after disease (R) states (Figure S1b). Therapeutic vaccination was not considered, so transition from unvaccinated to vaccinated strata was only possible for the uninfected, latent and recovered populations, at rate θ_S, θ_L and θ_R, respectively. Theta was equivalent to vaccine coverage (*cov*) multiplied by vaccine efficacy (VE) in the given age group and year. Protection was modelled as all-or-nothing, therefore unsuccessfully vaccinated individuals (1-VE) remained in the unvaccinated stratum, and successfully vaccinated individuals (VE) were transferred to the vaccinated stratum. Vaccines were modelled

as protective against disease, meaning vaccinated individuals could become infected, but were completely protected against development of disease. Individuals departed the vaccinated stratum either through all-cause mortality (u) or reaching the end of the duration of protection and returning to the unvaccinated stratum (d).

A)



B)

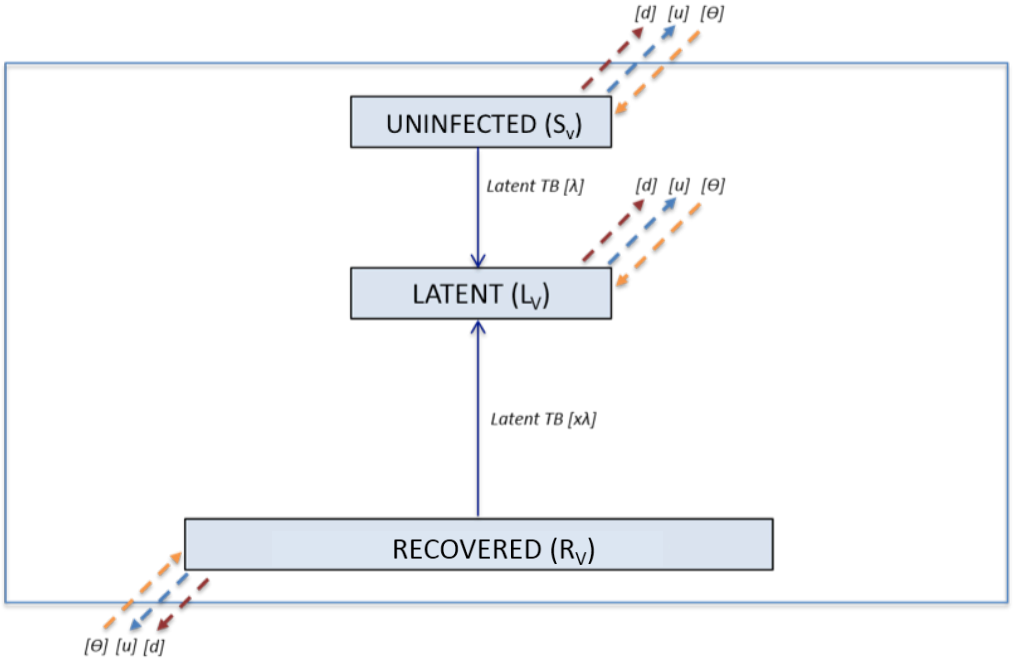


Figure S1: Model structure A) unvaccinated stratum, and B) vaccinated stratum, adapted from Knight et al. 2014[2]

4.3.2 Model equations

Difference equations describing the model populations in year *k*, time step *i* and age *j* are provided below. The size of time step (*dt*), was set to 0.5 years as a balance between sufficient granularity and computing time/capacity. The time step in the model is given by $i = \frac{k-(start\ year)}{dt} + 1$ for the first time step of the year and $i = \frac{k-(start\ year)}{dt} + 2$ for the second time step of the year.

The first set of equations is valid for all time steps *except* that at the start of each year. The second section provides equations for the first time step of each year, and includes ageing, vaccination and vaccine waning.

Births and ageing

Births entered the population as uninfected at the start of each year. Life expectancy was data-informed, but limited to no more than 100 years. Following the methodology of Schenzle (1984),[36] ageing was implemented in the model on an annual basis by transitioning the population in a given sub-population of age *j*, to the same sub-population of age *j*+1 at the end of each year.

Transmission

$$\lambda[i,j] = \eta cal \sum_{y=1}^{y=nygrp} \eta[m,y] z \left(\frac{I[i,y]}{T[i,y]}\right)$$

where $T[i,y] = \sum_{j=jmin}^{j=jmax} \begin{bmatrix} S[i,j] + L[i,j] + I[i,j] + NI[i,j] + R[i,j] \\ +S_V[i,j] + L_V[i,j] + R_V[i,j] \end{bmatrix}$ and *nygrp* was the number of contact age groups, *m* was the age group of the individual exposed to infection (including age *j*), *y* was age group of contacts, $\eta[m,y]$ was number of respiratory contacts of age group *m* with contacts of age group y, ηcal was the calibration factor for model fitting, z was the probability of transmission per respiratory contact between an infectious active case and a uninfected person (which is later scaled for protection afforded by latent infection), and *jmin and jmax* were the lower and upper bounds of age classes within a contact age group (*y*).

Vaccination

Vaccination (θ) and waning of protection (*d*) were assumed to occur in the first time step of a given year. As vaccine protects only against disease, those in the vaccinated stratum could become infected during all time steps.

Detection and treatment

CDR[k,j] was the case detection rate (proportion) for year *k* in age *j*, and *CoT[k]* was the proportion of detected cases that were treated successfully in a given year.

Equations for time steps NOT the first time step of the year

Unvaccinated

Uninfected (Susceptible)

$$S[i,j] = S[i-1,j] - (u[j] + \lambda[i-1,j])S[i-1,j]dt$$

Latent

$$L[i,j] = L[i-1,j] + \lambda[i-1,j](1-p[j])(S[i-1,j] + xR[i-1,j])dt$$
$$-(v[j] + \lambda[i-1,j]p[j]x + u[j])L[i-1,j])dt$$

New infectious active TB cases

$$new\_I[i,j] = \lambda[i-1,j]p[j]f[j](S[i-1,j] + xL[i-1,j] + xR[i-1,j])dt$$
$$+ v[j]f[j]L[i-1,j]dt + r[j]f[j]R[i-1,j]dt + wNI[i-1,j]dt$$

New non-infectious active TB cases

$$new\_NI[i,j] = \lambda[i-1,j]p[j](1-f[j])(S[i-1,j] + xL[i-1,j]$$
$$+ xR[i-1,j])dt$$
$$+ v[j](1-f[j])L[i-1,j]dt + r[j](1-f[j])R[i-1,j]dt$$

Infectious active TB cases

$$I[i,j] = I[i-1,j] + (1 - CDR[k,j] \times CoT[k])new\_I[i,j]$$
$$-(n[j] + u[j] + u_i[j])I[i-1,j]dt$$

Non-infectious active TB cases

$$NI[i,j] = NI[i-1,j] + (1 - e \times CDR[k,j] \times CoT[k]) new\_NI[i,j]$$
$$-(n[j] + u[j] + u_{ni}[j] + w) NI[i-1,j]dt$$

Recovered

$$R[i,j] = R[i-1,j] + n[j](I[i-1,j] + NI[i-1,j])dt$$
$$+(CDR[k,j] \times CoT[k])(new\_I[i,j] + e\ new\_NI[i,j])$$
$$-(r[j] + \lambda[i-1,j]x + u[j])R[i-1,j]dt$$

Vaccinated

Vaccinated uninfected (susceptible)

$$S_V[i,j] = S_V[i-1,j] - \lambda[i-1,j]S_V[i-1,j]dt$$
$$-u[j]S_V[i-1,j]dt$$

Vaccinated latent

$$L_V[i,j] = L_V[i-1,j] + \lambda[i-1,j](S_V[i-1,j] + xR_V[i-1,j])dt$$
$$-u[j]L_V[i-1,j]dt$$

Vaccinated recovered

$$R_V[i,j] = R_V[i-1,j] - x\lambda[i-1,j]R_V[i-1,j]dt$$
$$-u[j]R_V[i-1,j]dt$$

First time step of the year (ageing and vaccine delivery/waning)

Unvaccinated

Uninfected (susceptible)

If j=1: $S(i,1) = B$

If j≠1:

$$S[i,j] = S[i-1,j-1] - (u[j-1] + \lambda[i-1,j-1])S[i-1,j-1]dt$$
$$-\theta_S[k,j]S[i-1,j-1]$$
$$+d[k,j](1-\theta_S[k,j])S_V[i-1,j-1]$$

Latent

$$L[i,j] = L[i-1,j-1] + \lambda[i-1,j-1](1-p[j-1])(S[i-1,j-1]$$
$$+ xR[i-1,j-1])dt$$
$$-(v[j-1] + \lambda[i-1,j-1]p[j-1]x + u[j-1])L[i-1,j-1])dt$$
$$-\theta_L[k,j]L[i-1,j-1]$$
$$+d[k,j](1-\theta_L[k,j])L_V[i-1,j-1]$$

New infectious active TB cases

$$new\_I[i,j] = \lambda[i-1,j-1]p[j-1]f[j-1](S[i-1,j-1] + xL[i-1,j-1]$$
$$+ xR[i-1,j-1])dt$$
$$+ v[j-1]f[j-1]L[i-1,j-1]dt$$
$$+ r[j-1]f[j-1]R[i-1,j-1]dt + wNI[i-1,j-1]dt$$

New non-infectious active TB cases

$$new\_NI[i,j] = \lambda[i-1,j-1]p[j-1](1-f[j-1])(S[i-1,j-1]$$
$$+ xL[i-1,j-1] + xR[i-1,j-1])dt$$
$$+ v[j-1](1-f[j-1])L[i-1,j-1]dt + r[j-1](1-f[j$$
$$-1])R[i-1,j-1]dt$$

Infectious active TB cases

$$I[i,j] = I[i-1,j-1] + (1-CDR[k,j]\times CoT[k])new\_I[i,j]$$
$$-(n[j-1] + u[j-1] + u_i[j-1])I[i-1,j-1]dt$$

Non-infectious active TB cases

$$NI[i,j] = NI[i-1,j-1] + (1 - e \times CDR[k,j] \times CoT[k]) new\_NI[i,j]$$
$$-(n[j-1] + u[j-1] + u_{ni}[j-1] + w) NI[i-1,j-1]dt$$

Recovered

$$R[i,j] = R[i-1,j-1] + n[j-1](I[i-1,j-1] + NI[i-1,j-1])dt$$
$$+(CDR[k,j] \times CoT[k])\left(new_{I[i,j]} + e \times new\_NI[i,j]\right)dt$$
$$-(r[j-1] + \lambda[i-1,j-1]x + u[j-1])R[i-1,j-1]dt$$
$$-\theta_R[k,j]R[i-1,j-1]$$
$$+d[k,j](1 - \theta_R[k,j])R_V[i-1,j-1]$$

Vaccinated

Vaccinated uninfected (susceptible)

$$S_V[i,j] = S_V[i-1,j-1] + \theta_S[k,j]S[i-1,j-1]$$
$$-(\lambda[i-1,j-1] + u[j-1])S_V[i-1,j-1]dt$$
$$-d[k,j](1 - \theta_S[k,j])S_V[i-1,j-1]$$

Vaccinated latent

$$L_V[i,j] = L_V[i-1,j-1] + \theta_L[k,j]L[i-1,j-1]$$
$$+ \lambda[i-1,j-1](S_V[i-1,j-1] + xR_V[i-1,j-1])dt$$
$$- u[j-1]L_V[i-1,j-1]dt - d[k,j](1 - \theta_L[k,j])L_V[i-1,j-1]$$

Vaccinated recovered

$$R_V[i,j] = R_V[i-1,j-1] + \theta_R[k,j]R[i-1,j-1] - x\lambda[i-1,j$$
$$-1]R_V[i-1,j-1]dt$$
$$-u[j-1]R_V[i-1,j-1]dt - d[k,j](1 - \theta_R[k,j])R_V[i-1,j-1]$$

4.3.3 Model parameters, data sources and calibration

Justification of selection and sources for natural history, demographic, social and control parameters have been described previously in Chapter 3.

Natural history parameter range justifications are summarised in Chapter 3, and natural history parameter prior ranges, referred to as Supplementary Table S1 in the manuscript, are described in the same section. Demographic parameters are described in Chapter 3, as are data to inform heterogeneous social mixing patterns by age and existing control measures. As the focus of this research question was related to age-targeted vaccination, the age stratification in natural history parameters, social mixing patterns, control measures and demographics was the most important element of the design and calibration of this model.

As described in full in Chapter 3, a two-stage calibration process was employed, with the first stage a manual calibration to age-stratified UN population projections for 2010 and 2050,[29] and the second stage an adaptive rejection ABC MCMC-based method using a modified *easyABC* package to calibrate to 18 China-specific epidemiological data ranges. To represent age-structured TB epidemiology in China, the model was calibrated to eighteen epidemiological data points and their confidence or uncertainty intervals, representing four outcomes in 2010 (bacteriologically positive prevalence, mortality, notifications and incidence) for all-ages and/or age-stratified epidemiological outcomes. To capture temporal trends, prevalence was calibrated to data from both 2000 and 2010. Supplementary Figures S3-S8 referred to in paper 2 for demonstrating the model fit can be found in Chapter 3, section 3.4.1, figures 3.3-3.8.

4.3.4 Vaccine implementation and characteristics

To answer the research question exploring the impact of vaccine age targeting, the vaccine implementation scenarios modelled in this chapter differ from those in the previous work. Vaccine implementation assumptions include age targeting of vaccination, year of implementation and coverage. The subset of vaccine

characteristics modelled was informed by the results presented in Chapter 3. Modelled vaccine characteristics included the host infection status required for efficacy, the effect type, vaccine efficacy, and duration of protection. These are summarised in Table S3 and described below. The absolute impact of vaccines with these characteristics, as well as the relative impact by age targeting was explored.

4.3.4.1 Implementation: Age targeting of vaccination programmes

The target populations explored in this research were adolescents (15-19 year olds) and older adults (60-64 year olds). Adolescent vaccination was routinely delivered to 15 year olds annually from 2025 to 2050, and a catch up campaign for 16-19 year olds was implemented 2025-2027. The equivalent older adult campaign vaccinated 60 year olds routinely, and 61-64 year olds in the 3-year catch up campaign.

Adolescent vaccination was of interest given many current clinical trials, target product profiles and vaccine development strategies focus on vaccination in this age group. Previous modelling has shown that vaccinating adolescents and/or adults would provide greater and more cost effective impact than vaccinating in infancy over the 2024/25-2050 time horizon.[2,5] In many countries a rapid increase in TB rates occurs during adolescence, therefore vaccination of this age group would hopefully provide protection during the period during which infection and disease rates are known to be high. With a lifelong protective vaccine, vaccination of adolescents could eventually be age de-escalated in to infancy, but given the immediate need for protection in adolescence and adulthood, and the consideration that lifelong protection is unlikely with early candidates, vaccination of adolescents would provide greater impact more rapidly than infant vaccination. Vaccination of 15-19 year olds would allow delivery of vaccination through school-based platforms.

Vaccination of older adults was explored due to the ageing population structure in China and high prevalence of *M.tb* infection observed in older age groups in empirical studies.[1,22] The decision to vaccinate specifically 60-64 year olds took in to consideration immunosenescence, safety, and the age group for which protection

was required. Due to the high burden of disease in the elderly, this population would ideally be protected by new vaccines, but due to safety and immunogenicity concerns with regards to vaccination of elderly populations,[37] delivery to older adults was considered a more feasible approach for protecting older age groups.

Since lifelong protection is considered unlikely, vaccinating as late as possible without entering the main onset of immunosenescence was considered the best approach, as vaccination before immunosenescence increases the likelihood of a robust and sustained immune response to vaccines (Richard Aspinall, oral communication, 5th June 2015). This is supported by results of Zostavax trials against herpes zoster in older adults, where vaccine efficacy in 60-69 year olds (64%, 95% CI: 56-71%) was similar to 50-59 year olds (70%, 95% CI: 54-81%), but much higher than in 70-79 year olds (41%, 95% CI: 28-52%) and ≥80 years (18%, 95% CI: -29 to 48%).[38] Immunosenescence is a gradual process and varies between individuals, therefore there is no specific age that is certain to precede immune deterioration. However, based upon expert opinion and the Zostavax data, vaccination at the age of 60 would precede the onset of immunosenescence in most individuals (Richard Aspinall, oral communication, 5th June 2015).[38] Therefore, vaccination of 60-64 year olds was explored as a possible alternative target population for new TB vaccines in China.

4.3.4.2 Implementation: Timing and coverage

Vaccine was assumed to be widely introduced in 2025. At time of writing, several candidates were due to complete phase IIB/III clinical trials in 2017/2018; therefore if any of these vaccines were successful, vaccine could be implemented by 2025.

Vaccination coverage of 30% and 70% of the target populations were explored. Scale up was assumed to be immediate in routine vaccination (15 or 60 year olds), and catch up campaigns in 16-19 year olds or 61-64 year olds were assumed to occur over a period of three years from introduction, with a third of the desired coverage achieved in each year (e.g. 30% coverage assumes 10% per year for 3 years).

The higher coverage assumption (70%) was based upon school attendance levels, a vaccine acceptance study in China, and coverage achieved with HPV vaccines delivered in schools in South Africa. The gross secondary school enrolment ratio in China is 94.3%, and the HPV coverage achieved in 9 year olds in South Africa is 87%.[39,40] It was anticipated that some adolescents would be missed due to absence on vaccination days, and attendance up to the age of 19 would likely be lower than this, so 70% coverage was deemed a reasonable expectation of adolescent coverage. A qualitative study of acceptance of hypothetical adult hepatitis B vaccination indicated a 55-72% acceptance rate if the vaccine were offered free of charge with 0-100 Yuan compensation towards indirect costs.[41] Therefore, 70% coverage appears feasible in older adults given the upper end of the range from the qualitative study and would be likely in adolescents given school attendance is high.

Influenza vaccination rates were considered a suitable proxy for potential older adult TB coverage as influenza vaccination is provided free of charge to citizens ≥60years in government-run hospitals. Studies of self-reported coverage in the elderly indicate vaccine coverage of 36-49% in the 2008-2011 period.[42-44] These data were collected during the 2009 H1N1 pandemic, so may represent higher than usual coverage. Although it is hoped that acceptance and coverage of TB vaccines would be higher than this, 30% coverage was deemed a suitable pessimistic assumption to explore the lower bounds of possible coverage.

4.3.4.3 Characteristics: Host infection status for efficacy

Previous exposure of the immune system may impact the immune response to the vaccine antigen. Previous exposure has been found to reduce the efficacy of BCG,[45] whereas for some viral vectored MV85A vaccines immune priming by previous exposure was found to improve the antigen-specific immune response. Therefore, the efficacy of a given candidate may be reliant upon whether the vaccinated individual is *M.tb*-naïve, currently latently infected or ever-infected. Clinical trials tend to enrol either uninfected [46] or infected [7,47] populations, assessed by IGRA or

PPD, depending upon the anticipated pre-requisites for an immune response according to the antigens included in the vaccine.

Four infection status combinations were modelled (Table S2). Pre-infection vaccines were assumed to prevent development of active disease only in never-infected individuals ($\theta_S > 0, \theta_L = 0, \theta_R = 0$). Two types of post-infection (PSI) vaccines were modelled, the first (PSI-L) was only effective in those currently latently infected with *M.tb* ($\theta_S = 0, \theta_L > 0, \theta_R = 0$), including those who had recovered but were re-infected and enter the latent state. The second (PSI-L&R) was efficacious when delivered to either latently infected populations or those who had recovered from active disease ($\theta_S = 0, \theta_L > 0, \theta_R > 0$). Finally, a vaccine effective pre- and post infection (P&PI), producing immunity in all except those with active disease ($\theta_S > 0, \theta_L > 0, \theta_R > 0$), was modelled.

Therapeutic vaccines were not explored in this study, so no vaccines were effective in populations with active disease.

Table S2: Host infection status for vaccine efficacy

Vaccine	Never infected (S)	Latently infected (L)	Active Disease (I or NI)	Recovered from disease (R)
PRI	Yes	No	No	No
PSI-L	No	Yes	No	No
PSI-L&R	No	Yes	No	Yes
P&PI	Yes	Yes	No	Yes

4.3.4.4 Characteristics: Effect type

Vaccines with a prevention of disease (POD) effect type were modelled. This protection was modelled as 'all-or-nothing', meaning that vaccinated individuals were either completely protected against disease *or* received no protection, and that the distribution between these two subsets was equal to vaccine efficacy. Therefore,

upon vaccination of a population, the effectively vaccinated proportion transitioned to the vaccinated stratum and received complete protection against development of disease, whereas those in whom the vaccine was not effective remained in the unvaccinated stratum and remained fully susceptible to development of disease.

As vaccine did not protect against infection, in the vaccinated class it was still possible to become infected or reinfected and move to the vaccinated latent class. Prevention of disease was modelled as opposed to prevention of infection efficacy, as literature suggests that prevention of disease vaccines generally provide greater impact than prevention of infection vaccines in adolescent and adult populations.[6] In addition, most late-stage vaccines are currently being measured against disease endpoints as opposed to infection endpoints.[7]

4.3.4.5 Characteristics: Efficacy

Vaccine efficacies of 40%, 60% and 80% were modelled. Although examples exist exploring the extremes of vaccine efficacy (17.6-100%), the majority of TB vaccine models in the literature explore a VE range of 40-80%.[6] Given current challenges in developing a highly effective TB vaccine, the upper bound of vaccine efficacy considered technically feasible was 80%.

Although interesting to understand the lower limit of vaccine efficacy required to achieve minimum required levels of impact in Chapter 3, this was not the primary focus of this research question. In Chapter 3, a 40% to 80% efficacy POD P&PI vaccine providing 10 years protection achieved a substantial epidemiological impact of 40-49% up to 60-69% incidence rate reduction. Also of importance is that detection of vaccine efficacy as low as 20% would necessitate potentially unfeasibly large clinical trials to accrue sufficient disease endpoints. Therefore, a lower bound of 40% vaccine efficacy was selected.

When effective in multiple population groups (e.g. latent and recovered), efficacy was assumed to be the same in all groups.

In the model, coverage and efficacy were multiplied to calculate effective coverage (theta), which is the proportion of those in a given model state that are moved to the equivalent vaccinated state at the beginning of each vaccination year.

4.3.4.6 Characteristics: Duration of protection and waning

Mean duration of protection was assumed 10 or 20 years. The majority of models in the literature assumed 10 year or lifelong immunity.[6] Based upon expert consultation, lifelong immunity was considered unlikely with early candidates, therefore 10 and 20 years of protection were modelled.

A sigmoid-shaped waning profile was assumed, therefore waning was normally distributed around a mean of 10 or 20 years protection with a standard deviation of 1 or 2 years, respectively. Waning was exact, meaning that those individuals experiencing waning leave the vaccinated category and return to the equivalent unvaccinated state.

In addition to this basic waning, it was assumed that immunosenescent waning occurs in the elderly (≥65 years) due to ageing-induced deterioration of the immune system. Based upon expert consultation, immunosenescent waning was assumed to be 2% per year in addition to basic waning, starting from when an individual reaches 65 years old (Richard Aspinall, oral communication, 5th June 2015). However, there were few data to inform this assumption, therefore in sensitivity analyses 0% and 5% immunosenescent waning were also explored, based upon expert opinion and results from Zoster vaccine studies in the elderly.[11,38] Additional discussion on this issue can be found in Appendix B.

.

Table S3: Vaccine parameters

Parameter	Symbol	Description	Detail	References	Parameter
Vaccine implementation and characteristics	*c[k,j]*	Coverage of vaccine for those aged *j* in year *k*	Annual routine vaccination of 30% and 70% explored, age targeted to 15 or 60 year olds. In addition, three-year catch up campaign for 16-19 year olds or 61-64 year olds, with above coverages distributed over period of campaign (i.e. 30% would be 10% coverage of the age group in each of the 3 years of the campaign)	Secondary school attendance, qualitative studies, & self-reported influenza vaccine coverage.[39-44]	30% and 70%
	VE_a	Vaccine efficacy for preventing active TB disease	40%, 60%, 80%	Assumed	40%, 60%, 80%
	$\theta_S, \theta_L, \theta_R$ *[k,j]*	Proportion of Uninfected, Latent or Recovered aged *j* moving to vaccine strata in year *k*	Equal to coverage multiplied by vaccine efficacy. Where a is the population vaccinated (i.e. Uninfecteds, Latents or Recovereds): θ_a *[k,j] = c[k,j]* × VE_a	Assumed	PRI (θ_S), PSI-L (θ_L), PSI-L&R (θ_L, θ_R), P&PI ($\theta_S, \theta_L, \theta_R$)
	D	Duration of vaccine efficacy	Average waning of protection at 10 or 20 years, normally distributed around the mean with standard deviation of 1 or 2 years, respectively. Elderly (≥65 years) immunosenescence 2%/year.	Assumed	10 or 20yrs. 2%/year elderly senescence, varied 0-5% in sensitivity analyses.

4.3.5 Additional methods for calculating outcomes

4.3.5.1 Population Attributable Fraction (PAF)

The population attributable fractions (PAFs) of each age group to the annual number of *Mtb* infections in 2000, 2025 and 2050 were calculated in the baseline scenario using methods described previously in Orroth et al.[34] The age groups explored were defined by the four age groups in the contact matrix (0-4, 5-19, 20-64, ≥65 years).

The model was first run allowing transmission from all infected populations as a baseline. It was then re-run four times for each year of interest, blocking transmission from infected individuals in one age group per run by reducing the infectious individuals in that age group to zero. The total number of new infections was estimated in the year of interest for each of these runs. The percentage reduction in new infections compared to the baseline was calculated, giving an estimate of the PAF, indicating the contribution of that age group to *M.tb* transmission in the population as a whole.

4.3.5.2 Cumulative Number Needed to Vaccinate

As this is a dynamic transmission model and the NNV is calculated over an extended period of time (2025-2050) as opposed to annually, the reported measure is the cumulative NNV per case or death averted for the period 2025-2050. It is calculated by dividing the total number of vaccines delivered during that period by the total number of cases or deaths averted during the same period. Programmatically, this measure could be seen as an indicator of impact achieved per vaccine resource expended, therefore minimising the NNV is desirable.

4.3.6 Additional results and discussion

4.3.6.1 Epidemiological outcomes: Burden of disease and contribution to transmission in vaccinated age groups

At time of vaccine introduction (2025), 1·5% (UR: 1·3-1·7) of cases occurred in the 15-19 year-old target group for adolescent vaccination, whereas 8·4% (UR: 7·9-8·8)

occurred in the 60-64 year old age group targeted for older adult vaccination. Routine vaccination of 15 or 60 year olds with a vaccine of 10-years duration of protection would protect 15-24 or 60-69 year-olds. By 2050, 1·8% (UR: 1·5-2·3) and 21% (UR: 20-22) of cases were projected to occur in the 15-24 and 60-69 year-old age groups, respectively. This highlights the potentially avoidable burden of disease, which is clearly higher in the older age group.

Figure 1C in the main article demonstrates the population attributable fraction of each age group towards new infections over time in the no new vaccine baseline. These results suggest a transition from adult-dominated transmission in 2000 (74% from 20-64 year olds; UR: 72-75) to elderly-driven transmission in 2050 (57% from 65 year olds and above; UR: 51-65). Minimal transmission originated from children or adolescents. Transmission from children and adolescents aged 5-19 years old declined from 0·8% (UR: 0·7-0·9%) in 2025 to 0·4% (UR: 0·4-0·5) in 2050. This suggests that the potential indirect effects of an adolescent vaccine would be minimal and declining over the vaccination period.

The contribution of the 20-64 year old group, which includes the vaccination age for older adults also declined during this period. However, older adult vaccination with 10 or 20 years protection extends into the elderly years, and the contribution of the elderly to transmission increased from 39·9% (UR:35·5-46·8) in 2025 to 57·2 (UR: 51·4-65·1%) in 2050. Therefore, a vaccine in this age group could have a substantial indirect effect by reducing the number of cases in, and therefore transmission from, this elderly age group.

4.3.6.2 Epidemiological outcomes: Age-wise proportion of disease arising from reactivation versus new infection

Overall distribution of annual incident disease originating from new infections versus reactivation disease is shown in Figure 1D in the main article. A transition from new-infection driven to a reactivation driven epidemic was estimated to occur between 2014 and 2024. By 2050, 74·4% (UR:73·7-75·8) of incident cases were reactivation disease.

In those below 55 years of age, the majority of incident TB disease cases in the population were consistently caused by new infections over 2000-2050 (Figure S9). Historically, this was also the case in older age groups. However, for 55-64 year olds, the model suggests that a switch from new infections towards reactivation disease occurred in 2002/2003 (UR: <2000 to 2011). In the elderly, it appears the switch to reactivation driven disease occurred before 2000, but uncertainty ranges suggest the transition could have been as late as 2006 (Figure S9). In this elderly group, the model suggests that reactivation disease could constitute 84.3% (UR: 76.4-89.9%) of all annual incident cases in 2050.

The scale up of existing control measures has reduced burden of disease mostly by reducing transmission. The increasing contribution of the elderly to overall TB disease burden during the 2000-2050 period (main article, Figure 1B) combined with the growing contribution of reactivation disease in this age group, can account for the transition from new-infection-driven to a reactivation-driven epidemic in China over the modelled time frame.

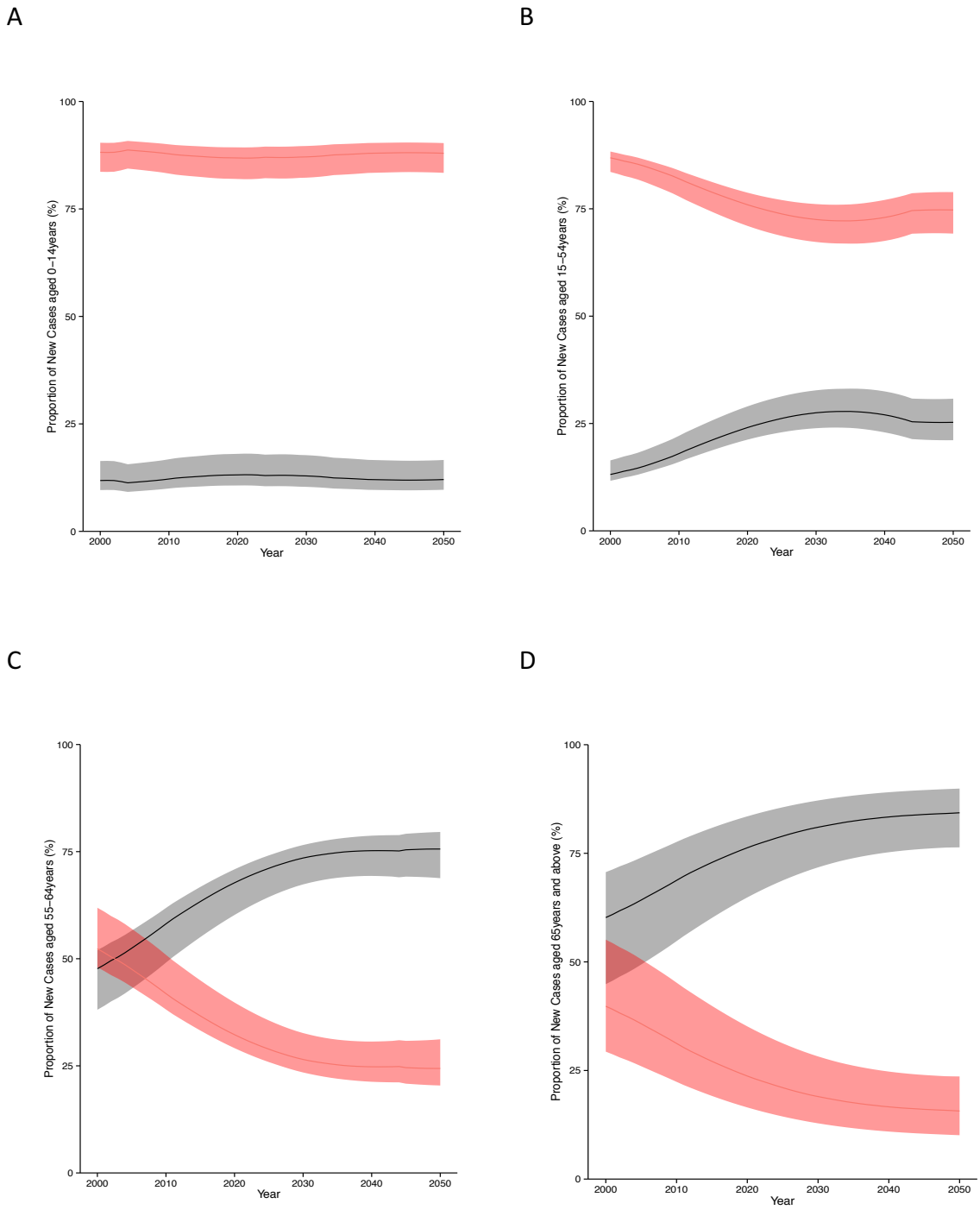


Figure S9: Proportion of annual new active cases emerging from new infections (red) versus reactivation of existing infection (black) from 2000-2050 in age group A. 0-14 year olds, B.15-54 year olds, C. 55-64 year olds, and D. ≥65 years old. Lines are median values and shaded areas represent uncertainty ranges.

4.3.6.3 Epidemiological outcomes: Prevalence of LTBI by age

The model was not calibrated to the Gao et al. prevalence of infection data,[1] as although the four locations selected for the study were considered representative of the social and epidemiological diversity across China, the study did not use random selection of sites or participants from a nationally-representative sampling frame; therefore the generalisability to our country-level model was uncertain.

A comparison to these data, however, found that the modelled age-stratified prevalence of infection matched the published Quantiferon (QFN) data very well (Figure S10). In our model, the overall prevalence of infection was 15.3% (UR: 13.8-17.0%), ranging from 1.3% (1.2-1.4%) in children aged 5-9 years, up to 36.9% (32.9-40.3%) in ≥70 years old. In the Gao study, prevalence of infection measured by Quantiferon was 19%, ranging from 3% in children 5-9 years to 32% in ≥70 years old. The model appeared to slightly overestimate infection prevalence in the >70 years group, but data suggest that QFN and TST tests are slightly less sensitive in the elderly.[48] TST-estimated prevalence was higher than QFN estimates for all except one age group due to lower test specificity. Alignment of model outcomes with the empirical Quantiferon data set speaks to the validity of this epidemiological model.

Prevalence of latent TB at vaccine introduction in 2025 showed a similar age distribution to 2013 (Figure S11), but a slightly lower overall prevalence of 13·6% (UR: 11·0-16·4%) due to declining transmission rates. Prevalence of infection was 1·8% (UR: 1·4-2·2%) in 10-19 year olds, which includes the adolescent 15-19 years age group receiving the vaccine, and 26·6% (UR: 21·6-31·6%) in 60-69 year olds, which includes the older adult 60-64 year old vaccinated group. Prevalence of infection is much higher in the older adult group, meaning that a greater proportion of the population receiving a post-infection vaccine would be effectively immunised in the older adult than the adolescent age group. Conversely, a greater proportion of the adolescent population would be effectively immunised with a pre-infection vaccine, but the risk of developing disease in the uninfected population is obviously much lower than the infected population.

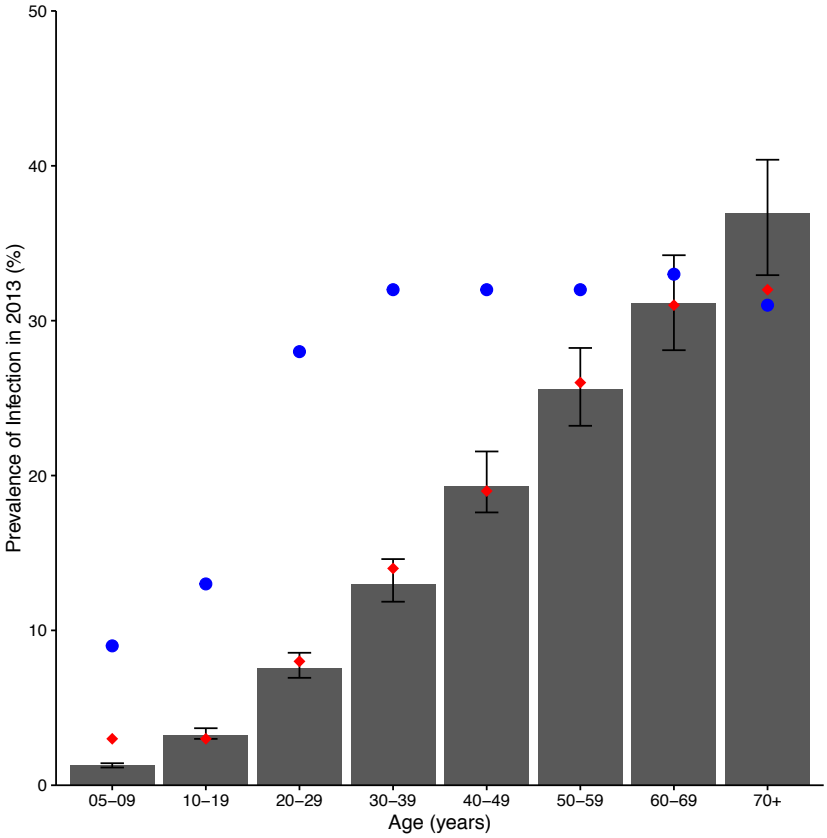


Figure S10: Age-stratified latent TB infection prevalence in 2013, modelled and empirical data. Grey bars and error bars denote modelled median estimates and uncertainty ranges. Empirical data from Gao et al. (2013) are TST prevalence (blue circles) and Quantiferon (red diamonds).[1]

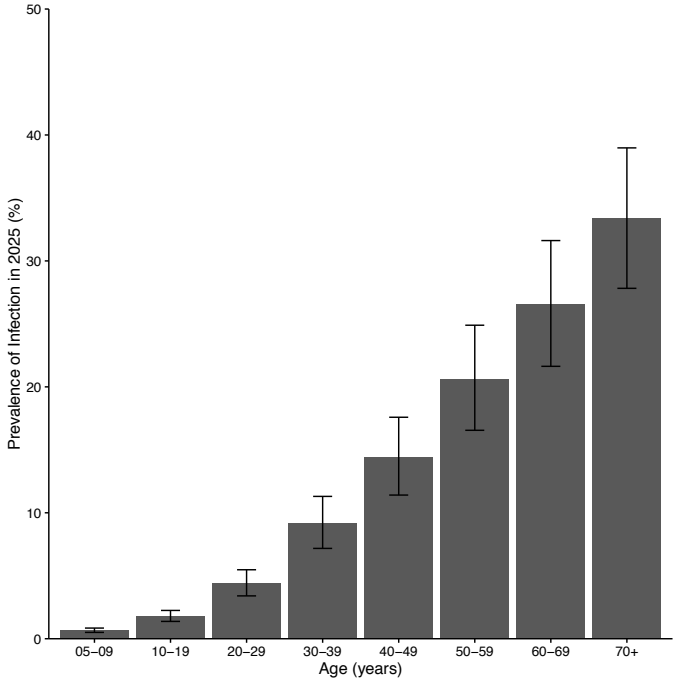


Figure S11: Age-stratified latent TB infection prevalence in 2025, modelled outcomes. Grey bars and error bars denote modelled median estimates and uncertainty ranges.

4.3.6.4 Vaccination outcomes: Comparison of post-infection vaccine effect types (PSI-L vs PSI-L&R)

Two types of post-infection (PSI) vaccine were compared: a vaccine effective only in latent populations (PSI-L) and a vaccine effective in both latent and recovered populations (PSI-L&R). Results in the main article (Figure 2) demonstrate a substantial increase in impact by protecting the recovered population in addition to the latently infected population. However, there was large variability in the amount of additional impact that could be gained by protecting the recovered population as a result of uncertainty in estimates of the impact of the latency-only vaccine (PSI-L).

Whereas uncertainty ranges on the PSI-L&R results were relatively narrow, uncertainty ranges on latency-only vaccines were wide. This was a result of the calibration process. Calibration targets allowed a relatively precise estimation of total reactivation disease (from both latent and recovered combined), but did not permit precise calibration of disease originating from the latently infected population alone. As discussed in more detail below, this led to wide uncertainty in the proportion of the TB burden contributed by the latently infected population, and therefore wide intervals for the latency-only post infection vaccines.

As shown in Figure 1B in the main text, during the 2025-2050 vaccination period, the majority of disease originated from the older adult and elderly population, and supplementary Figure S9 C&D demonstrate that this population experience mostly reactivation disease. However, as can be seen in Figure S12, the proportion of this reactivation disease coming from latent versus recovered populations showed wide variation between the 1000 fitted model runs, especially in the elderly age group.

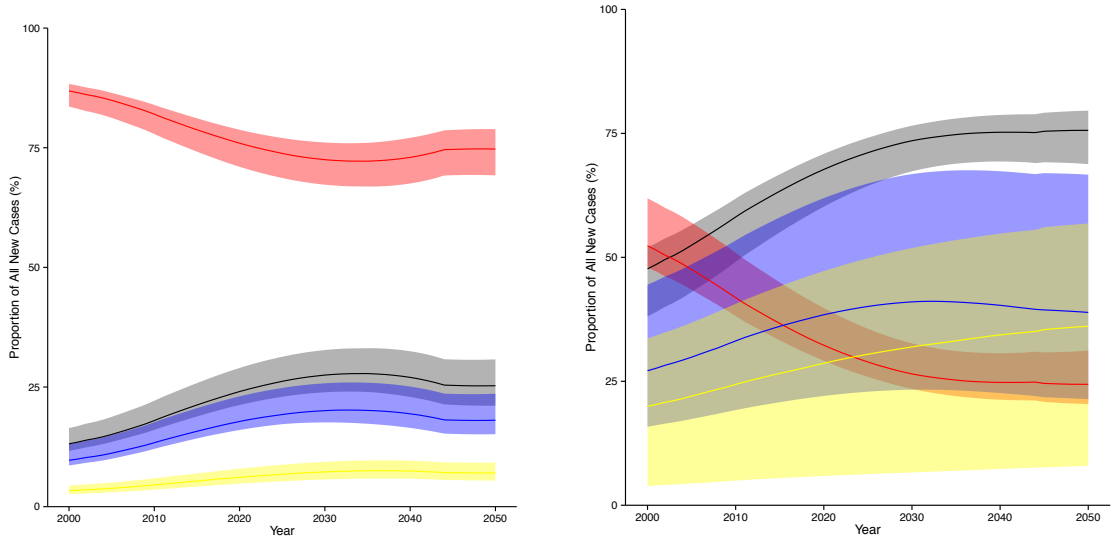


Figure S12: Origin of annual incident cases in A) 15-54 year olds and B) the elderly. Proportion of all annual incident cases originating from new infections (red) or reactivation (black). Reactivation disease is further broken down to show reactivation cases originating from the latent population (yellow) or recovered population (blue)

The proportion of disease from latent versus recovered populations in 2050 was inversely associated (Figure S13A).

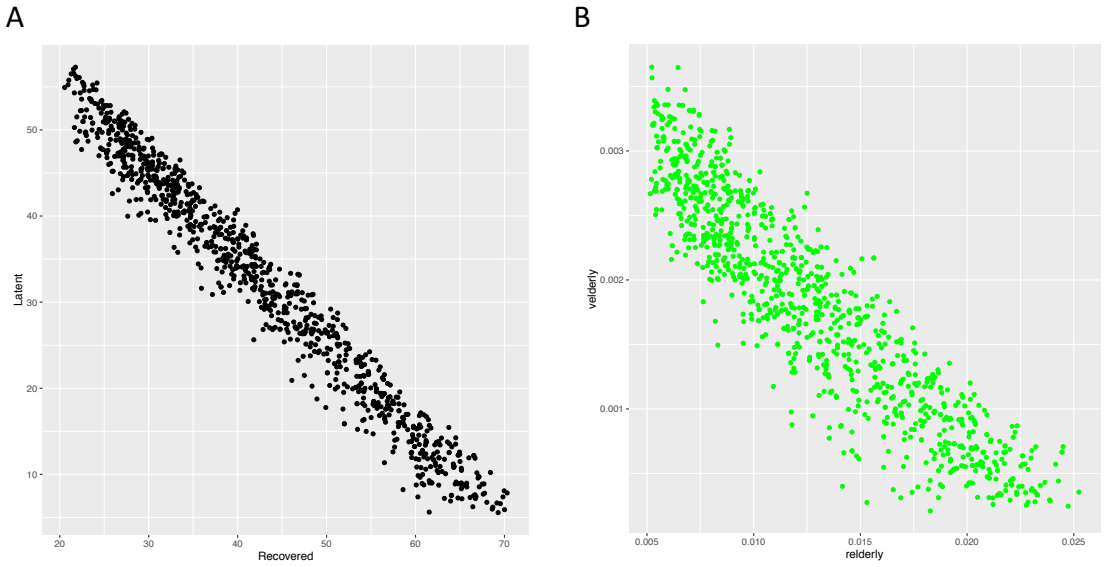


Figure S13: A) The proportion of reactivation disease from latent and recovered populations for 1000 model fits in 2050 and B) Association between risk of reactivation from the latent elderly population (*velderly*) and risk of reactivation from the recovered elderly population (*relderly*) in 1000 model fits.

All parameters from the 1000 model fits were plotted against the ratio of reactivation disease from latent:recovered populations to identify the drivers of this balance (not shown). Only two parameters were found to be associated with this ratio: the rate of reactivation from latency in the elderly (*velderly*) and the rate of reactivation from the recovered class in the elderly (*relderly*).

As the model was fitted to age-wise notification and prevalence rates, and the majority of elderly disease was reactivation disease, these reactivation parameters were the main drivers for fitting to the burden of disease in the elderly. The two parameters were inversely associated (Figure S13B), as the model produced a good fit to the elderly burden of disease data if a high value for one of these parameters was compensated for by a low value in the other parameter. Therefore, a narrow fit to overall burden of disease was achieved, but with a wide range of possible values for these two mutually compensatory parameters and therefore wide uncertainty ranges around the proportion coming from latency or recovered populations. Given the burden of disease originating from the latent population was so varied, the achievable impact with a latency-only post infection vaccine (PSI-L) was also highly variable within the 1000 model fits.

A better understanding of the contribution of recurrence versus latent reactivation would be required for calibration to fully understand the potential impact of a PSI-L vaccine, which may be important, as some TB vaccine clinical trials are recruiting only latently infected participants and excluding those with previous history of TB.[7]

Wide uncertainty ranges were observed in both adolescent and older adult PSI-L vaccine scenarios. However, uncertainty was proportionally greater in older adult vaccination due to the larger burden of reactivation disease in older age groups and the broader priors for older adult and elderly reactivation parameters. In addition, the contribution of reactivation from the recovered population in adolescents is limited by the small proportion of the population experiencing TB disease by adolescence.

Regardless of the uncertainties in the PSI-L vaccine impact, PSI-L&R vaccines provided consistently substantially higher impact, indicating that in vaccine development it may be important to explore whether a vaccine is safe and effective in both those latently infected and recovered from disease.

The wide intervals on the PSI-L results also affect the comparison with the pre-infection vaccines. Although median impact estimates were much higher for post-infection latency-only than pre-infection vaccines, overlap of estimate ranges indicated some uncertainty as to the difference in impact between these vaccines. For example, the latency-only post-infection vaccine with "intermediate" profile provided to older adults reduced 2050 incidence rates by 6·1% (UR: 1·3-8·7%) compared to baseline, indicating some uncertainty as to whether a latency-only vaccine would have greater impact than the equivalent pre-infection vaccine (3·3%, UR:2·3-5·3%).

4.3.6.5 Vaccination outcomes: Reduction in 2050 TB mortality rate in vaccine scenarios compared to the no new vaccine baseline

The reduction in population-level TB mortality rate in 2050 compared to the no new vaccine scenario was estimated by host infection status required for efficacy, age targeting, vaccine efficacy, coverage and duration of protection (Figure S14). Overall, trends observed were reflective of those observed for incidence rates (Figure 2, main article).

In all vaccination scenarios, in absolute terms more cases were averted than deaths. Logically this follows, given at least one case must be averted to avoid a death, and the vaccines were assumed not to affect disease severity.

At a population level, adolescent vaccination yields a greater percentage reduction in incidence rate than the percentage reduction in mortality rate, whereas the opposite was observed in older adult vaccination. This is a function of the relative TB mortality rate in the population directly protected by the vaccine. The model is

parameterised such that TB death rates in the elderly (≥65 years) are equal to or higher than in adolescents and adults. As older adult vaccines providing 10 or 20 years duration of protection protect up to the age of 69 or 79 years, respectively, higher case fatality rates in the elderly period (≥65 years) meant that more deaths were avoided per case averted through older adult vaccination than adolescent vaccination, as most cases averted in adolescent vaccination were in younger age groups experiencing lower case fatality rates. Consequently, older adult vaccination lowers the average population-level CFR, whereas adolescent vaccination increases the average population-level CFR.

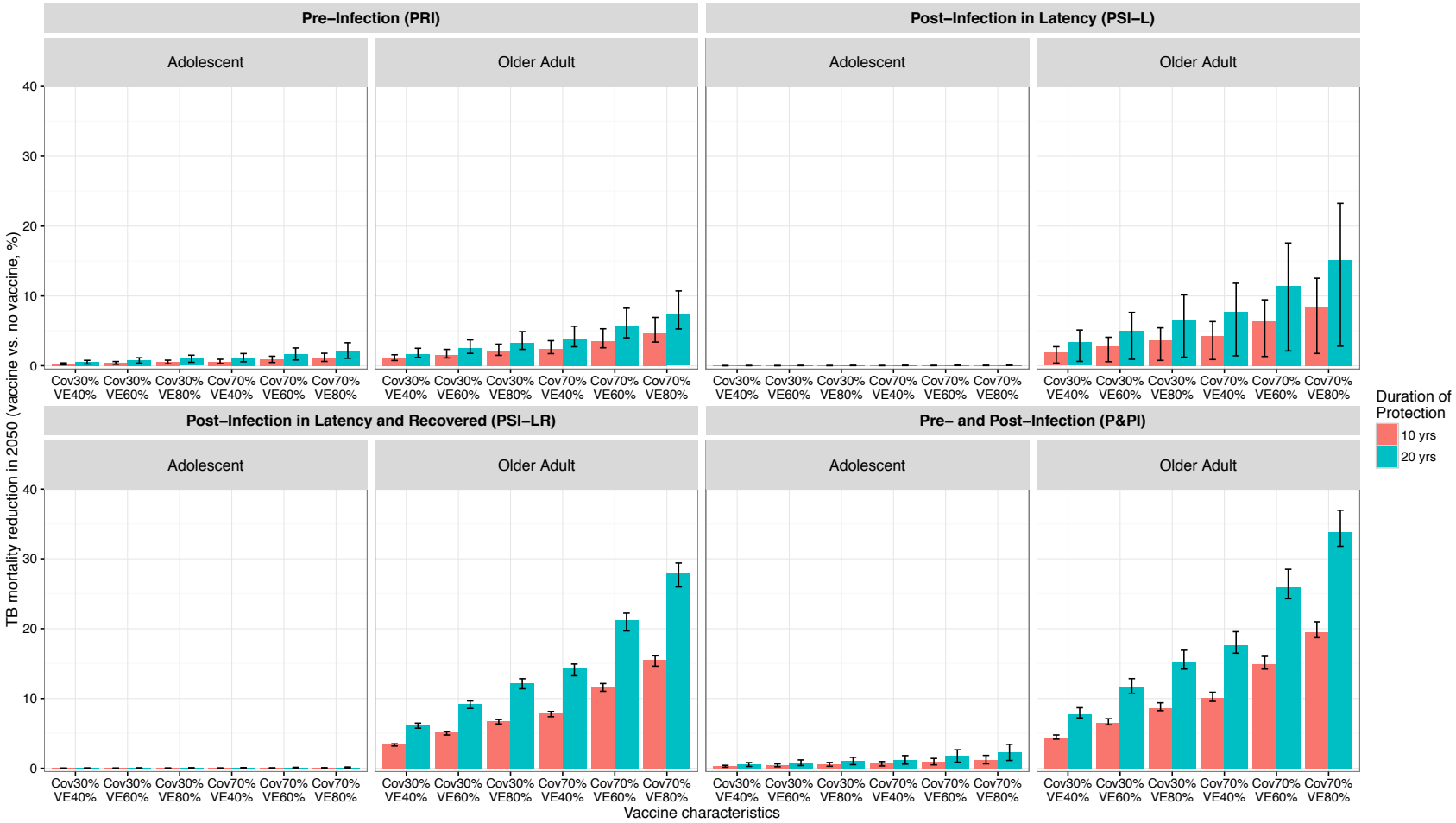


Figure S14: Reduction in population level TB mortality rate in 2050 compared to the no new vaccine scenario, by host infection status required for efficacy, age targeting, vaccine efficacy, coverage and duration of protection. Vaccination was implemented in 2025-2050, with adolescent vaccination delivered routinely to 15 year olds and a 3-year catch up campaign in 16-19 year olds 2025-2027, and older adult vaccination delivered routinely to 60 year olds with a 3-year catch up campaign in 61-64 year olds 2025-2027. Cov: vaccine coverage, VE: Vaccine Efficacy, Dur: Average duration of vaccine protection.

4.3.6.6 Vaccination outcomes: Cumulative number needed to vaccinate (NNV) per case or death averted 2025-2050

The cumulative number of vaccines delivered 2025-2050 and cumulative cases and deaths averted compared to the no new vaccine baseline during the same time period were estimated. These were used to calculate the cumulative number needed to vaccinate per case averted (Table S4) and cumulative number needed to vaccinate per death averted (Table S5) over the 2025-2050 period.

Due to the demographic structure of the population, vaccination of older adults required more doses than the equivalent coverage in adolescents (e.g. 70% coverage required approximately 378 million older adult vaccines or 317 million adolescent vaccines 2025-2050). Given the main demographic parameters were fixed in the first stage of the fitting process, uncertainty ranges are representative of the small differences in population size as a result of variation in TB natural history parameters.

The cumulative number of cases and deaths averted by adolescent vaccination 2025-2050 ranged from 2000 (UR: 2,000-3,000) to 502,000 (UR: 431,000-591,000) cases (Table S4), and 30 (UR: 10-60) up to 5,220 (2,760-8,690) deaths (Table S5). In older adult vaccination campaigns, many more cases and deaths were anticipated to be averted, estimated at 110,000 (UR: 85,000-151,000) to 3.0 million (UR: 2.5-3.5m) cases and 2,790 (1,400-6,410) to 73,730 (37,700-146,410) deaths through older adult vaccination campaigns.

The lowest achievable cumulative NNVs per case averted 2025-2050 with the adolescent and older adult vaccines explored were 590 (501-688) and 120 (103-141), respectively, reaching as high as 59,070 (43,503-77,369) to 1,505 (1,104-1,944) with the least effective vaccines explored. Cumulative NNV per death averted for adolescent vaccination ranged from 48,826 (37,877-71,140) to 3.8m (3.1m-6.3m), and older adult NNVs ranged from 5,464 (4,249-6,190) to 65,575 (53,816-75,898).

For the highest impact P&PI vaccines, adolescent campaigns required 5.3 times the number of vaccinated individuals to avert one case compared to older adult vaccination (Table 1, main article). The impact of adolescent versus older adult vaccination differed most for PSI-L&R vaccines; for the 'intermediate' vaccine profile, 91 times more adolescents would need to be vaccinated than older adults to avert the same number of cases, and 200 times more adolescents would need to be vaccinated to avert the same number of deaths. The most similar impact profiles were pre-infection vaccines, for which only 1.3 and 2.7 times more adolescent than older adult vaccines were needed per case or death averted, respectively, for the 'intermediate' vaccine profile.

Although a larger number of vaccines were delivered in older adult campaigns than equivalent adolescent campaigns, many more cases and deaths were averted, thus the number needed to vaccinate per case or death averted was lower with older adult vaccination. Given substantially fewer vaccines are required per case or death averted for older adult vaccination, assuming the number and size of doses for protection is the same as for adolescents, the impact per vaccine dose would be higher with older adult vaccination. To confirm whether older adult vaccination would be more cost effective would require full cost effectiveness analyses, as such differences could be offset by costs of delivery. The cost of school-based adolescent delivery is likely to be lower than older adult platforms, though influenza vaccines are delivered in China at 60 years old, so delivery alongside influenza vaccines could be a low-cost option if coverage could be improved.

Cumulative number needed to vaccinate per case or death averted 2025-2050 was found to be marginally higher with 70% coverage than 30% coverage. Although more cases are averted by increased coverage, this is suggestive that the large increase in number of vaccines required is not outweighed or even matched by the additional cases averted by increasing coverage. For example, increasing coverage of the older adult 'intermediate' P&PI vaccine from 30% to 70% increased the number of cases averted from 0·73 million to 1·64 million, but increased the number of vaccines required from 162 million to 378 million, so NNV over 2025-50 was found to increase

marginally from 223 (192-261) to 230 (UR: 199-269) with this increase in coverage (Table S4).

The increase in NNV with higher coverage levels was due to a combination of the vaccination time horizon and the changing disease burden over that period. Higher coverage reduces disease rates more rapidly initially after introduction, so over a short horizon the large additional number of vaccines for higher coverage averts many more cases. However, in the latter part of the vaccination period, the higher coverage levels had already brought down burden of disease to such levels that although the proportion of cases averted in the 70% coverage scenario was still higher, there were fewer cases remaining to be averted than in the 30% coverage scenario. Therefore, in later years, large volumes of additional vaccine were provided for limited additional epidemiological gain in the 70% coverage scenario. Over these longer periods the cumulative number of vaccines delivered for 70% coverage is much higher than with 30% coverage, but because the incremental gain is much lower in the later years while the same high level of vaccination is maintained, the cumulative NNV over this period is higher with 70% coverage than with 30% coverage. Therefore, although NNV is marginally higher with the higher coverage level explored, higher absolute reductions in disease burden were achieved and disease reduction targets could be met much faster with higher coverage levels, so is likely to still be considered programmatically favourable.

Table S4: Cumulative number needed to vaccinate per case averted 2025-2050

Vaccine Mechanism	Age targeting	Coverage (%)	Vaccine Efficacy (%)	Duration	Individuals Vaccinated 2025-2050, Thousands (Uncertainty range)	Cases averted 2025-2050, Thousands (Uncertainty range)	Cumulative NNV per case averted 2025-2050 (Uncertainty range)
PRI	Adolescent	30	40	10	135,917 (135,838-135,980)	75 (64-88)	1,822 (1,551-2,110)
				20	135,917 (135,838-135,980)	113 (97-134)	1,206 (1,015-1,407)
			60	10	135,917 (135,838-135,980)	111 (96-130)	1,227 (1,044-1,422)
				20	135,917 (135,838-135,980)	167 (143-198)	815 (686-950)
			80	10	135,917 (135,838-135,980)	146 (126-172)	930 (791-1077)
				20	135,917 (135,838-135,980)	220 (188-261)	619 (521-722)
		70	40	10	317,139 (316,955-317,287)	169 (146-199)	1,873 (1,593-2,170)
				20	317,139 (316,955-317,287)	254 (218-301)	1,249 (1,052-1,457)
			60	10	317,139 (316,955-317,287)	248 (214-292)	1,278 (1,087-1,481)
				20	317,139 (316,955-317,287)	370 (317-439)	858 (722-1,001)
			80	10	317,139 (316,956-317,288)	323 (279-380)	982 (834-1,137)
				20	317,139 (316,956-317,288)	479 (410-568)	663 (558-773)

Vaccine Mechanism	Age targeting	Coverage (%)	Vaccine Efficacy (%)	Duration	Individuals Vaccinated 2025-2050, Thousands (Uncertainty range)	Cases averted 2025-2050, Thousands (Uncertainty range)	Cumulative NNV per case averted 2025-2050 (Uncertainty range)
PRI	Older Adult	30	40	10	162,162 (161,417-162,745)	110 (85-151)	1,473 (1,077-1,907)
				20	162,162 (161,417-162,745)	148 (114-207)	1,093 (786-1,425)
			60	10	162,162 (161,417-162,745)	164 (127-224)	990 (725-1,281)
				20	162,162 (161,417-162,745)	220 (169-306)	736 (530-959)
			80	10	162,162 (161,417-162,745)	217 (168-296)	748 (549-967)
				20	162,162 (161,417-162,745)	291 (224-403)	558 (402-725)
		70	40	10	378,378 (376,639-379,739)	251 (195-343)	1,505 (1,104-1,944)
				20	378,378 (376,639-379,739)	337 (260-467)	1,123 (811-1,459)
			60	10	378,378 (376,640-379,740)	370 (287-504)	1,022 (752-1,318)
				20	378,378 (376,640-379,740)	494 (381-681)	766 (556-993)
			80	10	378,379 (376,640-379,740)	484 (377-658)	781 (576-1,005)
				20	378,379 (376,640-379,740)	644 (499-884)	587 (428-760)

Vaccine Mechanism	Age targeting	Coverage (%)	Vaccine Efficacy (%)	Duration	Individuals Vaccinated 2025-2050, Thousands (Uncertainty range)	Cases averted 2025-2050, Thousands (Uncertainty range)	Cumulative NNV per case averted 2025-2050 (Uncertainty range)
PSI-L	Adolescent	30	40	10	135,916 (135,838-135,980)	2 (2-3)	57,905 (42,648-75,818)
				20	135,916 (135,838-135,980)	4 (3-5)	35,941 (26,488-46,864)
			60	10	135,916 (135,838-135,980)	3 (3-5)	38,894 (28,645-50,932)
				20	135,916 (135,838-135,980)	6 (4-8)	24,170 (17,812-31,518)
			80	10	135,916 (135,838-135,980)	5 (4-6)	29,389 (21,644-38,490)
				20	135,916 (135,838-135,980)	7 (6-10)	18,285 (13,475-23,846)
		70	40	10	317,139 (316,955-317,287)	5 (4-7)	59,070 (43,503-77,369)
				20	317,139 (316,955-317,287)	9 (7-12)	36,781 (27,104-47,971)
			60	10	317,139 (316,955-317,287)	8 (6-11)	40,065 (29,505-52,492)
				20	317,139 (316,955-317,287)	13 (10-17)	25,019 (18,434-32,636)
			80	10	317,139 (316,955-317,287)	10 (8-14)	30,568 (22,508-40,058)
				20	317,139 (316,955-317,287)	17 (13-22)	19,142 (14,101-24,974)

Vaccine Mechanism	Age targeting	Coverage (%)	Vaccine Efficacy (%)	Duration	Individuals Vaccinated 2025-2050, Thousands (Uncertainty range)	Cases averted 2025-2050, Thousands (Uncertainty range)	Cumulative NNV per case averted 2025-2050 (Uncertainty range)
PSI-L	Older Adult	30	40	10	162,162 (161,417-162,745)	193 (39-317)	838 (511-4,212)
				20	162,162 (161,417-162,745)	280 (52-466)	579 (348-3,138)
			60	10	162,162 (161,417-162,745)	288 (57-473)	562 (342-2,824)
				20	162,162 (161,417-162,745)	418 (77-694)	389 (233-2,106)
			80	10	162,162 (161,417-162,745)	382 (76-628)	424 (258-2,130)
				20	162,162 (161,417-162,745)	553 (102-920)	293 (176-1,589)
		70	40	10	378,378 (376,641-379,740)	445 (89-730)	851 (518-4,275)
				20	378,378 (376,641-379,740)	643 (119-1,068)	589 (354-3,192)
			60	10	378,378 (376,642-379,740)	658 (131-1,081)	574 (350-2,886)
				20	378,379 (376,642-379,740)	950 (175-1,579)	399 (239-2,160)
			80	10	378,379 (376,643-379,741)	867 (173-1,424)	436 (266-2,192)
				20	378,379 (376,644-379,741)	1,247 (230-2075)	304 (182-1,644)

Vaccine Mechanism	Age targeting	Coverage (%)	Vaccine Efficacy (%)	Duration	Individuals Vaccinated 2025-2050, Thousands (Uncertainty range)	Cases averted 2025-2050, Thousands (Uncertainty range)	Cumulative NNV per case averted 2025-2050 (Uncertainty range)
PSI-L&R	Adolescent	30	40	10	135,916 (135,838-135,980)	4 (3-5)	38,771 (29,534-50,337)
				20	135,916 (135,838-135,980)	6 (4-7)	23,774 (18,156-30,782)
			60	10	135,917 (135,838-135,980)	5 (4-7)	26,042 (19,838-33,813)
				20	135,917 (135,838-135,980)	9 (7-11)	15,987 (12,210-20,701)
			80	10	135,917 (135,838-135,980)	7 (5-9)	19,679 (14,990-25,551)
				20	135,917 (135,838-135,980)	11 (9-15)	12,094 (9,237-15,661)
		70	40	10	317,139 (316,955-317,287)	8 (6-11)	39,554 (30,129-51,359)
				20	317,139 (316,955-317,287)	13 (10-17)	24,327 (18,581-31,503)
			60	10	317,139 (316,955-317,287)	12 (9-16)	26,831 (20,437-34,840)
				20	317,139 (316,955-317,287)	19 (15-25)	16,547 (12,638-21,429)
			80	10	317,139 (316,955-317,287)	15 (12-20)	20,471 (15,592-26,585)
				20	317,139 (316,955-317,287)	25 (19-33)	12,660 (9,669-16,396)

Vaccine Mechanism	Age targeting	Coverage (%)	Vaccine Efficacy (%)	Duration	Individuals Vaccinated 2025-2050, Thousands (Uncertainty range)	Cases averted 2025-2050, Thousands (Uncertainty range)	Cumulative NNV per case averted 2025-2050 (Uncertainty range)
PSI-L&R	Older Adult	30	40	10	162,162 (161,417-162,746)	380 (305-431)	426 (375-532)
				20	162,162 (161,417-162,746)	547 (432-625)	297 (259-376)
			60	10	162,162 (161,417-162,746)	567 (455-643)	286 (252-357)
				20	162,162 (161,418-162,746)	815 (643-931)	199 (174-252)
			80	10	162,163 (161,418-162,746)	752 (603-853)	216 (190-269)
				20	162,163 (161,418-162,746)	1,079 (852-1,233)	150 (131-190)
		70	40	10	378,380 (376,642-379,741)	874 (700-992)	433 (381-541)
				20	378,380 (376,642-379,741)	1,254 (990-1,432)	302 (264-383)
			60	10	378,381 (376,644-379,743)	1,295 (1,037-1,469)	292 (257-365)
				20	378,381 (376,644-379,743)	1,852 (1,461-2,117)	204 (179-259)
			80	10	378,382 (376,645-379,744)	1,704 (1,364-1,935)	222 (195-278)
				20	378,383 (376,646-379,744)	2,432 (1,919-2,781)	156 (136-197)

Vaccine Mechanism	Age targeting	Coverage (%)	Vaccine Efficacy (%)	Duration	Individuals Vaccinated 2025-2050, Thousands (Uncertainty range)	Cases averted 2025-2050, Thousands (Uncertainty range)	Cumulative NNV per case averted 2025-2050 (uncertainty range)
P&PI	Adolescent	30	40	10	135,917 (135,838-135,980)	78 (68-92)	1,743 (1,485-2,013)
				20	135,917 (135,838-135,980)	118 (102-139)	1,149 (976-1,339)
			60	10	135,917 (135,838-135,980)	116 (100-136)	1,174 (1,001-1,356)
				20	135,917 (135,838-135,980)	175 (150-206)	776 (659-905)
			80	10	135,917 (135,838-135,980)	153 (132-179)	890 (758-1,028)
				20	135,917 (135,838-135,980)	230 (198-271)	590 (501-688)
		70	40	10	317,139 (316,955-317,287)	177 (153-208)	1,792 (1,527-2,070)
				20	317,139 (316,955-317,287)	266 (229-314)	1,190 (1,011-1,388)
			60	10	317,139 (316,955-317,287)	259 (224-304)	1,223 (1,043-1,414)
				20	317,139 (316,955-317,287)	388 (333-456)	818 (695-954)
			80	10	317,139 (316,956-317,288)	338 (292-396)	939 (801-1,086)
				20	317,139 (316,956-317,288)	502 (431-591)	632 (537-737)

Vaccine Mechanism	Age targeting	Coverage (%)	Vaccine Efficacy (%)	Duration	Individuals Vaccinated 2025-2050, Thousands (Uncertainty range)	Cases averted 2025-2050, Thousands (Uncertainty range)	Cumulative NNV per case averted 2025-2050 (Uncertainty range)
P&PI	Older Adult	30	40	10	162,162 (161,417-162,746)	489 (418-566)	332 (286-387)
				20	162,162 (161,417-162,746)	693 (585-807)	234 (200-277)
			60	10	162,163 (161,418-162,746)	727 (622-841)	223 (192-261)
				20	162,163 (161,418-162,746)	1,028 (868-1,196)	158 (135-187)
			80	10	162,163 (161,418-162,746)	962 (822-1,111)	169 (146-197)
				20	162,163 (161,418-162,746)	1,355 (1,145-1,576)	120 (103-141)
		70	40	10	378,381 (376,643-379,742)	1,116 (954-1,289)	339 (293-396)
				20	378,381 (376,643-379,742)	1,571 (1,327-1,824)	241 (207-285)
			60	10	378,382 (376,645-379,744)	1,643 (1,403-1,893)	230 (199-269)
				20	378,383 (376,645-379,744)	2,300 (1,944-2,663)	164 (142-194)
			80	10	378,384 (376,647-379,745)	2,150 (1,837-2,473)	176 (153-206)
				20	378,384 (376,648-379,746)	2,995 (2,532-3,459)	126 (109-149)

Table S5: Cumulative number needed to vaccinate per death averted 2025-2050

Vaccine Mechanism	Age targeting	Coverage (%)	Vaccine Efficacy (%)	Duration	Individuals Vaccinated 2025-2050, Thousands (Uncertainty range)	Death averted 2025-2050, Thousands (Uncertainty range)	Cumulative NNV per death averted 2025-2050 (uncertainty range)
PRI	Adolescent	30	40	10	135,917 (135,838-135,980)	0.85 (0.45-1.40)	140,995 (108,261-201,462)
				20	135,917 (135,838-135,980)	1.18 (0.62-1.97)	102,019 (77,815-145,567)
			60	10	135,917 (135,838-135,980)	1.26 (0.67-2.08)	95,446 (73,282-136,383)
				20	135,917 (135,838-135,980)	1.74 (0.92-2.91)	69,254 (52,823-98,818)
			80	10	135,917 (135,838-135,980)	1.66 (0.89-2.74)	72,682 (55,800-103,857)
				20	135,917 (135,838-135,980)	2.29 (1.21-3.83)	52,880 (40,334-75,456)
		70	40	10	317,139 (316,955-317,287)	1.92 (1.03-3.17)	146,838 (112,726-209,823)
				20	317,139 (316,955-317,287)	2.65 (1.39-4.43)	107,023 (81,630-152,716)
			60	10	317,139 (316,955-317,287)	2.81 (1.51-4.64)	101,379 (77,813-144,867)
				20	317,139 (316,955-317,287)	3.85 (2.03-6.44)	74,340 (56,700-106,081)
			80	10	317,139 (316,956-317,288)	3.65 (1.96-6.04)	78,703 (60,397-112,465)
				20	317,139 (316,956-317,288)	4.98 (2.62-8.34)	58,048 (44,272-82,833)

Vaccine Mechanism	Age targeting	Coverage (%)	Vaccine Efficacy (%)	Duration	Individuals Vaccinated 2025-2050, Thousands (Uncertainty range)	Deaths averted 2025-2050, Thousands (Uncertainty range)	Cumulative NNV per death averted 2025-2050 (Uncertainty range)
PRI	Older Adult	30	40	10	162,162 (161,417-162,745)	2.79 (1.40-6.41)	64,045 (52,494-74,077)
				20	162,162 (161,417-162,745)	3.66 (1.76-8.71)	50,007 (40,335-56,999)
			60	10	162,162 (161,417-162,745)	4.14 (2.08-9.51)	43,078 (35,326-49,839)
				20	162,162 (161,417-162,745)	5.43 (2.61-12.90)	33,721 (27,189-38,416)
			80	10	162,162 (161,417-162,745)	5.47 (2.75-12.54)	32,596 (26,742-37,721)
				20	162,162 (161,417-162,745)	7.15 (3.44-16.98)	25,580 (20,617-29,126)
		70	40	10	378,378 (376,639-379,739)	6.34 (3.19-14.53)	65,575 (53,816-75,898)
				20	378,378 (376,639-379,739)	8.29 (3.98-19.64)	51,546 (41,536-58,671)
			60	10	378,378 (376,640-379,740)	9.31 (4.69-21.27)	44,613 (36,654-51,666)
				20	378,378 (376,640-379,740)	12.13 (5.83-28.61)	35,270 (28,397-40,097)
			80	10	378,379 (376,640-379,740)	12.15 (6.13-27.69)	34,136 (28,076-39,554)
				20	378,379 (376,640-379,740)	15.78 (7.59-37.07)	27,132 (21,832-30,816)

Vaccine Mechanism	Age targeting	Coverage (%)	Vaccine Efficacy (%)	Duration	Individuals Vaccinated 2025-2050, Thousands (Uncertainty range)	Deaths averted 2025-2050, Thousands (Uncertainty range)	Cumulative NNV per death averted 2025-2050 (Uncertainty range)
PSI-L	Adolescent	30	40	10	135,916 (135,838-135,980)	0.03 (0.01-0.06)	3,706,285 (2,958,178-6,096,566)
				20	135,916 (135,838-135,980)	0.04 (0.02-0.08)	2,462,677 (1,965,549-4,056,801)
			60	10	135,916 (135,838-135,980)	0.04 (0.02-0.09)	2,500,691 (1,996,083-4,113,801)
				20	135,916 (135,838-135,980)	0.06 (0.03-0.13)	1,663,455 (1,327,628-2,740,425)
			80	10	135,916 (135,838-135,980)	0.05 (0.03-0.11)	1,898,161 (1,515,252-3,122,865)
				20	135,916 (135,838-135,980)	0.08 (0.04-0.17)	1,264,068 (1,008,846-2,082,613)
		70	40	10	317,139 (316,955-317,287)	0.06 (0.03-0.13)	3,826,806 (3,054,996-6,296,245)
				20	317,139 (316,955-317,287)	0.09 (0.05-0.19)	2,550,356 (2,035,391-4,202,038)
			60	10	317,139 (316,955-317,287)	0.09 (0.05-0.20)	2,623,571 (2,094,819-4,317,445)
				20	317,139 (316,955-317,287)	0.13 (0.07-0.28)	1,753,138 (1,399,062-2,889,019)
			80	10	317,139 (316,955-317,287)	0.12 (0.06-0.26)	2,023,385 (1,615,897-3,330,456)
				20	317,139 (316,955-317,287)	0.17 (0.09-0.37)	1,355,760 (1,081,876-2,234,576)

Vaccine Mechanism	Age targeting	Coverage (%)	Vaccine Efficacy (%)	Duration	Individuals Vaccinated 2025-2050, Thousands (Uncertainty range)	Deaths averted 2025-2050, Thousands (Uncertainty range)	Cumulative NNV per death averted 2025-2050 (Uncertainty range)
PSI-L	Older Adult	30	40	10	162,162 (161,417-162,745)	4.60 (0.65-13.31)	45,342 (37,100-79,890)
				20	162,162 (161,417-162,745)	6.34 (0.85-19.17)	32,516 (27,191-59,563)
			60	10	162,162 (161,417-162,745)	6.85 (0.97-19.83)	30,446 (24,908-53,646)
				20	162,162 (161,417-162,745)	9.44 (1.27-28.51)	21,857 (18,274-40,037)
			80	10	162,162 (161,417-162,745)	9.08 (1.28-26.25)	22,999 (18,812-40,524)
				20	162,162 (161,417-162,745)	12.48 (1.68-37.71)	16,528 (13,815-30,275)
		70	40	10	378,378 (376,641-379,740)	10.54 (1.49-30.48)	46,217 (37,799-81,436)
				20	378,378 (376,641-379,740)	14.49 (1.95-43.76)	33,237 (27,778-60,881)
			60	10	378,378 (376,642-379,740)	15.55 (2.19-44.98)	31,324 (25,609-55,197)
				20	378,379 (376,642-379,740)	21.33 (2.87-64.41)	22,582 (18,864-41,362)
			80	10	378,379 (376,643-379,741)	20.41 (2.88-59.01)	23,879 (19,515-42,079)
				20	378,379 (376,644-379,741)	27.92 (3.76-84.29)	17,257 (14,408-31,607)

Vaccine Mechanism	Age targeting	Coverage (%)	Vaccine Efficacy (%)	Duration	Individuals Vaccinated 2025-2050, Thousands (Uncertainty range)	Deaths averted 2025-2050, Thousands (Uncertainty range)	Cumulative NNV per death averted 2025-2050 (Uncertainty range)
PSI-L&R	Adolescent	30	40	10	135,916 (135,838-135,980)	0.04 (0.02-0.09)	1,953,107 (1,814,214-3,607,241)
				20	135,916 (135,838-135,980)	0.06 (0.03-0.13)	1,264,352 (1,173,728-2,346,808)
			60	10	135,917 (135,838-135,980)	0.06 (0.03-0.13)	1,317,972 (1,224,238-2,434,448)
				20	135,917 (135,838-135,980)	0.09 (0.05-0.19)	854,077 (792,859-1,585,394)
			80	10	135,917 (135,838-135,980)	0.08 (0.04-0.17)	1,000,553 (929,387-1,848,334)
				20	135,917 (135,838-135,980)	0.12 (0.06-0.25)	649,058 (602,536-1,204,912)
		70	40	10	317,139 (316,955-317,287)	0.09 (0.05-0.20)	2,017,371 (1,873,873-3,726,978)
				20	317,139 (316,955-317,287)	0.14 (0.07-0.29)	1,309,585 (1,215,717-2,431,232)
			60	10	317,139 (316,955-317,287)	0.14 (0.07-0.29)	1,383,557 (1,285,122-2,556,696)
				20	317,139 (316,955-317,287)	0.20 (0.10-0.42)	900,378 (835,840-1,671,837)
			80	10	317,139 (316,955-317,287)	0.18 (0.09-0.38)	1,067,459 (991,497-1,973,094)
				20	317,139 (316,955-317,287)	0.26 (0.13-0.55)	696,435 (646,515-1,293,391)

Vaccine Mechanism	Age targeting	Coverage (%)	Vaccine Efficacy (%)	Duration	Individuals Vaccinated 2025-2050, Thousands (Uncertainty range)	Deaths averted 2025-2050, Thousands (Uncertainty range)	Cumulative NNV per death averted 2025-2050 (Uncertainty range)
PSI-L&R	Older Adult	30	40	10	162,162 (161,417-162,746)	9.83 (4.83-19.37)	19,036 (14,653-21,122)
				20	162,162 (161,417-162,746)	13.70 (6.59-27.68)	13,582 (10,382-15,247)
			60	10	162,162 (161,417-162,746)	14.64 (7.20-28.85)	12,785 (9,841-14,183)
				20	162,162 (161,418-162,746)	20.38 (9.80-41.18)	9,131 (6,980-10,249)
			80	10	162,163 (161,418-162,746)	19.38 (9.53-38.19)	9,659 (7,435-10,713)
				20	162,163 (161,418-162,746)	26.95 (12.96-54.45)	6,906 (5,279-7,750)
		70	40	10	378,380 (376,642-379,741)	22.50 (11.06-44.34)	19,413 (14,943-21,528)
				20	378,380 (376,642-379,741)	31.27 (15.03-63.19)	13,890 (10,617-15,585)
			60	10	378,381 (376,644-379,743)	33.21 (16.32-65.42)	13,164 (10,132-14,590)
				20	378,381 (376,644-379,743)	46.03 (22.11-93.00)	9,441 (7,216-10,589)
			80	10	378,382 (376,645-379,744)	43.57 (21.40-85.83)	10,040 (7,727-11,122)
				20	378,383 (376,646-379,744)	60.22 (28.93-121.71)	7,218 (5,517-8,092)

Vaccine Mechanism	Age targeting	Coverage (%)	Vaccine Efficacy (%)	Duration	Individuals Vaccinated 2025-2050, Thousands (Uncertainty range)	Deaths averted 2025-2050, Thousands (Uncertainty range)	Cumulative NNV per death averted 2025-2050 (Uncertainty range)
P&PI	Adolescent	30	40	10	135,917 (135,838-135,980)	0.89 (0.48-1.47)	131,098 (102,232-190,927)
				20	135,917 (135,838-135,980)	1.23 (0.65-2.05)	94,158 (73,049-137,194)
			60	10	135,917 (135,838-135,980)	1.31 (0.71-2.18)	88,762 (69,211-129,270)
				20	135,917 (135,838-135,980)	1.83 (0.97-3.04)	63,931 (49,596-93,150)
			80	10	135,917 (135,838-135,980)	1.73 (0.93-2.88)	67,604 (52,708-98,455)
				20	135,917 (135,838-135,980)	2.40 (1.27-3.99)	48,826 (37,877-71,140)
		70	40	10	317,139 (316,955-317,287)	2.00 (1.08-3.33)	136,595 (106,491-198,928)
				20	317,139 (316,955-317,287)	2.78 (1.47-4.62)	98,833 (76,667-143,998)
			60	10	317,139 (316,955-317,287)	2.93 (1.58-4.87)	94,346 (73,535-137,392)
				20	317,139 (316,955-317,287)	4.04 (2.14-6.72)	68,685 (53,275-100,065)
			80	10	317,139 (316,956-317,288)	3.81 (2.06-6.34)	73,274 (57,096-106,698)
				20	317,139 (316,956-317,288)	5.22 (2.76-8.69)	53,659 (41,615-78,168)

Vaccine Mechanism	Age targeting	Coverage (%)	Vaccine Efficacy (%)	Duration	Individuals Vaccinated 2025-2050, Thousands (Uncertainty range)	Deaths averted 2025-2050, Thousands (Uncertainty range)	Cumulative NNV per death averted 2025-2050 (Uncertainty range)
P&PI	Older Adult	30	40	10	162,162 (161,417-162,746)	12.57 (6.56-24.13)	14,698 (11,502-16,502)
				20	162,162 (161,417-162,746)	17.17 (8.74-34.06)	10,671 (8,303-12,097)
			60	10	162,163 (161,418-162,746)	18.68 (9.75-35.85)	9,897 (7,743-11,107)
				20	162,163 (161,418-162,746)	25.47 (12.97-50.51)	7,199 (5,600-8,159)
			80	10	162,163 (161,418-162,746)	24.66 (12.88-47.35)	7,496 (5,864-8,410)
				20	162,163 (161,418-162,746)	33.58 (17.10-66.59)	5,464 (4,249-6,190)
		70	40	10	378,381 (376,643-379,742)	28.59 (14.93-54.90)	15,091 (11,804-16,927)
				20	378,381 (376,643-379,742)	38.89 (19.81-77.10)	11,013 (8,564-12,474)
			60	10	378,382 (376,645-379,744)	41.94 (21.92-80.58)	10,292 (8,048-11,534)
				20	378,383 (376,645-379,744)	56.80 (28.98-112.67)	7,544 (5,863-8,538)
			80	10	378,384 (376,647-379,745)	54.71 (28.63-105.18)	7,893 (6,170-8,838)
				20	378,384 (376,648-379,746)	73.73 (37.70-146.41)	5,810 (4,514-6,571)

4.3.6.7 Vaccination outcomes: Vaccine efficacy by age

Vaccine efficacy was assumed to be invariable by age of administration. This was considered acceptable given the age of routine older adult vaccination (60 years) was selected to precede the main onset of immunosenescence,[49] and given that zoster vaccine trials have provided proof of concept for achieving robust efficacy and immunogenicity in older age groups.[11] However, to account for possible immunosenescence, an additional 2% waning per year (0-5% in sensitivity analyses) was assumed from the age of 65 years.

Given the higher proportion of co-morbidities in older adults than adolescents, and the experience with elderly vaccination against influenza,[50] vaccine efficacy could conceivably be lower in this older population. Therefore, an exploration of whether the conclusions would be robust to reduced vaccine efficacy in older age groups follows.

Although the absolute worst case scenario would be a complete lack of efficacy in older adults, this seems an unlikely scenario for the population of 60 year olds. For influenza vaccination, 70-90% efficacy has been observed in adults and 17-53% in the elderly.[50] Using this as a basis for a differential of efficacy, adolescents were assumed to experience 80% vaccine efficacy, and older adults towards the higher end of the elderly influenza vaccine efficacy (40% efficacy) given older adults are younger than the elderly. The longer vaccine duration (20 years) was assumed for adolescents, and for older adults 10 years was assumed, with the added disadvantage of immunosenescence from the age of 65 years. Coverage was also assumed as 70% in adolescents, and only 30% for older adults. These constituted the 'high' and 'low' scenarios for adolescents and older adults, respectively (Table S6).

Table S6: Comparison of impact of adolescent vaccination 'high' scenario (80% VE, 70% coverage, 20 years protection) versus older adult 'low' scenario (40% VE, 30% coverage, 10 years protection).
NNV: Number needed to vaccinate

Vaccine mechanism	Age group vaccinated (scenario)	Cases averted 2025-2050, Thousands (Uncertainty range)	Cumulative NNV per case averted 2025-2050 (Uncertainty range)
PRI	Adolescent (high)	479 (410-568)	663 (558-773)
	Older adult (low)	110 (85-151)	1,473 (1,077-1,907)
PSI-L	Adolescent (high)	17 (13-22)	19,142 (14,101-24,974)
	Older adult (low)	193 (39-317)	838 (511-4,212)
PSI-L&R	Adolescent (high)	25 (19-33)	12,660 (9,669-16,396)
	Older adult (low)	380 (305-431)	426 (375-532)
P&PI	Adolescent (high)	502 (431-591)	632 (537-737)
	Older adult (low)	489 (418-566)	332 (286-387)

Comparing the adolescent 'high' to older adult 'low' vaccine and coverage, older adult vaccination averted more cases and required fewer vaccinations per case averted for both post-infection vaccines. Therefore, even if trials were to show that a post-infection vaccine were much less efficacious and durable and anticipated achievable coverage were much lower in older adults than adolescents, older adult vaccination would likely still provide greater impact than vaccinating adolescents. Therefore, for vaccines with the potential for post-infection efficacy, ensuring trials are designed to examine safety and efficacy in these older age groups is essential.

When comparing identical vaccine efficacies and durations, pre-infection vaccines provided greater impact with older adult vaccination, though the age-wise difference is much smaller than was observed post-infection vaccines. Therefore, when comparison was made between the 'high' adolescent and 'low' older adult vaccine scenario, results suggested that the adolescent vaccine may provide substantially

higher impact, avoiding more than four times the number of cases, with an NNV of less than half the older adult estimate.

To explore this further, sensitivity to reductions in individual characteristics in the older adult vaccination scenario was estimated for the median estimate of cases averted. Coverage was the biggest influencing factor, as the median estimate of cases averted by older adult vaccination reduced by 23% by halving efficacy to 40%, 25% by halving duration of protection to 10 years, and by 55% by reducing coverage from 30% to 70%. The reduction in duration or efficacy was insufficient to bring the median cases averted by older adult vaccination below the adolescent vaccination estimate. The impact of reduced coverage in older adults was sufficient for the number of cases averted by this strategy to fall below the adolescent vaccination strategy with higher coverage, though the NNV per case averted was still lower with older adult vaccination. If both coverage and efficacy were reduced simultaneously in older adults, adolescent vaccination became preferable. Therefore, if clinical trials were to demonstrate substantially lower efficacy and/or duration of protection in older adults, or feasibility assessments were suggestive of low coverage in this older age group, further modelling would be required to assess whether adolescent vaccination or another approach to vaccination might be preferable. Understanding vaccine characteristics across age groups will be important to inform data-driven vaccination strategies with a pre-infection vaccine.

For a vaccine efficacious pre- and post-infection, the median number of cases averted for the 'high' adolescent vaccine scenario was marginally higher than the 'low' older adult vaccine, but with largely overlapping uncertainty ranges (Table S6). The 'high' scenario (70% coverage) in adolescents required almost double the number of vaccines needed for 30% coverage of older adults, therefore the NNV per case averted was much lower for older adult vaccination. Therefore, with a P&PI vaccine, even with the pessimistic assumption of halving vaccine efficacy and duration of protection for older adults compared to adolescent vaccination, assuming vaccine waning of 2% due to immunosenescence after the age of 65 years, and assuming much lower coverage (30% instead of 70%), older adult vaccination would still

provide a similar level of impact to adolescent vaccination, and would be a much more efficient use of resources as would require almost half the number of vaccines to achieve that impact.

The key recommendation from earlier in this chapter was the need for inclusion of older adults in clinical trials, to investigate efficacy and allow for the possibility of registration in this epidemiologically important age group. This additional discussion further strengthens this recommendation. For post-infection or pre- and post-infection vaccines, older adult vaccination would be recommended even with substantially lower efficacy or duration than observed in adolescents. Therefore, data to support registration in this population will be vital. For pre-infection-only vaccines, age-stratified estimates of efficacy will be important for designing impact-maximising implementation strategies, as the preference for vaccinating older adults could be diminished or negated if efficacy, duration or coverage were substantially lower than in adolescents.

4.3.6.8 Comparison of results to Chapter 3

In Chapter 3 the research question focussed on exploring the impact of vaccine characteristics, keeping implementation static as routine vaccination of 9 year olds alongside periodic mass vaccination of ≥10 year olds. Implementation scenarios (age and coverage) and the relative impact of vaccine characteristics with different age targeting were explored in this chapter.

Although differences exist between the two models, the most comparable vaccines were the 10-year protection and 60% efficacy against disease effective pre- and post-infection scenarios, with coverage of 80% in chapter 3 and 70% in this chapter. With this vaccine, routine plus mass vaccination in Chapter 3 reduced 2050 incidence rates by 53.3% (UR: 50.6-56.1%), averting 7.5 million cases 2025-50 (UR: 7.0-8.1). The adolescent vaccination policy in this chapter reduced 2050 incidence rates by 1.8% (UR: 1.5-2.4%), averting 0.26 million cases (UR: 0.22-0.30) during 2025-50, and the

older adult vaccination policy reduced incidence rates by 13.8% (UR: 12.9-15.2%), averting 1.64 million cases (UR: 1.40-1.89).

Although the mass versus age-targeted implementation scenarios are compared, age of vaccination was not the only contributory factor to impact differences. For example, routine coverage was 10% higher in Chapter 3, and protection was modelled as all-or-nothing instead of 'leaky' in this chapter. Therefore, although results provide a comparison of targeted versus mass campaigns, these additional distinctions must be considered.

From a country decision-maker perspective, both epidemiological impact and the resources required are important considerations, as discussed earlier in this chapter. Although number of vaccines delivered was not estimated in Chapter 3, based upon modelled population size an approximate estimate of the number required for the three 10-yearly mass vaccination campaigns with 70% coverage of 10-100 year olds could be made (approximately 2.6 billion). The three mass campaigns alone (i.e. not accounting for routine vaccination of 9 year-olds) would require more than six times the number of vaccines delivered in any of the 70% coverage scenarios in this chapter, which would also have implications for the number needed to vaccinate per case averted.

Feasibility of mass vaccination described in Chapter 3 would rely on the viability and practicality of such large-scale production. China is the world's largest vaccine producer, manufacturing in total more than one billion vaccine doses per annum,[51] so country-level capacity would be anticipated to be able to meet such demand. Achievability also relies on TB remaining a public health priority in China and the availability of funding for vaccine production and delivery.

Assuming these do not become limiting issues, broad mass vaccination of adolescents and adults would bring about a bigger reduction in 2050 incidence rates than targeted campaigns of 15 or 60 year olds, therefore would be the ideal scenario to

maximise vaccine impact if resources were not limited. However, the supply, logistical and cost implications of vaccinating the entire population aged ≥10 years every decade, in addition to routine vaccination of 9 year olds, would be substantial. Therefore, although the results of Chapter 3 support such wide-scope mass campaigns where feasible, the results of this chapter demonstrate that targeted routine vaccination of older adults could help maximise impact if resources are limited. Therefore, investigating efficacy in this older adult population is essential in clinical trials.

4.3.6.9 Additional discussion of context

Previous mathematical modelling research has demonstrated that vaccination of adolescents, or adolescents and adults, would be expected to provide greater impact and cost effectiveness than neonatal vaccination.[2,4,5] Neonatal vaccination with new TB vaccines has been predicted to provide limited impact in China, even with the optimistic assumption of 100% efficacy against infection with lifelong protection and 95% coverage,[4] therefore vaccination of infants was not explored in this research. Discussion of the results of this study in context are found in the main manuscript in this chapter, but some additional comparisons with the two most recent and relevant studies are provided below.

The Liu study comprised a TB model calibrated to China's TB notifications, comparing neonatal vaccination to neonatal plus periodic mass all-age vaccination ('mixed vaccination'). Over a 17-year time frame, 7.63×10^8 doses were delivered in the key mixed vaccination scenario.[4] In this chapter, the number of vaccines ranged 1.36-3.78×10^8, but were delivered to a narrower age group and over a longer time frame (26 years). In the Liu study, the time to achieving the 2035 WHO goal was halved, yet the number of vaccinations delivered did not double, by implementing the 25% coverage 'mixed' vaccination strategy compared to neonatal vaccination. Therefore, mass or adult campaigns appear to deliver more impact per vaccination than neonatal vaccination. However, the number needed to vaccinate per case averted

was not estimated, therefore direct comparisons cannot be made with the vaccination scenarios presented in this thesis.

The Liu model was calibrated to China's new TB notification data for 2004-2014, with a good fit to temporal trends achieved. Due to specifically fitting to these annual data, the Liu model may more closely track year-on-year notifications than the research presented in these two chapters, but this comes at the cost of having not calibrated to detailed age and epidemiological data (mortality, incidence and prevalence), which is essential to more accurately reflect the natural history and transmission dynamics. Such age-relevant dynamics were essential to the age targeting research question in this chapter, and were achieved in my research through significant age-specific structure, parameterisation and calibration.

Empirical studies have demonstrated the importance of recurrent TB, with one study in China indicating that populations recovered from a previous TB episode experienced disease rates of more than 18 times the observed rate in the general population.[52] Modelling recurrence is an important natural history transition, and in my research accounted for approximately 20-70% of all reactivation disease. The range on this estimate is wide due to the lack of appropriate calibration data to apportion cases between progression from latent and recovered populations, but even the lower end of the modelled estimate of recurrence demonstrates an important contribution. Vaccine protection of the population recovered from disease provided substantial additional impact beyond that achievable by protecting only latent populations. In the Liu et al. study, recurrent disease was not modelled.[4] Therefore, to calibrate to notification data, transmission or progression to disease rates may have been overestimated, and the future burden of TB may have been underestimated without the contribution of recurrence. If transmission was overestimated, the estimated impact of such a prevention of infection vaccine may be too high, and predictions of achieving 2035 goals would be optimistic. Inclusion of the recovered population is clearly important for predicting the future of China's epidemic and potential impact of future vaccines, thus the inclusion of relapse and

reinfection following recovery is an important element of the research presented here.

Another recently published TB vaccine study (Arregui et al., 2017),[5] explored the influence of age specificity in model parameterisation and fitting on the estimated impact of future TB vaccines. Similar to the Liu study, a prevention of infection vaccine with relatively optimistic vaccine assumptions was modelled, including 80% vaccine efficacy with lifelong protection and 100% coverage of the target population.

Vaccination of 15 year olds in the Arregui study was somewhat comparable to the 80% efficacy vaccine with 20-year protection delivered to adolescents in the study in this chapter, though it should be noted that the Arregui vaccine was for lifelong prevention of infection and the model was regional as opposed to China-specific.[5] Although not explicitly stated in the publication, it is assumed that China was included in the Western Pacific Region (WPR), as defined by WHO.[53] In the Arregui study, this vaccine was estimated to avert approximately 2.3 million cases in the WPR during 2025-2050 (extracted from Figure S16),[5] and the similar adolescent vaccination scenario in this thesis averted 0.5 million (UR: 0.4-0.6m) cases in China over the same period. Due to model and vaccine differences, this comparison is interpreted with caution.

Interestingly, in the Arregui WPR model, sensitivity analyses removing the age specificity in the demography and epidemiology of the model increased the number of cases averted by adolescent vaccination.[5] Therefore, without age-specificity in the model, the impact of adolescent vaccines and the difference between infant and adolescent vaccines was overestimated. In the light of results presented in Chapters 3 and 4 of this thesis, it is hypothesised that this difference derives from over-allocation of cases and transmission to younger age groups by heterogeneous contact patterns. This was supported by commentary in the Arregui publication indicating that the eldest age groups were the most impacted by the inclusion of heterogeneous contact patterns.[5]

In the model presented in this thesis, and in the Arregui 'complete' model,[5] fitting to age-stratified data and contact patterns ensured that the distribution of disease burden was reflective of age-structured epidemiology and demography. Importantly, in the model presented here, such age-structured models are valuable for more accurately estimating the future burden of disease, the impact of age-targeted vaccination, and most importantly for identifying the ages that should be targeted to provide the greatest population-level impact.

In both the Arregui model and this thesis, reactivation rates were higher in the recovered state than the latent state. In Arregui et al., re-infection leading to slow progression therefore remains in the recovered state so progression is reflective of the population's disease history[5] In the model presented in this thesis, re-infection leading to slow progression enters the latent state. Therefore, the natural history of reinfection in recovered populations could be improved in future work. Such a change would not be expected to alter the direction of impact or conclusions of this research. It would be anticipated to marginally increase the number and proportion of cases arising from the recovered population, which would be likely to increase the absolute estimates of vaccine impact, strengthen the conclusion that older adult vaccination with post-infection vaccines would provide greater impact than adolescent vaccination, and the PSI-L&R vaccine would have even greater impact compared to PSI-L vaccines.

Several limitations observed in the Arregui study that are important for exploring age-targeting were avoided in the research in this thesis.[5] For example, the Arregui study modelled 5-year age steps up to 70 years old.[5] Yet, in this study, I estimated that 38% of all active cases in 2025 occurred in the population ≥70 years; therefore it was vital to explicitly model the burden and transmission from these oldest age groups and explore vaccines protecting older ages. Therefore, I modelled 1-year age steps up to 100 years of age, and explored vaccines delivered to older adults capable of protecting into the elderly period.

Although contact patterns were not the most influential age-related factor, their inclusion has been demonstrated to change vaccine impact estimates.[5] Given the greatest change was observed in older age groups when switching from homogenous to heterogeneous contact patterns,[5] data-informed contact patterns were clearly important for the research question in this chapter. In the Arregui study, Polymod data from eight European countries was applied to the five modelled regions, and the influence of setting-specific demography was managed by scaling the contact matrix to regional demographic estimates.[5] However, this does not account for differences at the country level, nor does it account for the influence of social and cultural aspects. Therefore, using country-specific data, as was the case in the study presented in this thesis, ensured contact patterns were representative of the modelled population.

The Arregui study demonstrated that epidemiological, demographic and social age dependencies substantially changed vaccine impact estimates, supporting the importance of the substantial age structure and calibration in the model presented in this thesis. The Arregui study explored just one vaccine profile targeted to infants or adolescents, whereas this thesis provides the research needed to inform TPP decision making through an in-depth exploration of vaccine profiles and consideration of vaccination of older adult populations.

4.4 Conclusion

The research presented in this chapter investigated the impact of age-targeted vaccination programmes on population-level TB epidemiology (objective 3). The results clearly demonstrated that, with all vaccine characteristics explored, older adult vaccination delivered greater impact than adolescent vaccination. Older adult vaccination delivered larger reductions in 2050 TB incidence rates, averted more cases and deaths, and required fewer vaccinations per case averted than adolescent vaccination; demonstrating that older adult vaccination is likely to maximise impact and prove a more efficient use of resources. Previous research has recommended adolescent/adult over infant vaccination to maximise impact.[2,4,5] This research

develops this argument for vaccinating adolescent/adult age groups further by demonstrating that in ageing epidemics, most impact could be achieved by targeting older adults, potentially minimising the resources required per case averted.

The relative impact of adolescent versus older adult vaccination schedules differed between pre-infection and post-infection vaccines. Post-infection vaccines provided substantial impact in older adults, whereas impact was negligible when this vaccine was delivered to adolescents. Although pre-infection vaccines provided relatively similar levels of impact when vaccinating either age group, the impact was substantially lower than that achievable with vaccination of older adults with a post-infection vaccine. In this setting, a vaccine efficacious post-infection or pre- and post-infection delivered to older adults will be critical to maximise population-level impact. Previous research was suggestive of this importance of post-infection vaccines for reducing the burden of disease in China,[23] and is supported by the results of Chapter 3 and the older adult vaccination campaign in this chapter. However, if delivered to adolescents, post-infection vaccines would provide much lower impact than pre-infection vaccines. Therefore, the research presented here provides a critical clarification: that maximising population-level impact with post-infection vaccines is contingent on delivery to older age groups.

When planning vaccination programmes, there may be other considerations that would influence the decision regarding age of vaccination, such as the resources required or the accessibility of the populations. School-based platforms would provide relatively easy access to adolescents, whereas strengthening the influenza vaccine platform for vaccination of 60 year olds may require more investment and may achieve lower coverages. However, almost all vaccines delivered higher population-level impact with older adult vaccination even when comparing as little as 30% coverage of older adults to 70% coverage of adolescents. Therefore, even with reduced coverage, older adult vaccination would be the preferred vaccination age.

More vaccines were delivered in the older adult campaigns, but the number needed to vaccinate (NNV) per case averted was consistently lower than for adolescents, demonstrating that the increased impact far outweighs the additional vaccines required. NNV is a useful measure of the balance between resource use and benefit, and by this measure clearly demonstrates that older adult vaccination remains the most efficient use of resources as well as providing greatest impact. However, differences may exist in the costs of vaccine delivery between these two age groups and the age of averted cases is not taken into consideration in the NNV measure. Therefore, future work could develop more detailed health economic analyses, by incorporation of more detailed costings of vaccine and delivery, and by estimating quality adjusted life years gained by vaccination.

When compared to the results of Chapter 3, as would be expected, the impact of routine targeted vaccination was less than that achieved by broad mass vaccination. However, the results of this chapter demonstrate that it is essential that older adult age groups are included in whichever vaccination campaign is conducted. If resources are limited, targeted vaccination of older adults may help maximise the achievable impact.

Previous literature has focussed on infant, adolescent or mass vaccination, therefore this study is the first to show that older adult vaccination with new TB vaccines is likely to be critical for maximising impact in settings such as China, and strengthens the conclusions from previous studies suggesting that vaccines effective post-infection will be needed in this setting. In terms of modelling methodology, age specificity in demographic and epidemic dynamics has not previously been addressed in China vaccine models, therefore research in this thesis is the first to apply such age-specific methodology. This research is one of only two TB models to incorporate heterogeneous social contact patterns,[5] and the only TB model to apply country-specific contact patterns to a model of that country.

There are several limitations of this work, discussed throughout the chapter. These included the dearth of elderly natural history data, the assumption that contact patterns do not change over time and that no new interventions would be introduced. These could be explored in further epidemiological and modelling research. Relative impact of vaccines could change over longer time horizons, but model uncertainty would become too large much beyond 2050. This research does not explore all possible routine vaccination ages as those modelled were considered appropriate given the anticipated future epidemiology and the existence of platforms for vaccine delivery. Once a vaccine is available and the characteristics are known, modelling could be used to further optimise the older adult vaccination age, and explore different mass campaign options.

This research clearly demonstrates the importance of including older adults in vaccine implementation plans. With the exception of influenza and zoster vaccines, the majority of routine vaccines are delivered at young ages. Therefore, older adult platforms may need to be developed or strengthened for delivery of vaccines to older age groups. To allow on-label use following vaccine registration, older age groups need to be included in clinical trials, ideally from phase IIB, but at a minimum by phase III, to ensure that sufficient participants have been exposed to the investigational product to demonstrate safety and immunogenicity to the regulators. Clinical trial outcomes should ideally be disease endpoints. Studies would ideally enrol both pre- and post-infection populations of sufficient sample size to allow stratification of efficacy estimates by infection status. However, the required sample size would likely be infeasible, so at a minimum trials should recruit post-infection populations to demonstrate safety, and potentially conduct stratified analyses as secondary outcomes.

4.5 Overall conclusions from modelling research (Aim 1)

The first aim of this thesis was to “generate mathematical models of age-stratified demography and TB epidemiology to explore the population-level epidemiological

impact of vaccine characteristics and implementation strategies for potential new TB vaccines, using China as a case study".

In Chapter 2, I presented a comprehensive summary of the existing new TB vaccine modelling literature to identify available literature and research gaps for informing vaccine development. This study confirmed that modelling had not comprehensively explored the impact of combinations of efficacy for prevention of infection and/or disease, efficacy pre- and/or post-infection and duration of protection, which are needed to inform vaccine Target Product Profiles. Several studies compared prevention of infection to prevention of disease vaccines, and consistently found prevention of disease vaccines to provide greater impact and more rapidly; however, results were confounded by simultaneous changes to vaccine delivery, so findings were interpreted with caution. Additionally, studies exploring vaccines efficacious against both infection and disease did not explore varying the relative efficacy against infection and disease. The available literature was divided as to whether vaccines effective pre- or post-infection would provide greatest impact, without clarity as to the underlying causes of the differences in results. In addition, no models explored the potential impact of vaccinating older adults, a possible vaccination population of interest in ageing populations such as China. The TB vaccine modelling literature for China explored a very limited subset of vaccines with minimal age structure in the modelled populations. This lack of in-depth modelling of new TB vaccines in China was a significant gap in the literature given its priority as a high burden country, ageing population, and upcoming phase III TB vaccine clinical trial results.

These research gaps were addressed in Chapters 3 and 4, in which I developed an *M.tb* transmission model calibrated to age-stratified demographic and epidemiological data from China, and applied this model as a case study to explore the absolute and relative epidemiological impact of different vaccine characteristics and implementation strategies identified as research gaps in Chapter 2. Results suggest that, in this setting, efficacy for prevention of disease would provide the greatest contribution towards impact. The majority of the achievable impact was through post-infection protection of those latently infected or recovered from active

disease. At least 5 years duration of protection was predicted as necessary to achieve a minimum of 20-29% reduction in incidence rate compared to the no new vaccine baseline in 2050. The median characteristics for efficacy against infection and disease, and duration of protection were estimated to inform the minimum and ideal characteristics in TB vaccine target product profiles.

The results of the second modelling study exploring vaccine implementation indicated that current development plans focussing on adolescents may have minimal impact on the burden of TB in China. Even at much lower coverage, older adult vaccination provided greater population-level impact than vaccination of adolescents. Developing vaccines for older adults to protect through to old age will be imperative for the future of TB control in this setting. In particular, a vaccine able to protect older adult populations already latently infected will be key for maximising population-level impact in ageing, reactivation-driven tuberculosis epidemics similar to China.

The results presented are specific to China, but may also be generalisable to several other high TB burden countries with ageing epidemics, and as China is the third largest contributor to global TB incidence, measures to maximise TB control in China will be of global importance.

In summary, this research demonstrates an important potential contribution of new TB vaccines towards WHO 2050 goals of TB elimination in China, but has important implications for current vaccine development strategy and planning vaccine platforms for implementation. These include the need for development and testing of candidates to explore efficacy against disease, particularly in post-infection populations, the inclusion of older adults in clinical trials, and planning delivery strategies for this age group. These recommendations advocate for the importance of including the above endpoints and populations in TB vaccine development, particularly those currently neglected in many development strategies; but should not be interpreted as requiring the exclusion of infection endpoints, pre-infection

populations, or younger trial populations, as they may provide some additional benefit if sufficient resources are available to include them.

Target product profiles, ideally informed by mathematical modelling as described above, are translated in to clinical development plans (CDPs) designed to develop appropriate candidates ready for delivery to identified target populations. As highlighted in the TB vaccine Blueprint's second epidemiological research need, designing clinical trials and selecting trial sites and recruitment populations requires appropriate epidemiological data.

As has been discussed in the literature, there is a dearth of trial-ready research sites with known sufficient TB burden for vaccine trials.[54] Country and regional level programmatic tuberculosis data, and prevalence surveys where available, can be used to identify regions experiencing a trial-suitable burden of disease. However, data of greater granularity and specific to communities where recruitment is planned is needed, to calculate a sample size that is a data-informed balance of minimising costs and the number of participants exposed to the investigational product, against minimising risk of failure due to insufficient endpoints. Such data can also help direct recruitment through identification of geographical hotspots in the intended recruitment populations (e.g. HIV-negative populations or specific age groups). Given the lack of sufficiently granular available data, and the expense of large prospective prevalence or incidence studies, new methods are required to acquire the data needed for design of TB vaccine efficacy trials in low-income, high-burden settings. The development and evaluation of a novel tool to meet this research need for TB vaccine development is described in Chapter 5.

4.6 Chapter 4 References

1. Gao L, Lu W, Bai L, et al. Latent tuberculosis infection in rural China: baseline results of a population-based, multicentre, prospective cohort study. *Lancet Infect Dis* 2015; **15**(3): 310-9.
2. Knight GM, Griffiths UK, Sumner T, et al. Impact and cost-effectiveness of new tuberculosis vaccines in low- and middle-income countries. *Proc Natl Acad Sci U S A* 2014; **111**(43): 15520-5.
3. Lietman T, Blower SM. Potential Impact of Tuberculosis Vaccines as Epidemic Control Agents. *Clinical Infectious Diseases* 2000; **30**(Supplement 3): S316-S22.
4. Liu S, Li Y, Bi Y, Huang Q. Mixed vaccination strategy for the control of tuberculosis: A case study in China. *Math Biosci Eng* 2017; **14**(3): 695-708.
5. Arregui S, Sanz J, Marinova D, et al. A data-driven model for the assessment of age-dependent patterns of Tuberculosis burden and impact evaluation of novel vaccines. *bioRxiv* 2017: Online first.
6. Harris RC, Sumner T, Knight GM, White RG. Systematic review of mathematical models exploring the epidemiological impact of future TB vaccines. *Hum Vaccin Immunother* 2016; **12**(11): 2813-32.
7. GlaxoSmithKline. Study to Evaluate the Efficacy of GlaxoSmithKline (GSK) Biologicals' Candidate Tuberculosis (TB) Vaccine in Adults (NCT01755598). May 2017 2012. https://clinicaltrials.gov/show/NCT01755598 (accessed 3rd May 2017).
8. Dartmouth-Hitchcock Medical Center. DAR-901 TB Booster Vaccine to Prevent TB in Adolescents (DAR-PIA). 2016. https://clinicaltrials.gov/ct2/show/NCT02712424 (accessed 26th July 2017).
9. Anhui Zhifei Longcom Biologic Pharmacy Co. Phase III Clinical Study of Efficacy and Safety of Vaccae™ to Prevent Tuberculosis. 27th December 2016. https://clinicaltrials.gov/show/NCT01979900 (accessed 3rd January 2017).
10. Lal H, Cunningham AL, Godeaux O, et al. Efficacy of an adjuvanted herpes zoster subunit vaccine in older adults. *N Engl J Med* 2015; **372**(22): 2087-96.
11. Cunningham AL, Lal H, Kovac M, et al. Efficacy of the Herpes Zoster Subunit Vaccine in Adults 70 Years of Age or Older. *New England Journal of Medicine* 2016; **375**(11): 1019-32.
12. Wang L, Zhang H, Ruan Y, et al. Tuberculosis prevalence in China, 1990-2010; a longitudinal analysis of national survey data. *The Lancet* 2014; **383**(9934): 2057-64.
13. World Health Organization. Global Tuberculosis Report 2016. 2016. http://www.who.int/tb/publications/global_report/en/ (accessed 21st December 2016).
14. Lin H, Wang L, Zhang H, Ruan Y, Chin DP, Dye C. Tuberculosis control in China: use of modelling to develop targets and policies. *Bulletin of the World Health Organization* 2015; **93**: 790-8.
15. Houben RM, Menzies NA, Sumner T, et al. Feasibility of achieving the 2025 WHO global tuberculosis targets in South Africa, China, and India: a combined analysis of 11 mathematical models. *The Lancet Global health* 2016; **4**(11): e806-e15.
16. Huynh GH, Klein DJ, Chin DP, et al. Tuberculosis control strategies to reach the 2035 global targets in China: the role of changing demographics and reactivation disease. *BMC medicine* 2015; **13**: 88.

17. Dye C, Williams BG. Eliminating human tuberculosis in the twenty-first century. *Journal of the Royal Society Interface* 2008; **5**(23): 653-62.
18. Xu K, Ding C, Mangan CJ, et al. Tuberculosis in China: A longitudinal predictive model of the general population and recommendations for achieving WHO goals. *Respirology* 2017: Online first.
19. Aeras. Aeras Annual Report. 2015. http://www.aeras.org/annualreport2015 (accessed 10th November 2016).
20. Aeras and TBVI. TB Vaccine Research and Development: A Business Case for Investment. http://www.aeras.org/pdf/TB_RD_Business_Case_Draft_3.pdf (accessed 14th January 2017 2017).
21. Negin J, Abimbola S, Marais BJ. Tuberculosis among older adults – time to take notice. *International Journal of Infectious Diseases* 2015; **32**: 135-7.
22. United Nations Population Division. World Population Prospects: The 2012 Revision, Highlights and Advance Tables. ESA/P/WP.228. 2012. https://esa.un.org/unpd/wpp/publications/Files/WPP2012_HIGHLIGHTS.pdf (accessed 23rd June 2014).
23. Dye C, Glaziou P, Floyd K, Raviglione M. Prospects for tuberculosis elimination. *Annu Rev Public Health* 2013; **34**: 271-86.
24. R Core Team. R: A language and environment for statistical computing. . R Foundation for Statistical Computing, Vienna, Austria; 2014.
25. Pathai S, Bajillan H, Landay AL, High KP. Is HIV a Model of Accelerated or Accentuated Aging? *The Journals of Gerontology: Series A* 2014; **69**(7): 833-42.
26. Read JM, Lessler J, Riley S, et al. Social mixing patterns in rural and urban areas of southern China. *Proceedings of the Royal Society B: Biological Sciences* 2014; **281**(1785).
27. WHO. TB treatment outcomes. 2015. http://www.who.int/tb/country/data/download/en/ (accessed 24th August 2015).
28. WHO. WHO TB burden estimates. 2015. http://www.who.int/tb/country/data/download/en/ (accessed 3rd July 2016).
29. United Nations Department of Economic and Social Affairs Population Division. World Population Prospects: The 2015 Revision, custom data acquired via website. 2015. https://esa.un.org/unpd/wpp/Download/Standard/Population/.
30. Jabot F, Faure T, Dumoulin N, Albert C, Adapted by Funk S and Knight G. EasyABC (R package, adapted version). 2014.
31. Marjoram P, Molitor J, Plagnol V, Tavare S. Markov chain Monte Carlo without likelihoods. *Proc Natl Acad Sci U S A* 2003; **100**(26): 15324-8.
32. WHO. Global Tuberculosis Report 2013. 2013. http://www.who.int/tb/publications/global_report/en/.
33. Zhang H, Huang F, Chen W, et al. Estimates of tuberculosis mortality rates in China using the disease surveillance point system, 2004-2010. *Biomed Environ Sci* 2012; **25**(4): 483-8.
34. Orroth KK, White RG, Korenromp EL, et al. Empirical observations underestimate the proportion of human immunodeficiency virus infections attributable to sexually transmitted diseases in the Mwanza and Rakai sexually transmitted disease treatment trials: Simulation results. *Sexually transmitted diseases* 2006; **33**(9): 536-44.

35. Ragonnet R, Trauer JM, Denholm JT, Geard NL, Hellard M, McBryde ES. Vaccination Programs for Endemic Infections: Modelling Real versus Apparent Impacts of Vaccine and Infection Characteristics. *Scientific reports* 2015; **5**: 15468.
36. Schenzle D. An age-structured model of pre- and post-vaccination measles transmission. *IMA J Math Appl Med Biol* 1984; **1**(2): 169-91.
37. Amanna IJ. Balancing the Efficacy and Safety of Vaccines in the Elderly. *Open Longevity Science* 2012; **6**(2012): 64-72.
38. Merck. Zostavax Package Insert. Whitehouse Station: Merck; 2009. p. 9.
39. UNESCO Institute for Statistics. Education: gross enrolment ratio by level of education. 2016. http://data.uis.unesco.org/?queryid=142 (accessed 3rd January 2017).
40. HPV information centre. South Africa: Human Papillomavirus and Related Cancers, Fact Sheet 2016 (2016-12-15). 2016.
41. Zhu D, Wang J, Wangen KR. Hepatitis B vaccination coverage rates among adults in rural China: Are economic barriers relevant? *Vaccine* 2014; **32**(49): 6705-10.
42. Harouna Djingarey M. Roll out of the meningococcal A conjugate vaccine through mass vaccination campaigns in countries of the African meningitis belt. 2014. http://www.who.int/immunization/sage/meetings/2014/october/2.DJINGAREY_Session6_SAGE_Oct2014_FINAL_21Oct2014.pdf?ua=1 (accessed 28th November 2016 2016).
43. Wu S, Yang P, Li H, Ma C, Zhang Y, Wang Q. Influenza vaccination coverage rates among adults before and after the 2009 influenza pandemic and the reasons for non-vaccination in Beijing, China: a cross-sectional study. *BMC Public Health* 2013; **13**: 636.
44. Zheng Y, Yang P, Wu S, et al. A cross-sectional study of factors associated with uptake of vaccination against influenza among older residents in the postpandemic season in Beijing, China. *BMJ Open* 2013; **3**(11).
45. Mangtani P, Abubakar I, Ariti C, et al. Protection by BCG Vaccine Against Tuberculosis: A Systematic Review of Randomized Controlled Trials. *Clinical Infectious Diseases* 2014; **58**(4): 470-80.
46. Tameris MD, Hatherill M, Landry BS, et al. Safety and efficacy of MVA85A, a new tuberculosis vaccine, in infants previously vaccinated with BCG: a randomised, placebo-controlled phase 2b trial. *The Lancet* 2013; **381**(9871): 1021-8.
47. Statens Serum Institut. A Safety and Immunogenicity Trial With an Adjuvanted Tuberculosis (TB) Subunit Vaccine in Purified Protein Derivative (PPD) Positive Volunteers (THYB-02) (NCT00929396). January 2013 2009. https://clinicaltrials.gov/show/NCT00929396 (accessed 3rd May 2017).
48. Schaaf HS, Collins A, Bekker A, Davies PDO. Tuberculosis at extremes of age. *Respirology* 2010; **15**(5): 747-63.
49. Aspinall R, Del Giudice G, Effros RB, Grubeck-Loebenstein B, Sambhara S. Challenges for vaccination in the elderly. *Immunity & ageing : I & A* 2007; **4**: 9.
50. Goodwin K, Viboud C, Simonsen L. Antibody response to influenza vaccination in the elderly: a quantitative review. *Vaccine* 2006; **24**(8): 1159-69.
51. Wang J. [Small vaccines contain big market] (In Chinese). *Guide of China Medicine* 2005; **12**: 98-9.

52. Shen X, Yang C, Wu J, et al. Recurrent tuberculosis in an urban area in China: Relapse or exogenous reinfection? *Tuberculosis (Edinburgh, Scotland)* 2017; **103**: 97-104.
53. World Health Organization. Global Tuberculosis Report 2016. 2016. http://apps.who.int/iris/bitstream/10665/250441/1/9789241565394-eng.pdf?ua=1 (accessed 10th May 2017).
54. McShane H. Need for more TB vaccine field sites. *Indian journal of experimental biology* 2009; **47**(6): 445-6.

CHAPTER 5: Informing clinical trial design - the electronic PArticipant Locator (ePAL) app for spatial mapping of TB

Summary of Chapter 5

This chapter reports the development and evaluation results (objective 4) of ePAL (electronic PArticipant Locator), a novel, low cost, rapidly implementable and easy to use app for spatial mapping of the place of residence of TB cases registering for treatment in Blantyre, Malawi. The development and evaluation results are reported as research paper 3, which is in preparation for submission. Preliminary findings of this work were presented at the 2016 TSRU conference. The results of 12 months data collection by NTP TB officers using ePAL in Blantyre, Malawi are summarised in this chapter. This includes spatial heatmaps (objective 4c) of all TB, and demographic and clinical subgroups relevant to TB trial recruitment populations or outcomes.

Citation (paper 3): Harris RC, Kaswaswa K, Choko AT, Molineux A, MacPherson P, Webb E, Shonga W, Mapunga J, Matchaya M, White RG*, and Corbett EL*. Development and evaluation of novel software to identify place of residence for clinic-based TB patients in Blantyre, Malawi. *(proposed authors, author order and title. In review with co-authors. *Joint senior authors)*

This chapter fulfilled objective 4 of the thesis:

a) Develop and implement a novel, low resource, rapidly implementable and easy to use app for spatial mapping of the place of residence of TB cases registering for treatment in Blantyre, Malawi.

b) Evaluate the accuracy of patient coordinates measured via the app compared to the gold standard of measurements taken using GPS at the patient's place of residence.

c) Based upon place of residence collected using the app, generate spatial maps of disease numbers and rates in populations potentially relevant for TB vaccine trial recruitment.

CHAPTER 5 Informing clinical trial design - the electronic PArticipant Locator (ePAL) app for spatial mapping of TB registrations

5.1 Introduction

Translation of tuberculosis (TB) vaccine Target Product Profiles (TPPs) into implementable clinical trials requires local epidemiological data to inform trial design and recruitment. Available prevalence and incidence studies, and country- and regionally-aggregated National Tuberculosis Programme (NTP) data can assist shortlisting of sites with sufficiently high TB rates to conduct a clinical trial. However, as highlighted by the TB Vaccine Blueprint epidemiological research needs (Box 1 Chapter 1), following these preliminary regional-level assessments, local data stratified by likely recruitment characteristics and trial outcomes are required. Surprisingly few trial-ready research sites have sufficiently granular data available to inform identification of high TB risk recruitment populations and to assist estimation of sample sizes that balance cost and risk.[1] Appropriate epidemiological data are essential to ensure that sufficient participants are recruited in to the study to accrue an adequate number of endpoints over a reasonable timeframe.

The mathematical modelling research in Chapters 3 and 4 demonstrated the importance of prevention of disease vaccines for maximising new TB vaccine impact. It was recommended that late-phase clinical trials should focus on disease endpoints, therefore TB disease incidence data are required to inform trial design. In China, older adults were identified as an age group of interest for vaccine implementation targeting. This may vary by epidemiological setting, but knowledge of burden of disease and identification of geographical disease hotspots by age could support development of studies in a given setting. Efficacy trials often exclude HIV-positive individuals, so identifying where high burden of disease exists in HIV-negative populations is also important for trial design and recruitment.

Prospective prevalence and incidence studies are the gold standard to inform these trial design aspects. Where available, such data are the gold standard for informing trial design. However, existing studies may become very rapidly outdated, and the cost and time implications of prospectively initiating these studies ahead of phase IIB and phase III trials would be substantial, especially given that these late phase trials require involvement of multiple research sites (e.g. eight recruitment centres across three countries in the M72 phase IIB trial).[2] Therefore, inexpensive and rapidly implementable alternative methods are needed to collect sufficient data to design late-phase clinical trials.

Existing alternatives and their strengths and limitations are discussed in detail in Chapter 2. Health and demographic surveillance sites can provide high quality, granular data, but there are a limited number of these sites, which are not necessarily in high TB burden locations and would be prohibitively expensive to set up solely for the purpose of TB vaccine clinical trials. NTP data are readily available, but publicly available summaries are generally aggregated at the district or regional level. Data can be extracted from NTP records to achieve clinic-level granularity, but clinic catchment areas may be large (e.g. 100,000 population in India) and many areas of high TB burden are in low-income, urban settings that may be missing municipal address systems.[3] Although a verbal description of place of residence is generally captured in NTP records, these may prove insufficient to identify the patient's area of residence in practice.[4]

Conducting home visits for GPS coordinate collection for patients recruited at NTP clinics could provide the required geospatial information. This could be conducted by travelling directly home with the patient, or recruiting them at the clinic and arranging a later follow up visit at the home. However, the former requires substantial human resourcing and logistics, whereas the latter may experience substantial loss to follow up given the challenges of contacting and finding the participants.

To remove the need for home visits in an HIV testing cluster-randomised trial in Malawi, a map book system for clinic-based capture of cluster residency was developed.[4,5] The map book consisted of a series of paper maps marked with points of interest (POIs) and the cluster boundaries from the parent study to allow identification of study cluster. An evaluation of the tool in 40 participants by conducting home visits to collect GPS coordinates demonstrated excellent study cluster agreement (97.5%).[4] This tool provided proof of concept for the use of point-of-interest annotated maps for remote identification of place of residence. The map book format, however, does not provide the resolution required for designing new individual-level trials, and long-term sustainability would also benefit from a switch to electronic based data collection. The concept was therefore employed as the basis for the design of a new tool, an electronic tablet-based application (app) to more accurately identify a patient's place of residence when registering for TB treatment.

The map book was developed in Blantyre, Malawi, an urban, high TB burden setting, lacking the municipal address system needed for identification of patient place of residence. Malawi has one of the top-20 fastest growing urban populations globally,[6] with an absolute increase of approximately 4% per year.[6] In Blantyre city, the population density has increased from 1,514 per square kilometre in 1987 to 3,006 per square kilometre in the last census in 2008.[7] Overcrowding and poor ventilation may contribute to TB transmission in high density urban slums.[8]

Blantyre's high TB burden, rapid urbanisation, and lack of addresses for geolocating TB cases provided the ideal setting for development and implementation of an electronic clinic-based spatial mapping tool for TB patients. It was hoped that in addition to supporting existing studies in the site, it could gather appropriate data in the hope of conducting TB vaccine trials in this setting in the future.

In this chapter I present the development, implementation and evaluation of a TB spatial mapping app in Blantyre, Malawi, with the aim of mapping the place of residence of populations experiencing disease outcomes relevant to TB vaccine

clinical trials. Based upon place of residence collected using the app, I present maps of trial-relevant outcomes in potential recruitment populations for TB vaccine trials.

5.2 Paper 3: Development and evaluation of novel software to identify place of residence for clinic-based tuberculosis patients in Blantyre, Malawi

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	Rebecca Claire Harris
Principal Supervisor	Richard White
Thesis Title	Development and evaluation of novel software to identify place of residence for clinic-based tuberculosis patients in Blantyre, 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?	TBC
Please list the paper's authors in the intended authorship order:	Rebecca C. Harris, Kruger Kaswaswa, Augustine T Choko, Peter MacPherson, Andrew Molineux, Emily Webb, James Mpunga, Medson Matchaya, Wisdom Shonga, Richard G White, and Elizabeth L Corbett,
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)	See next page

Student Signature: R Harris **Date:** 29/9/17

Supervisor Signature: Richard White Date: 29/9/17.

Authors' Contributions: RCH was the study principal investigator. KK and AC were local study leads, and Frank Chipaka and George Sinjani provided project management support. RCH, LC and PM developed the research question. RCH, LC, RW, PM and EW conceptualised and designed the study. AM developed the ePAL software. Data and POI management was conducted by WS and RCH. MLW server/data management was supported by Lingstone Chiume, David Matiya, Vincent Phiri and McEwan Khundi. RCH and EW conducted data cleaning and analysis. Production of monthly reports was conducted by RCH until May 2015, beyond which Vincent Phiri assumed responsibility for generating the reports. POI collection, monitoring and evaluation visits were conducted by the field team: Ken Kaswaswa, Hygiene Kumwenda, Elizabeth Chazungu, Eluby Kaunda and Singatiya Nkute. Data collection was conducted by TB officers and nurses of the NTP in Blantyre, Malawi. RCH wrote the first draft of the manuscript. I am incredibly grateful to all of the TB officers, community health workers and ePAL study participants for contributing to this research.

Development and evaluation of novel software to identify place of residence for clinic-based tuberculosis patients in Blantyre, Malawi

Authors: Rebecca C. Harris,[a,*] Kruger Kaswaswa,[b,c] Augustine T Choko,[a,c] Peter MacPherson, [d] Andrew Molineux,[e] Emily Webb,[a] James Mpunga,[f] Medson Matchaya,[g] Wisdom Shonga,[c] Richard G White,[a,**] and Elizabeth L Corbett,[a,c,**]

[a] London School of Hygiene and Tropical Medicine, Keppel Street, London, UK

[b] National Tuberculosis Control Programme, Community Health Sciences Unit, P/Bag 65, Lilongwe, Malawi

[c] Malawi-Liverpool-Wellcome Trust Clinical Research Programme, PO Box 30096 Chichiri, Blantyre 3, Malawi

[d] Liverpool School of Tropical Medicine, Pembroke Place, Liverpool, UK

[e] Tripod Software, Salford Innovation Forum, 51 Frederick Road, Salford, M6 6FP

[f] National Tuberculosis Control Programme, Community Health Sciences Unit, P/Bag 65, Lilongwe, Malawi

[g] District Health Office, P/Bag 66, Blantyre, Malawi

*Corresponding author

** Joint senior authors

Keywords: tuberculosis, spatial, epidemiology, slum health, Malawi, App, ePAL

Abstract

Background: In many low-income, urban settings with high tuberculosis (TB) burden, the absence of municipal address systems poses a significant obstacle to clinical trial design, patient follow up, and targeting of public health interventions. We developed and evaluated software for use by Blantyre District Health Office (DHO), Malawi, to remotely capture place of residence for registering tuberculosis patients.

Methods: High-resolution satellite maps were annotated with points of interest (POIs) captured by community health workers and loaded onto electronic tablets for use with a bespoke 'electronic PArticipant Locator' application (ePAL app; Tripod Software, Salford) combining touch-screen Global Positioning System (GPS) coordinate capture with an electronic tuberculosis register. Following training, ePAL was used by tuberculosis officers from 11 primary care clinics for collection of demographics, health status and coordinates of place of residence for consenting adults (≥18 years) initiating TB treatment. Case capture and data quality were monitored against DHO registers. Accuracy of ePAL-recorded co-ordinates was evaluated against GPS coordinates obtained at the place of residence.

Findings: Community health workers defined 3,243 POIs. Over 12 months from 12th February 2015, 1,899 TB patients (98.7% participation) were registered using ePAL. Case-capture increased from 61% to 95% from months 1 to 12 of the study. GPS coordinates captured at evaluation home-visits with 121 participants were median 84 metres (IQR: 35-317 metres) from those captured through ePAL. Advantages of the ePAL system included real-time availability of high-resolution spatial data, and ease of use by DHO staff. ePAL continues to be used as part of routine TB registration in September 2017, supported by monthly DHO meetings.

Interpretation: ePAL enabled sustainable and reasonably accurate capture of place of residence of patients presenting at TB clinics in areas lacking formal addresses. Data generated by ePAL could have both clinical trial and programmatic utility,

through facilitating patient follow up and informing spatially targeted public health interventions and clinical trials. With local annotation, ePAL is readily adaptable for spatiotemporal mapping and home tracing in other settings.

Funding: UK Medical Research Council and Wellcome Trust.

Article

Background

Tuberculosis (TB) was the cause of death of 1.8 million people, including 400,000 people living with HIV (PLHIV) in 2015.[9] Despite being such a major cause of death, the high costs and development risks associated with TB prevention studies, such as trials of new TB vaccines or complex community-based interventions, are prohibitive and rate-limiting.[1] This reflects the combination of the suboptimal and costly diagnostics, TB incidence and prevalence rates that even in high burden settings necessitate large sample sizes, and a prolonged and highly variable interval between transmission and disease.[1]

Robust epidemiological data could help identify high burden populations to direct trial recruitment, assist sample size minimisation, and help de-risk study endpoint accrual.[1] Public health programmes could also benefit from such data for design and targeting of local TB prevention and care programmes. However, the prohibitive factors affecting clinical trials often also preclude formal prospective cohort studies to inform clinical trial design and site selection ahead of trial initiation. They also mean that available disease burden estimates are prone to being incorrect, out-dated, or not available for otherwise suitable trial sites.[1,10]

National Tuberculosis Programmes (NTPs) collect an abundance of relevant data, but often report TB disease burden as data aggregated by administrative units that are either too large (e.g. districts) or poorly defined (e.g. facility catchment areas) to fully inform planning. Methods of recording patient place of residence in NTP registers often lack the clarity and reproducibility required to support disaggregation.[4,10,11]

Africa and Asia are urbanising rapidly (>1% per year).[6] Inadequate town planning capacity has left more than half of city streets in sub Saharan Africa without a formal address system, leaving residents to rely on imprecise and poorly-reproducible

landmark-based directions to communicate where they live.[3,4] Poverty, overcrowding and poor ventilation in urban, high-density, informal settlements may contribute to increased risk of *M.tb* transmission and TB disease.[8,12] This confluence of urbanisation and TB risk has resulted in higher tuberculosis prevalence in urban than rural areas at the global level.[12] For example, a study in the Philippines estimated the TB risk in children in 'slum' areas was nine times higher than in 'non-slum' areas.[13] These high burden urban areas present a challenge and an opportunity: a challenge to manage TB control effectively, but with improved data could present an opportunity to target public health interventions and clinical trial recruitment to high burden settings to maximise impact and reduce cost.

Research approaches outside of sites with pre-existing demographic health surveillance systems have included locally drawn street maps, and home visits, which are expensive and time consuming.[10] The current study builds on a demonstration study in 2011,[4,5] which provided proof of concept for the clinic-based use of maps annotated with points of interest (POIs) for identification of place of residence in urban settings. In this study, a paper-based "map book" was used by DHO staff to assist patients in identifying whether they were resident in one of 28 research study catchment areas (~1,300 adults) in Blantyre, Malawi.[4] The study found excellent agreement between cluster-residency defined in the clinic using the "map book" and cluster residency obtained through home-visits (97.5%, n=40).[4]

The current study sought to broaden the utility of this tool beyond cluster randomised studies and improve sustainability, by seeking to digitise the system and support collection of specific coordinates of the place of residence. Solutions to allow real-time electronic data collection and improved data on place of residence as part of National Tuberculosis Programmes could provide long-term benefit to the public health system through informing patient follow up, improving efficiency of data collection, and targeting of TB control programmes,[10] and generate the information required to select clinical trial locations and populations to minimise cost and risk.

Here we describe the development, implementation and evaluation of the accuracy of a novel electronic PArticipant Locator (ePAL) application for use on electronic tablets for collection of Global Positioning System (GPS) coordinates of the place of residence of TB patients registering for TB treatment in urban Blantyre, Malawi. The aim was to provide an electronic TB register that could support long-term, real-time capture of high-resolution spatial information as part of routine TB registration activities in urban Blantyre, Malawi.

Methods

Study setting and design

The study was conducted in urban Blantyre, Malawi, a city with large areas of high-density informal residential settlements and an estimated population of approximately 750,000 in 2016. All 11 TB registration centres that served Blantyre City participated in the study, including two public and three private hospitals, and six local primary care centres. TB diagnosis, registration and care was provided exclusively through these centres, which relied on paper-based registers with brief recording of verbal descriptions given by patients of their place of residence.

The study was conducted prospectively (see appendix A for protocol) in three phases: 1) app design, 2) implementation and run-in, and 3) the main study and evaluation phase. The data reported here relates to the implementation, monitoring and evaluation aspects of the study.

ePAL App Design

The ePAL app combined an electronic case report form (eCRF) with high resolution satellite maps of urban Blantyre (approx. area: 254km^2) and community-identified points of interest (POIs) (Figure 5.1). When the location was selected via touch-

screen, it was converted in-app and stored as the corresponding WGS84 (EPSG: 4326) coordinates.

Maps were purchased from European Space Imaging.[14] POIs were collected by study staff and Ministry of Health (MoH) Community Health Workers (CHW). Staff used Garmin eTrex 20 Global Positioning System (GPS) handsets to capture POIs that were identified by CHWs as being used for local address-finding. A pilot study of one CHW area informed the target POI density, and set a minimum density of 10 to 20 POIs, per catchment area.

To ensure uninterrupted ePAL use offline, the app, maps and POIs were stored on the tablet hard-drive. TB officers used the app to collect participant demographic and health information, and select the participant's place of residence on the annotated maps. Following several iterations, features of the app to facilitate prompt and accurate identification of the place of residence included high resolution zoom, searchable community-identified POIs, colour coding of POIs by type (e.g. place of worship, shop), and a drop-down menu of city wards.

The app runs on Android devices: Asus Google Nexus 7-inch or 9-inch tablets (2GB RAM, 16GB eMMC) were used for the implementation study. Data entry was facilitated through use of simple instructions in a choice of English or Chichewa, and large buttons for data entry. Verbal consent and checks against inclusion and exclusion criteria were in-built. Sputum sample barcodes were scanned using the tablet's camera to allow linkage to subsequent laboratory results.

Patient data, specimen bar-code and GPS coordinates were stored under encryption on the tablet, and uploaded daily by the TB officers via the 2G/3G mobile network using comma delineated files. The central database was held on a server at the Malawi-Liverpool-Wellcome Trust clinical research site, accessible in real-time from any location via a password-protected web interface.

Implementation and run-in period

Following a one-day training session for all 44 TB officers and nurses working in Blantyre's TB clinics, implementation of the ePAL app began in November 2014. Each clinic received a password-protected and encrypted electronic tablet with ePAL installed. The study commenced with a 10-week run-in period preceding the main study. During the run-in period, app user interface improvements were incorporated based upon suggestions gathered at regular meetings with field users.

Main study

The main study and evaluation period began on 12th February 2015 and ran for 12 months. All adult patients (≥18 years old) registering for TB treatment in any of the 11 clinics serving Blantyre city were invited to participate in the main ePAL study, regardless of place of residence. For those providing verbal consent to participate, demographic, clinical information, and place of residence were collected using ePAL.

Data quality monitoring was continuous throughout the study. Uploads to the ePAL database were regularly monitored via the web interface. Any unusual gap in uploads triggered a call or visit by study staff. Clinics were visited monthly to check for hardware, software or mobile data connectivity issues. During this visit, the number of ePAL uploads was checked against the NTP paper registers, and seven fields (e.g. sex, smear status) were checked against the NTP register for a random 10% sample of ePAL records (minimum one record per clinic) selected prospectively using Stata. Monitoring data were collected on a Nexus 5 tablet using Open Data Kit (ODK).

Monthly meetings were held with TB officers to provide a forum for collaborative troubleshooting hardware or software issues and feedback of clinic-level ePAL summary reports, including data quality targets such as case capture and zoom level.

Evaluation cohort

Recruitment to the evaluation cohort was limited to patients who reported living in one of three high-density neighbourhoods of Blantyre (Figure 5.2), described elsewhere.[5] Evaluation in these settlements tested accuracy in the highest density, and therefore most challenging, areas in which to ascertain location of residence using ePAL. GPS coordinates collected at the participant's place of residence were used to evaluate the spatial accuracy of the ePAL-collected coordinates. A study research assistant was guided to the residence by the participant, and collected an averaged waypoint GPS coordinate location using Garmin eTrex 20 handsets. The evaluation cohort consisted of patients providing written (or thumbprint) informed consent and registered in ePAL during the first four months of the main study. Home visits were conducted in the run-in period and after the first four months of the study for monitoring purposes.

Statistical analyses

Statistical analyses were conducted in Stata (version 13.1) and R (version 3.3.1).[15] Descriptive analyses of the recruitment population for the main study were reported for the 12 months of the study. For the evaluation cohort, the 'spDists' command within the 'sp' R package was used to calculate the distance in metres between the ePAL and evaluation GPS spatial point for each participant using the great circle distance method (WGS84 ellipsoid). ePAL accuracy was assessed by the median and interquartile range of the distances in metres between the paired ePAL and GPS coordinates.

The cumulative proportions accurate to within 20m, 50m, 100m, 200m, 500m and 1000m were estimated with 95% confidence intervals (CI). The target sample size for the evaluation cohort was 196, which would provide 5-7% precision when estimating proportion agreement with a point estimate of 50-85%. Time between date of ePAL record and date of evaluation data upload (latest possible follow up date) were

calculated. Predictors of accuracy and missing data were explored using Spearman's correlation, Kruskal-Wallis rank sum, and Chi-squared tests, as appropriate.

Ethical considerations

Ethics approval for this study was granted by the research ethics committees of the University of Malawi College of Medicine (COMREC) and the London School of Hygiene and Tropical Medicine (LEO). All participants in the main study provided verbal informed consent, and participants in the evaluation cohort also provided written or witnessed thumbprint informed consent.

Results

During September to November 2014, 194 CHWs in Blantyre identified 3,243 POIs (mean of 17 per CHW) following community engagement. In the ePAL app, 962 of these POIs were prioritised to be visible at all levels of map zoom, and the remainder visible only once further zoomed in, to avoid obscuring underlying map features when zoomed out. Feedback from TB officers was used to improve functionality and ePAL user experience during two months of pilot data capture in clinics.

Study population

Flow of study participants is shown in Figure 5.3. For 12 months from 12th February 2015, overall case capture with ePAL was 77% of the patients recorded in the NTP register (1,924/2,489), with capture increasing from 61% to 95% over months 1 to 12, once logistical challenges had been resolved relating to weekend registrations and need for additional electronic tablets in the busiest clinics. 1,899 (98.7%) of 1,924 eligible adults agreed to participate, with the main reasons for non-participation being not wishing to participate in research (n=12), time pressure (n=3), not wanting data recorded (n=1) or place of residence (n=1) captured electronically. Eight provided no reason for non-participation.

The majority of registrations occurred in the two public hospitals (50.7%, n=962 in Queen Elizabeth Central Hospital [QECH] and 13.6% in Mlambe Hospital), 4.4% (n=83) of registrations were from private hospitals, and 32% from six primary care clinics. Thirty-four TB officers collected data in ePAL over this period, with more than half of all records collected by six TB officers.

Participants were mostly male (n=1,167, 61.4%), and mean age was 37.9 years. Sputum smear microscopy results were available for 56.9% (n=1,080) of participants, of whom 58.8% (n=636) were smear positive. Overall, 43% of patients had microbiologically confirmed disease (microscopy, Xpert or culture) at time of registration for TB care. The majority were classified as previously untreated 'new' TB cases (n=1,661, 87.5%), with 5.6% relapse, 6.3% 'other', and <1% were treatment failure or return after default. For those reporting cough, median duration was 4 weeks (IQR: 3-5 weeks).

For all patients registered with ePAL, self-reported residence in urban Blantyre was 73% (1,385/1,899), with 19% of patients residing in rural Blantyre (353/1,899), and 8% out-of-District (161/1,899). A total of 1,240 participants provided coordinates in ePAL (Figure 5.3), of which 1,229 reported residence in urban Blantyre. Monthly estimates of coordinate capture for those self-reporting residency in Blantyre ranged 77% (95% CI: 68-85%) to 95% (95% CI: 89-98%). Two percent of those reporting residence outside of urban Blantyre had ePAL coordinates falling within urban Blantyre boundaries. Of those missing coordinates (n= 659), 503 were resident outside of urban Blantyre, and 156 reported as resident in Blantyre but were missing coordinates.

Of the 1,240 participants with ePAL coordinates, 514 (41%) self-reported residence in the evaluation neighbourhoods. Only 13 (3.4%) of these participants refused participation in an evaluation visit. A total of 252 evaluation visits were conducted, of which 121 participants were recruited in to ePAL within the 4 month evaluation

period, and the remainder were conducted for monitoring purposes, either during the run-in period (n=48) or after the evaluation period (n=83).

Monitoring

Ninety-six percent of coordinates were collected at the two highest resolution zoom levels. Monthly capture at the highest zoom level ranged from 38% to 84%, increasing over time. Monitoring checks of seven variables for 131 patient records identified a 1.5% error rate, with 79% of errors relating to erroneous allocation of gender. In addition to the planned checks, minor errors were reported in 25 NTP registration IDs.

Evaluation

In the four month evaluation period, 171 participants self-reporting as resident in the evaluation area provided written consent to participate. Follow up of these participants with successful collection of GPS coordinate at the home was 70.8% (n=121).

The median distance between the ePAL-recorded coordinates of place of residence and the GPS reading taken at the evaluation visit for these 121 patients was 84 metres (Interquartile range (IQR): 35m-317m). The minimum estimated distance was 4m, and the greatest was 11.6km, with the participant having GPS-recorded residency outside of the evaluation area. Accuracy to within 50m was 35.5% (95% CI: 27.0-44.8%). Accuracy to within other ranges are shown in Table 5.2.

In the 4-month evaluation period of the study, the target sample size of 196 home visits was not reached. With a reduced sample size of 121, the study was powered to give precision of 6.5%-9%. Combining the evaluation and monitoring visits during the 12 months of the study, 204 participants were followed up.

In the evaluation cohort, median distance by registration centre was 54m (IQR: 24-139m) and 89m (IQR: 36-219m) in the two primary care clinics, 88m (IQR: 48-420m) at QECH, 292m (IQR: 135-951m) at Mlambe hospital and 16m for the one patient recruited at Adventist Hospital. Kruskal-Wallis rank sum tests indicated no evidence of a difference in the evaluation distance distributions between health centres (p=0.11), although there was evidence of a difference between data collectors (p=0.03).

Fourteen percent of evaluation visits identified a place of residence more than 1km from the location identified in ePAL. Pearson's chi-squared tests were conducted to identify characteristics predictive of a large displacement (>1km) between the ePAL and evaluation GPS coordinates. None of 20 possible explanatory variables were significantly associated with large displacement, including study month. A Spearman's rank test for correlation between month and distance was also non-significant (p=0.13).

In the larger monitoring data set for the full 12 months of the study, a significant relationship between large errors (>1km) and patient-level factors, notably HIV status, raised the question of whether some patients had deliberately misinformed TB officers about the true location of their household, motivated by TB and/or HIV stigma. Informal interview during home visits identified various patient-level explanations, for instance staying with relatives for diagnosis. However, these data were not fully systematically captured.

Median time lag between ePAL entry and evaluation data upload (latest possible follow up date) was 26 days (IQR: 14-37 days). The Spearman's correlation between the evaluation distance and data collection lag estimate was 0.12 (p=0.17), suggesting negligible correlation between delays in follow up and discrepancy between the ePAL and GPS measurements.

Chi-squared tests suggest a difference in availability of data from evaluation visits by TB officer (p=0.01) and number of people in household (p=0.04). Chi-squared tests were also conducted with the full 12-month data set, in which the significance of any differences between the TB officers became weak.

Discussion

The main findings of this study, in which a novel spatial application, ePAL (electronic PArticipant Locator), was developed, implemented and evaluated in collaboration with the National TB Programme and District Health Office in Blantyre, Malawi, was that this provided a rapidly implementable and sustainable approach to remote capture of place of residence for patients with no formal address system.

Monthly feedback meetings supported troubleshooting and near complete case-capture within a few months of implementation. Real-time collection of GPS coordinates was sufficiently accurate for most patients to support real-time spatial analysis of disease burden, and interventions such as household contact screening. Home visits showed substantial errors in household identification for about 1 in 7 patients, however, reflecting issues such as multiple places of residence, but also likely deliberate misidentification by some patients. Data generated by ePAL has high potential research and programmatic utility, and could facilitate spatial epidemiology, design and implementation of public health interventions and clinical trials, and patient follow-up.

Strengths of ePAL for collection of spatial data include its accuracy, ease of use, low cost, integration into routine NTP data collection with QAQC processes in place, real time data accessible via a web interface, and ability to function offline with infrequent access to mobile data and electricity. A key feature of the app was tailoring of content and the user experience through stakeholder involvement: for example, community engagement for identification of locally-relevant POIs in the initial development, and ongoing feedback and improvement of the app and data collection through regular

meetings with the TB officers. This may have also contributed to very high participation rates. The app interface and training were designed for ease of use by individuals without experience of electronic data collection. Error rates were low for single-entry electronic data capture. Implementation was rapid and maintenance was relatively low in cost and staff time.

This tool can be adapted for use in different settings, with the main investment for introduction into new locations being purchase of high resolution satellite maps (~USD$ 1,000 per 100 km^2), and collection of local points of interests needed to fully annotate the map and make it more readily usable. In this study POI capture involved Ministry of Health CHWs, plus study staff time over seven weeks of POI collection. One full time field staff member was employed for ongoing monitoring and evaluation visits, which is likely to be essential for maintaining as well as understanding accuracy and case-capture. Our use of tablets instead of computers kept hardware costs for electronic data capture low, and avoided reliance on unstable sources of electricity and internet. With the support of the District Health Office and NTP, most data were collected routinely by TB officers during patient registration activities. TB officers received a small per diem for attending monthly feedback meetings, but no other compensation. One staff member was employed to assist data collection in the busiest central hospital.

ePAL-based data collection was successfully incorporated in to the National Tuberculosis Programme through this integration in to routine patient registration. Initial computer and map literacy was low for some TB officers, with many never having used an electronic device before, but with one full-day training, and regular support through meetings twice per month (one in the clinic and one at the DHO), NTP staff were able to easily and effectively use the ePAL app for collection of patient data and location of residence. Initial differences between TB officer in terms of ePAL accuracy diminished over time. Results of the study, including health centre-level reports, were provided to the NTP and TB officers on a monthly basis.

We estimate median distance between the ePAL coordinates and GPS-measured coordinates to be 84m (IQR: 35-317m). The median would correspond to approximately 6 to 8 dwellings from the patient's residence in the highest density areas, but as close as the adjacent structure in low-density suburbs. The accuracy of GPS measurements at the home was maximised by taking averaged waypoints, reducing the inherent inaccuracy of measurement to a few metres, although this will have been increased by cloud cover, rain, and proximity to buildings. Home-visit and ePAL GPS readings used slightly different reference points (from the front of the building at home-visit, but centre of the building image on the satellite maps), introducing unavoidable 'inaccuracy' even when fully correct. The evaluation was conducted in the highest density slum areas of the city and included patients registered in clinics remote from these neighbourhoods, with no obvious effect of proximity from the clinic (site of ePAL capture) on accuracy.

Evaluation was focused on patients registering in the first 4 months of the study, with smaller numbers recruited for on-going quality assurance thereafter. As such, we cannot comment on accuracy trends over time, which are likely to have improved with increasing user-familiarity with ePAL and periodic improvements to the software over the course of the first year. Conducting home visits was more time consuming than anticipated, affecting both timeliness and final numbers followed up. There was no obvious effect of longer intervals between registration and home visit on accuracy.

A displacement of >1km was recorded between the ePAL and home visit coordinates for 14.0% of evaluation participants. Feedback from the field team suggested that such large differences were mostly due to relocation, for instance temporary residence with a relative to seek TB care, and also as a result of heavy rainfalls and flooding in early 2015. Over the longer study period, large displacement was also significantly associated with HIV status, raising deliberate misidentification because of HIV stigma as a possible cause.

Limitations

The main limitation was the difficulty of conducting home visits in high density informal settlement areas, which followed verbal description given to TB officers and recorded in the paper NTP register (the standard approach before ePAL). Use of ePAL coordinates was avoided to prevent observer bias. Tracking down the correct households proved very time-consuming, leading to some long delays between registration and home visit (IQR: 14-37, maximum 182), and was completely unsuccessful in 50 (29.2%) of patients registering in the four-month evaluation period. While these issues underscore the need for alternatives such as ePAL, it is possible that incomplete capture may have biased our estimates of accuracy if, for instance, deliberately incorrect information had been provided both for ePAL and the verbal address. A more systematic classification of underlying issues resulting in large distance errors, ideally including qualitative interviews to probe for HIV/TB stigma as well as to identify multiple residencies, would have provided better understanding of the cause of these differences and extent to which they could be minimized. ePAL case capture issues early in the study were rectified by provision of additional support and electronic tablets for data collection in the QECH hospital.

Conclusion

In summary, ePAL provides a pragmatic digital tool that can be integrated into the National Tuberculosis Programme to collect place of residence for patients living in settings without a municipal address system. These data provide broad programmatic and research benefits, including design of preventative clinical trials for TB. The system is readily adaptable to other diseases, and can be implemented in any setting where satellite maps can be purchased for offline use and annotated with local points of interest.

List of abbreviations

App	Application
CHW	Community Health Worker
CI	Confidence intervals
cRCT	cluster Randomised Controlled Trial
DHO	District Health Office
eCRF	electronic Case Report Form
ePAL	electronic PArticipant Locator
GPS	Global Positioning System
IQR	Interquartile range
M.tb	*Mycobacterium tuberculosis*
MoH	Ministry of Health
NTP	National Tuberculosis Programme
ODK	Open Data Kit
POI	Point of Interest
TB	Tuberculosis
UR	Uncertainty Range

Figures and Tables

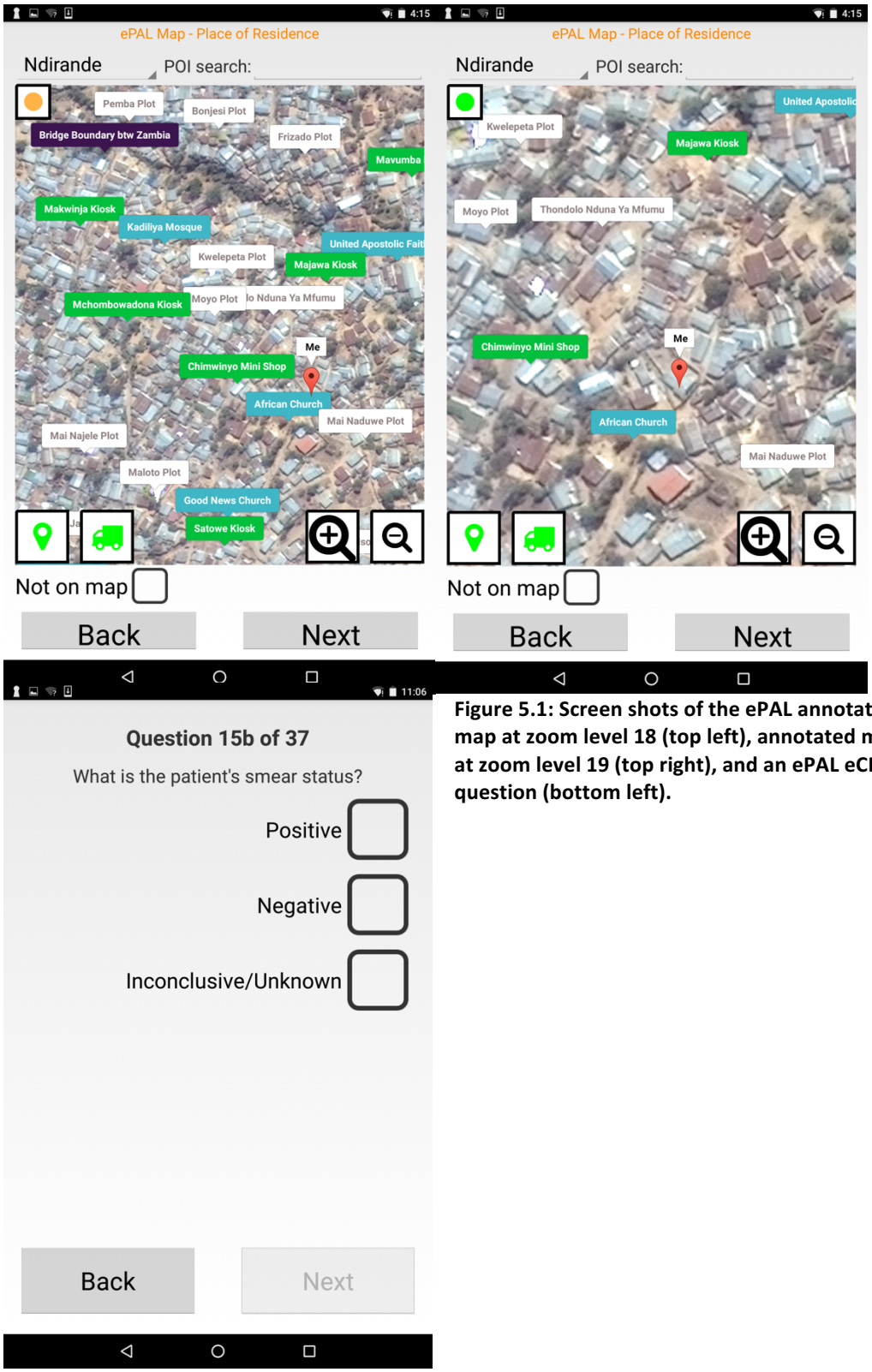


Figure 5.1: Screen shots of the ePAL annotated map at zoom level 18 (top left), annotated map at zoom level 19 (top right), and an ePAL eCRF question (bottom left).

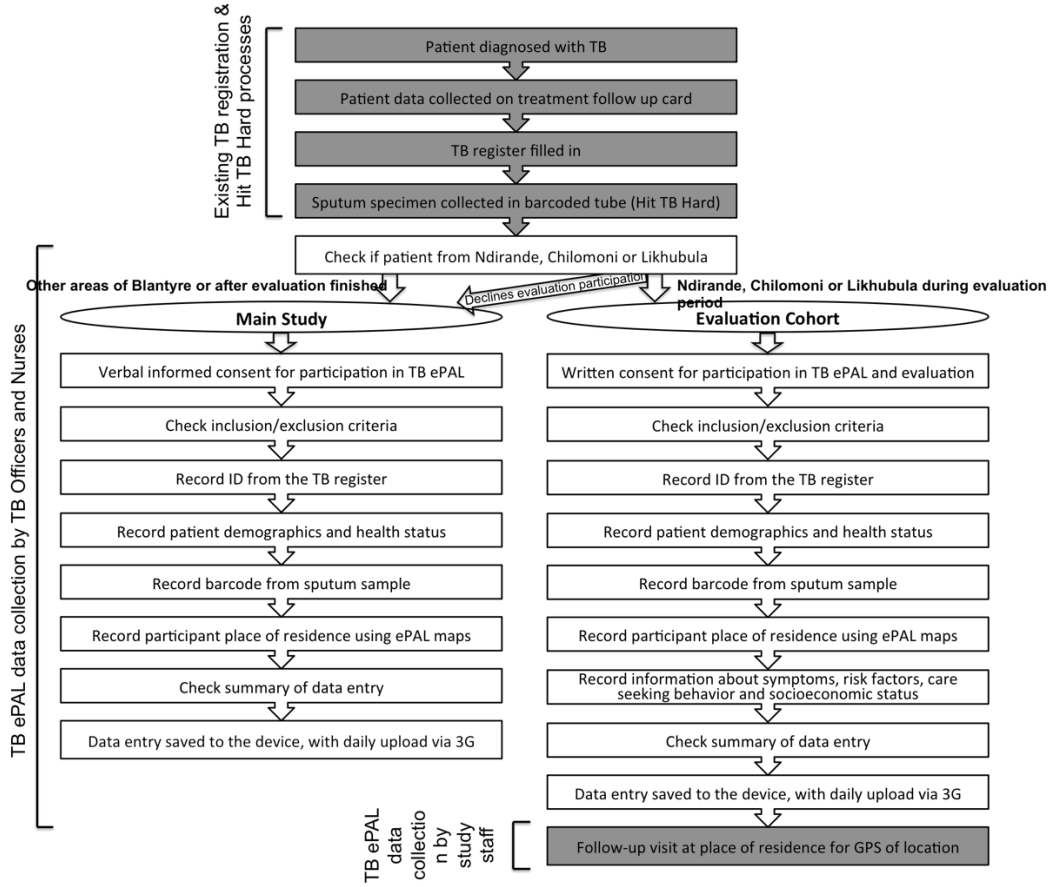


Figure 5.2: Outline of study processes for main study and evaluation cohort

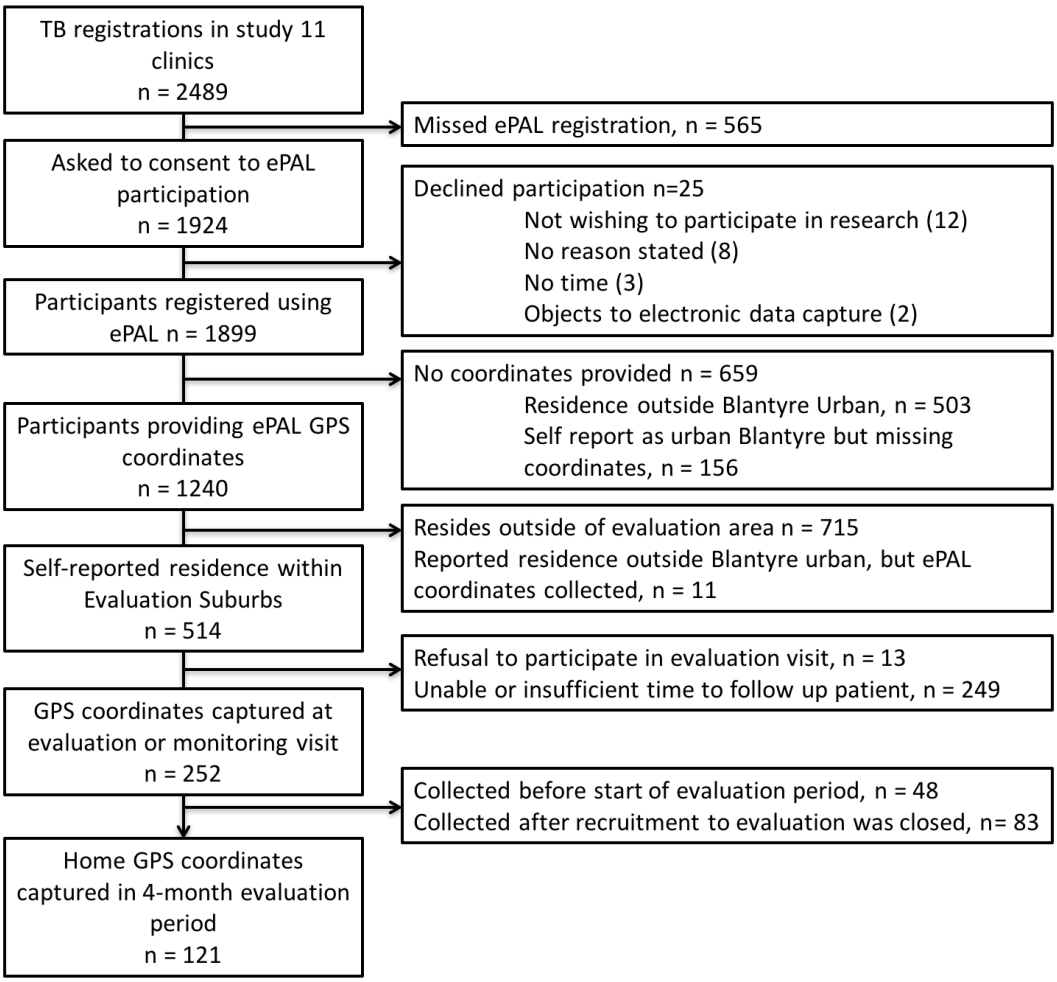


Figure 5.3: ePAL participant study flow

Table 5.2: Proportion of evaluation participants by estimated distance between ePAL and GPS measurement

Distance range (metres)	Number within distance range	Proportion within distance range (95% CI)	Cumulative number within distance	Cumulative proportion within distance (95% CI)
0 ≤ m < 20	16	13.2% (7.8-20.5)	16	13.2% (7.8-20.5)
20 ≤ m <50	27	22.3% (15.2-30.8)	43	35.5% (27.0-44.8)
50 ≤ m < 100	24	19.8% (13.1-28.0)	67	55.3% (46.0-64.4)
100 ≤ m < 200	18	14.8% (9.1-22.5)	85	70.2% (61.2-78.2)
200 ≤ m < 500	12	9.9% (5.2-16.7)	97	80.2% (71.9-86.9)
500 ≤ m < 1000	7	5.8% (2.4-11.6)	104	86.0% (78.5-91.6)
1000 ≤ m	17	14.0% (8.4-21.5)	121	100.0% (97.0-100*)

* One-sided 97.5% confidence interval

References

1. World Health Organization. Global Tuberculosis Report 2016. 2016. http://apps.who.int/iris/bitstream/10665/250441/1/9789241565394-eng.pdf?ua=1 (accessed 10th May 2017).
2. McShane H. Need for more TB vaccine field sites. *Indian journal of experimental biology* 2009; **47**(6): 445-6.
3. Seebregts C, Hosseini M, Bleed D, Williams BG, Dye C. Developing electronic tuberculosis recording and reporting systems: summary report of a WHO expert consultation meeting. 2006. http://www.who.int/tb/err/err_expertmeeting_report_jul06.pdf (accessed 10th September 2017).
4. Subbaraman R, Thomas BE, Sellappan S, et al. Tuberculosis patients in an Indian mega-city: Where do they live and where are they diagnosed? *PloS one* 2017; **12**(8): e0183240.
5. MacPherson P, Choko AT, Webb EL, et al. Development and validation of a global positioning system-based "map book" system for categorizing cluster residency status of community members living in high-density urban slums in Blantyre, Malawi. *American journal of epidemiology* 2013; **177**(10): 1143-7.
6. United Nations Department of Economic and Social Affairs Population Division. World Urbanization Prospects: The 2014 Revision, (ST/ESA/SER.A/366). 2015. https://esa.un.org/unpd/wup/CD-ROM/ (accessed 20th August 2017).
7. Farvaque-Vitkovic C GL, Leroux H, Verdet F, and Chavez R. Street Addressing and the Management of Cities. 2005. http://siteresources.worldbank.org/CMUDLP/Resources/461753-1160058503655/Street_Addressing_Manual.pdf?resourceurlname=Street_Addressing_Manual.pdf (accessed 2nd March 2017).
8. Hargreaves JR, Boccia D, Evans CA, Adato M, Petticrew M, Porter JDH. The Social Determinants of Tuberculosis: From Evidence to Action. *American journal of public health* 2011; **101**(4): 654-62.
9. Stop TB Partnership. Key Populations Brief: Urban Populations. http://www.stoptb.org/assets/documents/resources/publications/acsm/KP_Urban_Spreads.pdf (accessed 2nd September 2017).
10. Fry S, Cousins B, Olivola K. Health of children living in urban slums in Asia and the near east: Review of existing literature and data. 2002. http://www.ehproject.org/PDF/Activity_Reports/AR109ANEUrbHlthweb.pdf (accessed 2nd September 2017).
11. Corbett E. Intensified HIV/TB prevention linking home- based HIV testing, including the option of self-testing, with HIV care. 2012. http://www.controlled-trials.com/ISRCTN02004005/ (accessed 13th August 2017).
12. European Space Imaging. European Space Imaging website. 2014. http://www.euspaceimaging.com/ (accessed 14th October 2014).
13. R Core Team. R: A language and environment for statistical computing. . R Foundation for Statistical Computing, Vienna, Austria; 2014.

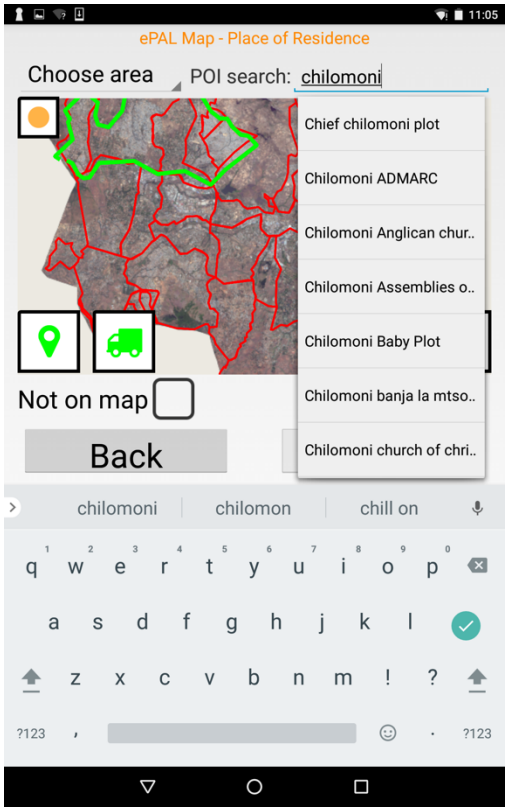
5.3 Additional methods

5.3.1 Additional ePAL app features

Some additional app details and map features not discussed in section 5.2 are summarised in Table 5.3 below.

Table 5.3: Summary of additional ePAL app features

ePAL App Feature	Function	Description
POI control	Button to turn POIs on/off	On the ePAL maps, pressing the  button turns the POI labels on and off. This is of use if a participant's place of residence lies underneath one of the POI labels.
Road controls	Button to turn roads and area boundaries on/off	On the ePAL maps, pressing the  button turns the road and boundary markings on and off.
Zoom	Grid-based zooming	Map zooming is grid-based using the  buttons (as opposed to 'pinch' action). This was selected as the most straightforward option for inexperienced users.
Zoom traffic lights	Indicate sufficient zoom level	Although proportion collecting coordinates at a high level of zoom (at least zoom 18) was good throughout the study, to encourage collection at zoom 19, a traffic light system was introduced in early 2016. The  symbol in the top left corner of the map changes to  at zoom 19 (see Figure 5.4).
Summary screen	Allow check of collected data	The final screen of the app contains a summary of all of the collected data, allowing the TB officer to return to the eCRF if any data are incorrect
Settings section	Language and TB centre selection	Language can be switched between English and Chichewa in the settings section of the app. TB centre is selected in the settings to allow correct auto-generation of the ePAL participant reference (as this starts with the centre code). Initials of new data collectors to appear as options in-app can be added in the settings section.



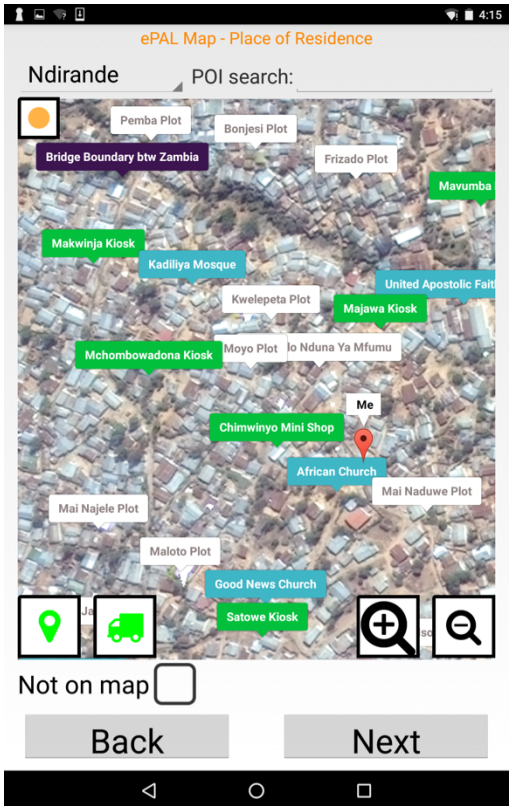


Figure 5.4: Screen shots of ePAL app

Top left: ePAL map fully zoomed out (zoom 12) demonstrating POI search function. Red boundaries are wards, and the green boundary is the ePAL evaluation area.

Top right: ePAL map at zoom level 18. Different POI categories are represented by different colours (e.g. green = shop, blue = place of worship, purple = bridges and public structures, white = plots and other).

Bottom left: ePAL map at full zoom (level 19) demonstrating the high density of structures and high coverage of POIs. When a place of residence is selected by the user, the red pointer labelled "Me" shows the selected location, and can be moved if placed incorrectly. If the participant doesn't live in urban Blantyre, "not on map" can be selected.

5.3.2 Pilot of POI density

In the paper-based map book study, an average of 9 POIs (range: 3-13) were marked on maps of each cluster with a mean adult population of 1,324.[4] Higher target resolution in the ePAL study was expected to require a higher POI density. A pilot was conducted with 'map book'-style maps to identify an appropriate POI density for the ePAL maps.

Printed maps with one of three POI densities (10, 15 and 20 POIs per HSA area) were field tested. Eighteen randomly selected adults attending a market place in the annotated area were invited to participate. Those verbally consenting to participate were asked to identify and mark their place of residence on a printed map (six participants for each POI density), then accompanied by a field worker to their place of residence for collection of GPS coordinates (Garmin eTrex 20, Garmin International, Inc., Olathe, Kansas). The paper-based location and GPS coordinates were compared visually, and proximity (yes/no) agreed by consensus between two reviewers (RH and WS). All participants using maps with 20 POIs were assessed to have achieved close proximity, whereas only 4/6 of participants using maps with 10 POIs were considered to have achieved close proximity. Therefore, target density of 20 POIs (minimum 10) per HSA area was set.

POIs were collected by a team of field workers between 23rd September and 14th November 2014 using Garmin eTrex 20 handsets (Garmin International, Inc., Olathe, Kansas). Additional POIs were added as necessary to areas identified by TB officers. In total, ePAL currently contains 3,243 POIs, averaging 17 per HSA area. ePAL POIs can be managed via the web interface, allowing future POI updates.

5.3.3 GPS accuracy

Inherent inaccuracies in GPS receivers and methods of GPS collection affect the accuracy of the gold standard GPS measurement used for comparison to ePAL coordinates.

The U.S. government guarantees satellite error will remain below 7.8m, and in May 2016 was estimated as below 0.7m with 95% probability.[16] Inherent error in Garmin receivers are estimated as "typically accurate to within 10 meters".[17] Combined, these two contribute to an overall GPS reader error of approximately 10m. Environmental factors such as signal blockage by proximity to buildings and atmospheric conditions can increase this error further. Given many of the evaluation visits were conducted during Malawi's wet season, atmospheric conditions are likely to have unavoidably affected evaluation visit GPS readings.

Study design and standard operating procedures endeavoured to minimise these errors. Firstly, to maximise accuracy of the GPS measurements, waypoint averaging was employed. This functionality takes a measurement every second until the sample size is large enough that the estimate of the mean remains stable with additional sampling (defined on the handset as "100% confidence"). This process takes 3-5 minutes depending on the weather conditions and satellite position, but greatly improves accuracy compared to single waypoint measurement. Secondly, to avoid signal blocking by buildings, standard operating procedures stipulated measurement of GPS from 1m in front of the main dwelling entrance.

In the clinic, ePAL users selected the centre of the participant's dwelling as doors were not visible on the maps. Therefore, the GPS data collection method ensured adequate signal strength, but contributed to a consistent offset of several metres from the location of the dwelling if correctly identified on ePAL.

The evaluation sub-study was powered to measure a proportion of 85% of ePAL measurements, with a lower limit of the 95% confidence interval of 80%, within 20m

of the GPS reading. The distance was based upon the assumptions of 10m inaccuracy in the GPS reading plus 5m inaccuracy due to the selection of the centre versus the door of the building, allowing for only 5m inaccuracy in the selection of the location in ePAL. Considering the limitations in both GPS measurement and use of ePAL experienced once in the field, this was soon identified as an unrealistic expectation. Given the environmental conditions, 10m was likely an overestimate of the accuracy of the GPS reader. The sources of inaccuracy from the practical use of ePAL were also underestimated, such as using fingers to select the place of residence could introduce inaccuracy. Therefore, the median estimate of 84m between the evaluation and ePAL coordinates was considered a success, especially when considered in context (see Figure 5.4). To note, ePAL is now being used with stylus screen pens to allow more accurate selection of the map location.

5.3.4 Additional information on statistical methods

For sustainability of ongoing use by the local data management team, data cleaning and data analysis for monthly monitoring was conducted in Stata, version 13.1 (StataCorp LP, College Station, Texas).

The map tiles in the ePAL app were converted to the Google Earth (ESPG: 3857) projection. When a location is selected on the screen, the application programming interface accounts for location and zoom level to record the coordinates in WGS84 (EPSG: 4326) projection in the ePAL database in comma separated value format. The GPS handsets collected data in the WGS84 setting, therefore direct calculation of the distance between the spatial points collected by ePAL in the clinic and by the Garmin handset at the evaluation visit was possible.

Several options exist for calculating the distance between two spatial points in R. To ensure that choice of package/command did not significantly affect the estimated distance, the distance between the GPS and the ePAL data set were calculated using the ‘raster’ package ‘pointDistance’ command, ‘geosphere’ package ‘distm’ command, and the ‘sp’ package ‘spDists’ command. The difference between these

methods of distance estimation did not exceed 0.2% of the estimated distance, therefore all were considered suitable methods. The spDists calculates the distance between the two points using Great Circle distance (WGS84 ellipsoid) and provides convenient to use outputs, so was selected for use in this study.

5.4 Additional monitoring and evaluation results

5.4.1 Monitoring

Given the primary aim of the study was capture of patient place of residence, one of the key monitoring aspects was appropriate capture of coordinates. The overall coordinate capture of those reporting urban Blantyre residence was 89%. Monthly point estimates of the proportion with available coordinates ranged 77% to 95% (Table 5.4). Through monitoring and feedback to the TB officers at monthly meetings, coordinate capture was maintained at high levels.

Limitations of this measurement should be recognised, as the 'gold standard' measure relies on participant self-report of residency in urban Blantyre. Errors in self report could lead to over- or under-estimation of coordinate capture.

Comparison of case capture in ePAL to the NTP register was an important monitoring measure during the study. NTP records provided a robust denominator. Monthly case capture for the study and continued use of ePAL are presented in Figure 5.5.

Table 5.4: Proportion coordinate capture by month in participants reporting residency in urban Blantyre

Study Month	Proportion of urban Blantyre residents with coordinates (95% confidence intervals)
1	81% (72-88%)
2	87% (79-93%)
3	77% (68-85%)
4	93% (86-97%)
5	92% (85-96%)
6	94% (87-98%)
7	95% (89-98%)
8	93% (86-97%)
9	91% (84-96%)
10	84% (75-91%)
11	86% (78-92%)
12	91% (84-96%)

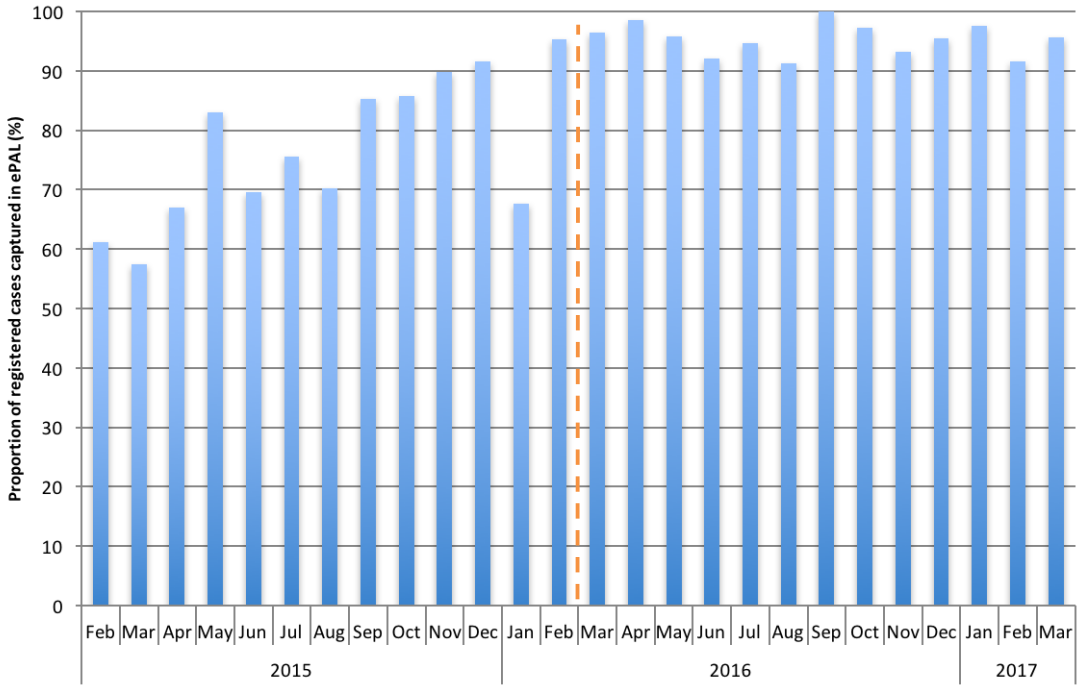


Figure 5.5: Monthly case capture with ePAL as compared to NTP registers from February 2015 to March 2017.

5.4.2 Evaluation

The main evaluation outcome was defined as distance between the evaluation and ePAL coordinates estimated in metres. Median distance between the ePAL and evaluation coordinates was 84m (IQR: 35m-317m). However, as a quality assurance measure, the latitude offset and longitude offset between ePAL and evaluation visit coordinates were also calculated to ensure there was no systematic offset (Figure 5.6). There is no visible skew in the data, so no issue was identified.

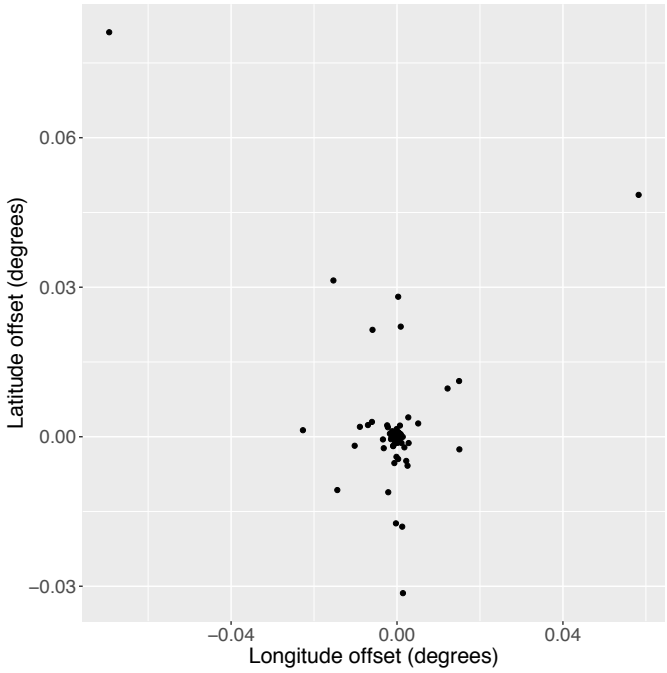


Figure 5.6: Longitude and latitude offset (WGS84 degrees) between pairs of ePAL and evaluation coordinates

Exact time between clinic registration and evaluation follow up visit was not recorded. However, date of upload of evaluation visit data was available, so lag between date of collection and latest possible follow up date was estimated. As discussed in section 5.2, mean lag between ePAL entry and latest possible follow up was 26 days (IQR: 14-37 days), ranging from 2 days to 182 days (Figure 5.7). Spearman's correlation between the evaluation distance and data collection lag estimate was 0.12 (p=0.17), suggesting negligible correlation.

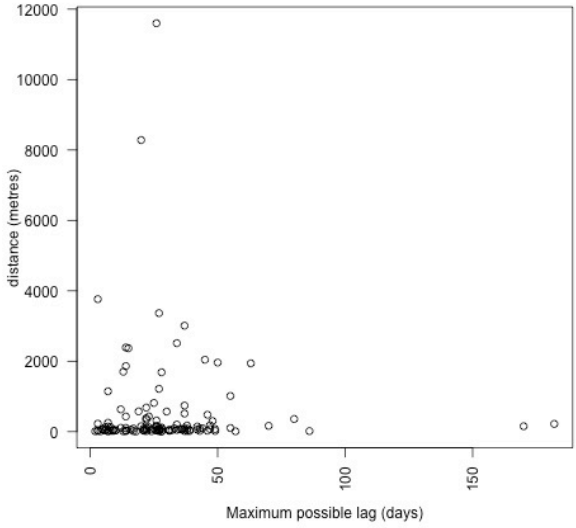


Figure 5.7: Comparison of time elapsed between ePAL registration and evaluation data upload ('maximum possible lag', days) plotted against evaluation distance (metres).

In total, 204 evaluation and monitoring visits were conducted during the 12 months of the main study. Evaluation/monitoring results by month are presented in Figure 5.8.

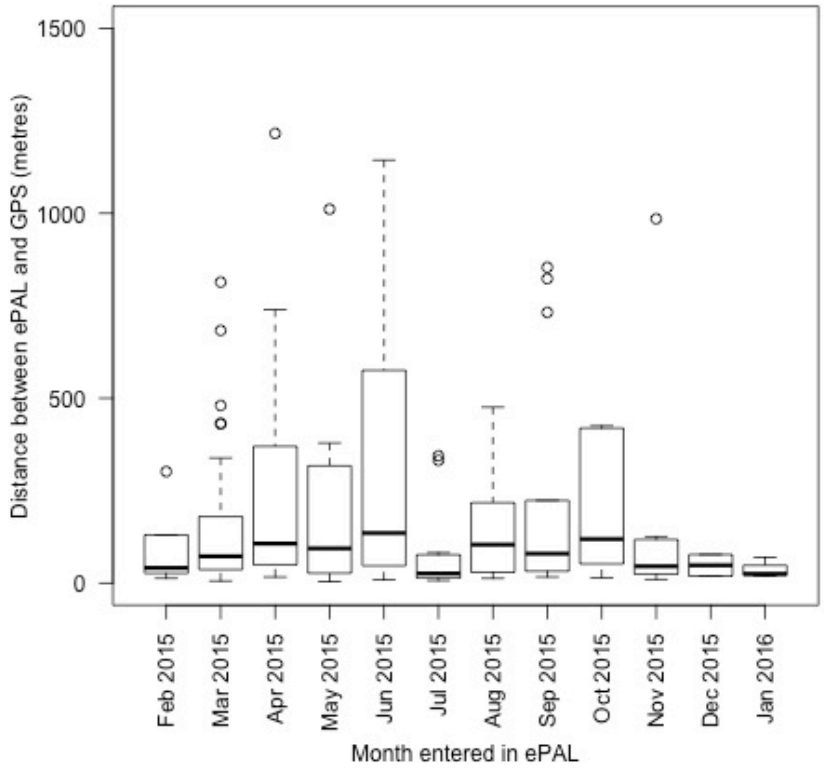


Figure 5.8: Estimated distance between ePAL- and GPS- measured place of residence by study month. To allow visualisation of the data trends, the y axis was cropped at 1500m. A small number of points lay above this cut off: Feb 2015 (n=3), Mar (n=2), Apr (n=3), May (n=3), June (n=3), Aug (n=2), Sept (n=2), Oct (n=2), Nov (n=1).

All patients reporting residence in the evaluation area during the first 4 months of the study were invited to participate in the evaluation visit at the home. Consent rates for evaluation visits were very high, and if 100% follow up were achievable, closer to the 196 needed for the evaluation cohort would have been achieved. However, as was expected, the challenges of lack of addresses and incomplete mobile phone ownership for contacting participants led to loss to follow up. Due to these challenges, the time per evaluation visit was much greater than expected.

Characteristics of participants with missing data have been described elsewhere section 5.2. Characteristics explored were suggestive of a difference in missing data

by TB officer, though this appeared diminish with only evidence of a difference over the 12 month period. ePAL data were not used to assist evaluation visits, so did not bias the results. Theoretically, an association could exist between ability to describe place of residence and both ease of evaluation follow up and of locating place of residence in ePAL. Therefore, those located for evaluation visits could be those with more accurate ePAL locations. Data were not available to test this possible risk of bias, but given evaluation visits could be arranged by meeting at a well-known point of interest, this bias is considered fairly unlikely.

The 12 months of the study was, by design, initiated at the same time as the launch of the searchable POI function as this was the last major change to the app. However, it took some time for the TB officers to uptake and become comfortable using the POI search function. Therefore, in the early phase of the study, some TB officers were probably not using the full functionality of the tool. The POI search function greatly increased the ease of finding the correct area on the map. Therefore, evaluation should have perhaps been initiated a month or two after the implementation of the app update. However, this most likely means estimates of accuracy are conservative, as the POI search function may have improved the user experience.

5.4.3 Study continuation

In moving forward with ePAL after the 12-month study period, evaluation transitioned to monitoring visits. These followed the same format as evaluation visits, but were conducted for 5% of registrations (at least 1 per centre per month). The written consent and evaluation questions were removed from the app as the evaluation period had finished. A small number of key risk factor questions were moved into the case report form.

Ethics approval has been sought and granted for de-escalating the enrolment age for ePAL data collection as part of the "Hit TB Hard" study to include children. For children aged <8 years old, verbal consent was sought from the guardian for the child's participation. For children aged 8-17 years, verbal consent was sought from

the guardian and verbal assent procedures were put in place for the participant. The child was only enrolled if both patient assent and guardian consent were given.

5.5 Spatial mapping with ePAL to inform clinical trial design

5.5.1 Utility of ePAL to inform design of prevention of disease TB vaccine clinical trials

ePAL provides a low-cost, rapidly implementable tool that can be employed to assess the spatial distribution and burden of disease in specific populations. There is a shortage of research sites with both clinical trial capacity and appropriate available epidemiological data. As described previously, spatiotemporal data can help inform identification of TB disease hotspots, both overall in the population and in potential subsets for trial recruitment. This can inform data-driven sample size calculations to help balance the minimisation of costs against risk of achieving insufficient endpoints. From a safety perspective, minimising the number exposed to investigational products is preferable, and from an ethical perspective, if new vaccines are effective, enrolling those populations most at risk improves the risk-benefit of participation.

Without an alternative source of epidemiological data, NTP data would be employed to inform sample size calculations. However, publicly available data are aggregated at the regional or district level,[18] and do not differentiate whether patients registering in Blantyre are city residents. Results of the ePAL study demonstrate that a large proportion (27%, n= 514) of those registering in Blantyre are rural Blantyre or out-of-district residents. A TB vaccine trial with a sample size based upon NTP data would be substantially under-powered, as these non-residents would be included. Therefore, even the most basic use of ePAL to identify Blantyre residents would be an important contribution to informing appropriate sample size calculations for new TB vaccine efficacy studies. A simpler questionnaire-based approach could be taken to collecting this level of data, but there are many additional advantages to the high resolution data ePAL provides.

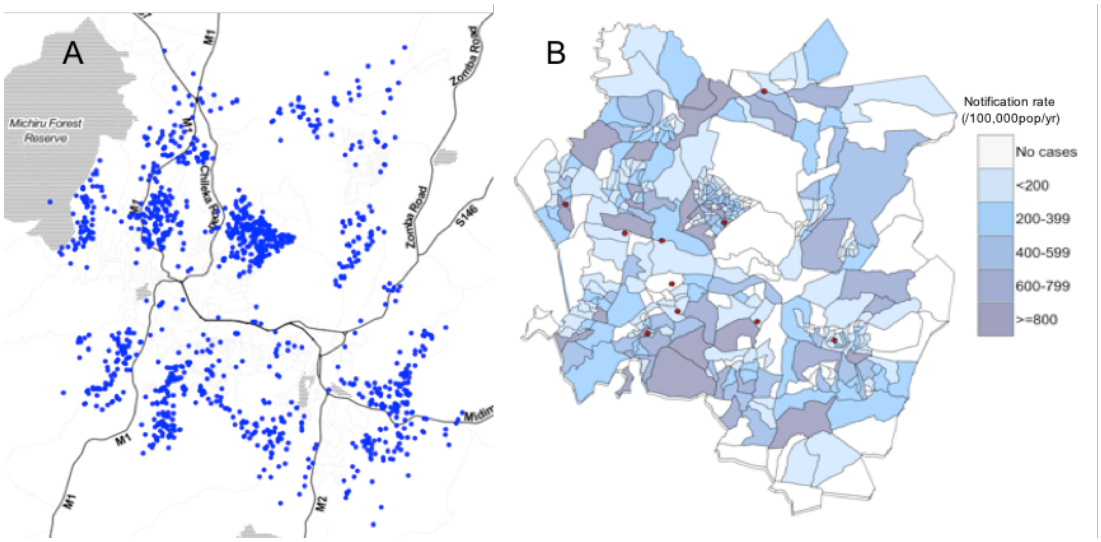


Figure 5.9: All TB registrations collected using ePAL over a 12 month period (February 2015-2016) provided as A) individual spatial points (blue dots), and B) notification rates estimated by census enumeration area (red dots are health centres)

As can be seen in Figure 5.9A above, the majority of incident TB cases in Blantyre were focussed in central and central-west areas of the city. This high concentration of cases in these areas is indicative of potential suitability for trial recruitment. In Figure 5.9B, these ePAL registered cases are combined with census enumeration area population estimates to calculate TB notification rates.[7] This provides much higher granularity than the NTP-reported data, where rates are reported for urban Blantyre. Once converted to rates, even though the areas with a large number of cases were also areas of high population density, the number of cases in this area was so high that rates remained high, indicating that this could be a suitable population for TB vaccine trial recruitment. However, several additional areas are highlighted once rates as opposed to numbers of cases are considered. Several areas that did not appear high burden when considering the point processes data were identified as having high notification rates. Given clinical trial sample sizes are based upon rates, these areas with fewer residents but high numbers of reported cases could be ideal for TB vaccine clinical trials.

For ongoing clinical trials, the exact definition of 'disease' varies between studies, with some measuring effect against purely microbiologically confirmed endpoints,

and others using broader clinically-defined endpoints.[2,19] The results in Figure 5.9 would most likely be directly relevant for studies measuring clinically defined endpoints. Some studies limit the outcome to pulmonary disease, therefore mapping exclusively of pulmonary disease may be more informative. Sixty percent of ePAL participants (1,135/1,899) reported pulmonary disease, which is mapped in Figure 5.10 below. 785 participants resident in urban Blantyre with available coordinates reported pulmonary disease.

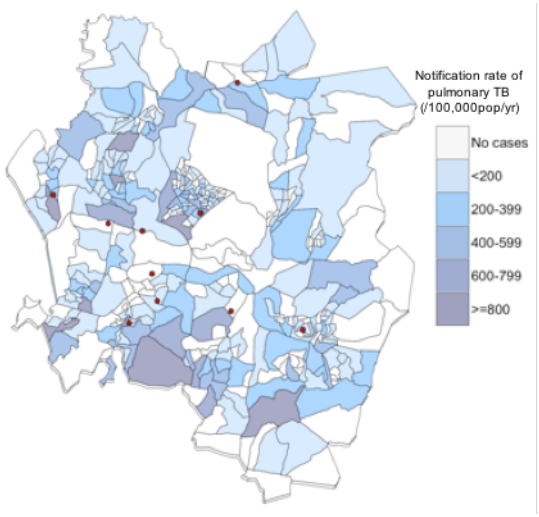


Figure 5.10: Estimated notification rates by census enumeration area of pulmonary TB disease using ePAL data

This would not be an accurate estimate of the disease rates and, potentially, the location of patients in studies where outcomes are limited to microbiologically confirmed disease. In this study only 43% of patients had microbiologically confirmed disease at time of registration. In the subset of patients resident in urban Blantyre with coordinates available, 48% were microbiologically positive at time of registration. Therefore, exploring the distribution and rates of microbiologically positive disease could help inform trial design and recruitment tailored more precisely to the study endpoint (Figure 5.11).

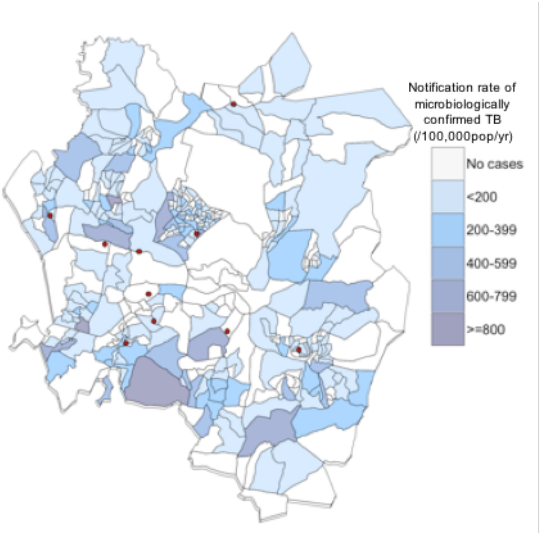


Figure 5.11: Estimated notification rates by census enumeration area of microbiologically confirmed TB at time of registration using ePAL data

One limitation of this approach is the delay in culture results becoming available. Planned future analyses include merging the ePAL database with culture results from the parent study (Hit TB Hard), to include patients with culture results becoming available after the treatment initiation visit.

Mathematical modelling in chapters 3 and 4 also demonstrated that targeting vaccination to groups most at risk of developing disease could help maximise impact of new TB vaccines. In particular, age targeting of vaccination may be of importance. Clinical trials currently ongoing have recruited adult populations of ages 15-65 and 18-50 years. The maps below (Figure 5.12 maps B and C) are a demonstration of the distribution of disease in these populations. Other studies have recruited narrower age ranges (e.g. 13-15 years). Map A in Figure 5.12 demonstrates the anticipated rates in 18-25 years olds as an example of the distribution that could be expected in more limited young adult age groups. The age for inclusion in ePAL has now been widened to include birth upwards, with guardian consent and participant assent processes in place. Therefore, data relevant for these younger age groups will shortly be available, and could be assessed in future analyses.

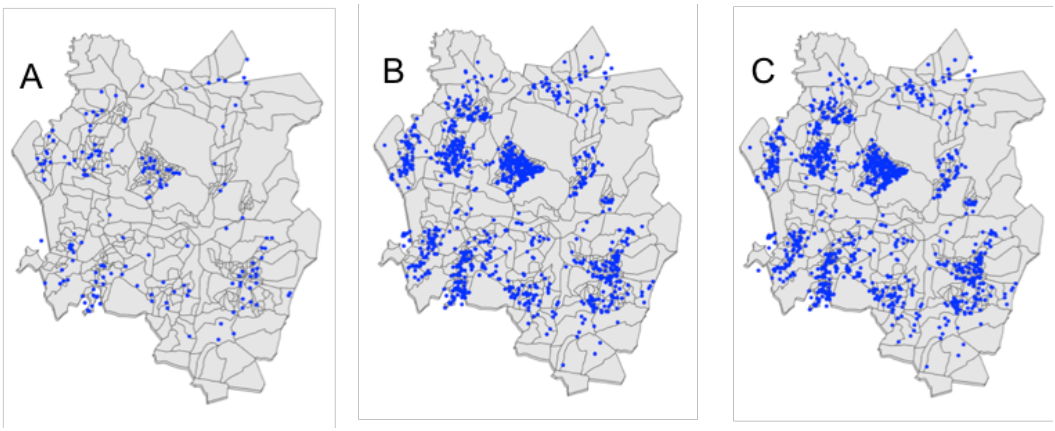


Figure 5.12: Distribution of cases in A) 18-25 year olds, B) 18-50 year olds, and C) 18-65 year olds

Finally, studies may want to include or exclude certain risk groups, based upon safety or likelihood of mounting an effective immune response to a vaccine. For example, immune compromise caused by HIV infection increases the risk of development of TB disease, but also may impact an individual's ability to develop a sufficient immune response to vaccination. Therefore, inclusion of such populations in efficacy trials could dilute measures of efficacy, risking a no-go decision on a vaccine that could otherwise be effective in a healthy population. These risk groups are often studied in separate safety and immunogenicity studies. Therefore, spatial maps are presented in Figure 5.13 limiting the data set to HIV negative patients.

Figure 5.13: Spatial distribution of TB cases in Blantyre by HIV status A) HIV-positive (yellow) and HIV-negative (red) cases, B) HIV-negative cases, and C) estimated notification rates of TB in HIV-negative populations, assuming homogenous distribution of 18.2% population-level HIV prevalence.

Trial endpoints may affect the anticipated rates of disease. As a demonstration of the impact of endpoint selection, Figure 5.14A below shows the enumeration area TB rates for all TB, and Figure 5.14B demonstrates the TB rates and changes in hotspots when the primary outcome is limited to microbiologically-positive pulmonary disease. In these first two figures all participants are included, whereas in Figure 5.14C it was assumed that recruitment criteria would exclude HIV positive populations. Enumeration-area level HIV prevalence was not available to estimate these rates, so the 2015 HIV prevalence estimate for Blantyre city (18.2%) was applied to the population estimate of each enumeration area.[20]

Figure 5.14: Estimated enumeration area notification rates for A) all TB cases, B) microbiologically-confirmed pulmonary TB cases, and C) microbiologically-confirmed pulmonary TB cases in HIV-negative populations. HIV negative rates were estimated by adjusting the population level denominators based upon 18.2% HIV prevalence in Blantyre city[20] *highlight examples of areas particularly affected compared to the map immediately to the left. Red circles are TB clinics or hospitals

Limiting the primary outcome definition and the HIV inclusion criterion substantially reduced notification rates. Therefore, without this adjustment to rate estimates, the TB vaccine trial sample size would be too small and the study would most likely be underpowered. Importantly, several enumeration areas marked with a * were particularly affected by the change in outcome or recruitment criteria. These areas may be suitable for a trial with a broad outcome measure and recruitment population, but become unsuitable if the trial design is more restrictive.

These analyses serve as an example of mapping to inform trial design with the ePAL data set. Once future clinical trials have defined the recruitment population of interest, data generated using ePAL can help direct study recruitment to areas where disease burden is highest in those specific populations, and can help provide a more accurate estimate of notification rates to inform sample size calculations.

5.5.2 Possible utility of ePAL for prevention of relapse studies

In recent years, the costs and challenges of developing new TB vaccines has led to innovation in approaches to proof of concept studies. One such innovation is conducting prevention of relapse (POR) studies for phase IIB proof of concept of efficacy before progressing to phase III POD studies. POR is effectively a subset of POD, but in a population at elevated risk of developing disease compared to the general population, therefore minimising the study sample size.

In the ePAL database, relapse patients comprised 6% (n=120) of TB patient records over a 12-month period, of which 70 were from urban Blantyre and had coordinates available. ePAL data can provide an approximate relapse proportion based upon the ratio of new to relapse cases (1 relapse to every 14 new patients), which can inform sample size by giving an approximation of the anticipated relapse rates that could be expected from recruitment of patients with history of disease. These data would also be available from the NTP registers, so most importantly, the spatial data available from ePAL would also help study staff to locate people with history of TB disease for targeted recruitment (Figure 5.15). Ethical considerations may mean that it is not appropriate to use the data to seek and recruit patients at their home, but patients could be recruited at the end of treatment at TB clinics, and ePAL used to identify their place of residence to facilitate clinical trial follow up visits and reducing loss to follow up.

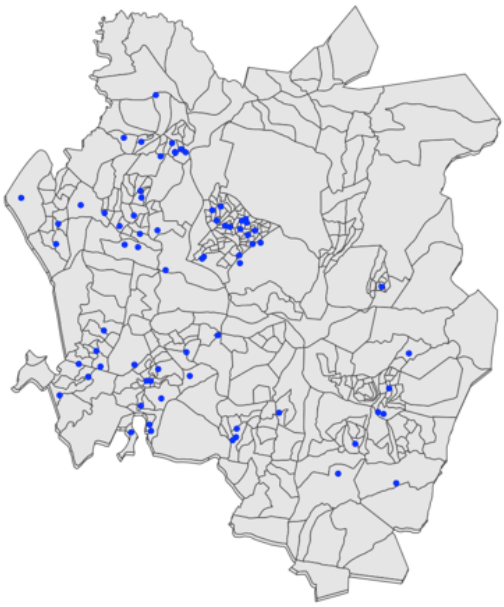


Figure 5.15: Spatial location of relapse cases recorded in ePAL with coordinates available between 12th February 2015 to 11th February 2016

5.5.3 Possible utility of ePAL for pre- versus post-infection studies

Mathematical modelling presented in Chapters 3 and 4 demonstrated that, in China, vaccines effective post infection will be important to maximise vaccine impact. Therefore, recruitment of latently infected populations in TB vaccine trials for demonstration of safety, and ideally efficacy, will be essential. Testing for latency is generally either an inclusion criterion or part of initial participant assessments to ensure infection status is known at time of vaccination. The M72 and *M.vaccae* TB vaccine clinical trials recruited exclusively IGRA-positive or TST-positive participants.[2,19]

If recruiting for a post-infection (latency) study, recruiting from populations with high prevalence of infection can reduce the number of prospective participants that are excluded at the screening stage. However, the ePAL study is *not* designed to identify latently infected (or uninfected) populations. The burden of disease in the population identified using ePAL is likely a reasonable proxy for the prevalence of infection, on the condition that the underlying cause of high disease rates includes high infection

risk, not just rapid progression to disease. Therefore, data collected using ePAL to inform sample size calculations and recruitment populations could also be indicative of the populations with high prevalence of infection and therefore may reduce the number of exclusions at the screening stage.

5.5.4 Strengths and Limitations of ePAL

The main strengths and limitations of ePAL as a data collection tool have been discussed in detail in section 5.2.

The main strength of this tool for informing TB vaccine clinical trials is the ability to rapidly and inexpensively generate up-to-date information on the spatial distribution of burden of TB disease in high burden urban settings without municipal address systems. Additional questions with regards to demographic or risk factors can be included in the eCRF to allow stratification of the data by potential recruitment characteristics.

One of the main limitations of ePAL for the estimation of notification rates is the availability, quality and relevance of the denominator data. In Blantyre, census data were readily available to allow estimation of notification rates by census enumeration area (population size: 134-4,360).[7] However, the census was last conducted in 2008, so is relatively outdated. Population growth estimates applied to the data provided population sizes reflective of the 2015 population to estimate more accurate rates of disease. It could be expected that the population may grow faster in high density areas of the city, therefore rate estimates for high density settlements using these denominators may be over-estimated. An enumeration study has recently been conducted in Blantyre to improve these estimates, so planned future analyses will be conducted with these updated enumeration figures. Another limitation was the lack of age and risk-factor stratified denominator data, limiting the availability of denominators for analyses of these sub groups for clinical trials.

Availability of denominator data could potentially be a limitation for use of ePAL for trial design in other settings. However, most settings have census data available, or alternatively progress in the use of satellite imagery for population enumeration could allow the high-resolution satellite maps to be used for enumeration.[21]

As mentioned in section 5.5.3, ePAL would have limited utility in informing recruitment specifically of uninfected (pre-infection) or infected (post-infection) populations, or informing design of prevention of infection studies. There may be a role for ePAL in design of prevention of relapse studies.

As ePAL is embedded in the NTP, this tool does not avoid the fundamental limitation of missed burden through passive case finding. From the perspective of designing TB trials, this would lead to conservative sample size estimates, which unfortunately does not minimise resource use, but ensures the study is not under-powered.

5.6 ePAL adaptations for other research studies

ePAL is now in routine use at the Blantyre research site as an integral part of the Hit TB Hard study, and several other research studies.

The map aspect of ePAL has been adapted for compatibility with Open Data Kit (ODK) for use in a pneumococcal burden of disease study led by Todd Swarthout. With this adaptation, the ePAL annotated maps can be called in from an ODK eCRF. Similarly to the core ePAL app, the coordinates of the location selected on the map are recorded in the study database. This adaptation has the advantage of CRF adaptability without the need for support from app developers. However, the ODK CRF interface is not as user-friendly as the ePAL CRF, thus it is suitable for use by experienced research staff, but may not be appropriate for integration into the public health system.

In a typhoid study led by Franziska Olgemoeller and Nick Feasey, ePAL has been adapted to collect multiple locations. This allows collection of both the place of

residence and key water sources, used for following up patients at the home, and also exploring possible sources of transmission.

I am currently exploring the potential use of ePAL in other TB- or vaccine-related projects in South Africa and India.

5.7 Conclusions and future work

The ePAL study successfully delivered the design, implementation and evaluation of a low-cost novel tool for clinic-based high-resolution mapping of the place of residence of TB patients in Blantyre, Malawi. These data can be used to inform design of clinical trials for new tools such as new TB vaccines. The data generated allows for a more accurate estimation of the burden of disease relevant to the intended trial outcomes and in the populations of interest for recruitment in to TB vaccine clinical trials. If participants are recruited at a location other than their place of residence, ePAL would provide an accurate tool for identifying place of residence for subsequent follow up visits.

The data presented here establishes the spatial accuracy of ePAL, and demonstrates examples of the utility of such data to inform spatially targeted recruitment in to TB vaccine clinical trials and appropriate sample size calculations.

A wealth of data has been generated by ePAL in Blantyre. Therefore, in future work I plan to explore the dataset further to improve our understanding of the spatial distribution of TB in Blantyre and associated risk factors, to help inform current public health measures, and in the hope of informing future clinical trials. These analyses will include exploring risk factor data, employing the new enumeration data set as a denominator, and conducting observed versus expected analyses to identify statistically significant TB hotspots. In addition, several TB vaccines in the pipeline target paediatric populations; therefore, given paediatric ePAL data are now

available, it could be of value to explore the rates of disease in children in Blantyre to explore suitability for paediatric TB vaccine trials.

It is also hoped that ePAL will continue to be employed in other studies in Blantyre, and I am currently exploring the possibility of implementing ePAL in studies in South Africa and India.

5.8 Chapter 5 References

1. McShane H. Need for more TB vaccine field sites. *Indian journal of experimental biology* 2009; **47**(6): 445-6.
2. GlaxoSmithKline. Study to Evaluate the Efficacy of GlaxoSmithKline (GSK) Biologicals' Candidate Tuberculosis (TB) Vaccine in Adults (NCT01755598). May 2017 2012. https://clinicaltrials.gov/show/NCT01755598 (accessed 3rd May 2017).
3. Farvaque-Vitkovic C GL, Leroux H, Verdet F, and Chavez R. Street Addressing and the Management of Cities. 2005. http://siteresources.worldbank.org/CMUDLP/Resources/461753-1160058503655/Street_Addressing_Manual.pdf?resourceurlname=Street_Addressing_Manual.pdf (accessed 2nd March 2017).
4. MacPherson P, Choko AT, Webb EL, et al. Development and validation of a global positioning system-based "map book" system for categorizing cluster residency status of community members living in high-density urban slums in Blantyre, Malawi. *American journal of epidemiology* 2013; **177**(10): 1143-7.
5. Corbett E. Intensified HIV/TB prevention linking home- based HIV testing, including the option of self-testing, with HIV care. 2012. http://www.controlled-trials.com/ISRCTN02004005/ (accessed 13th August 2017).
6. United Nations Department of Economic and Social Affairs Population Division. World Urbanization Prospects: The 2014 Revision, (ST/ESA/SER.A/366). 2015. https://esa.un.org/unpd/wup/CD-ROM/ (accessed 20th August 2017).
7. National Statistical Office of Malawi. Malawi Population and Housing Report 2008. 2008. http://www.nsomalawi.mw/images/stories/data_on_line/demography/census_2008/Main Report/Census Main Report.pdf (accessed 18th September 2016).
8. Hargreaves JR, Boccia D, Evans CA, Adato M, Petticrew M, Porter JDH. The Social Determinants of Tuberculosis: From Evidence to Action. *American journal of public health* 2011; **101**(4): 654-62.
9. World Health Organization. Global Tuberculosis Report 2016. 2016. http://apps.who.int/iris/bitstream/10665/250441/1/9789241565394-eng.pdf?ua=1 (accessed 10th May 2017).
10. Seebregts C, Hosseini M, Bleed D, Williams BG, Dye C. Developing electronic tuberculosis recording and reporting systems: summary report of a WHO expert consultation meeting. 2006. http://www.who.int/tb/err/err_expertmeeting_report_jul06.pdf (accessed 10th September 2017).
11. Subbaraman R, Thomas BE, Sellappan S, et al. Tuberculosis patients in an Indian mega-city: Where do they live and where are they diagnosed? *PloS one* 2017; **12**(8): e0183240.
12. Stop TB Partnership. Key Populations Brief: Urban Populations. http://www.stoptb.org/assets/documents/resources/publications/acsm/KP_Urban_Spreads.pdf (accessed 2nd September 2017).
13. Fry S, Cousins B, Olivola K. Health of children living in urban slums in Asia and the near east: Review of existing literature and data. 2002.

http://www.ehproject.org/PDF/Activity_Reports/AR109ANEUrbHlthweb.pdf (accessed 2nd September 2017).
14. European Space Imaging. European Space Imaging website. 2014. http://www.euspaceimaging.com/ (accessed 14th October 2014).
15. R Core Team. R: A language and environment for statistical computing. . R Foundation for Statistical Computing, Vienna, Austria; 2014.
16. National Coordination Office for Space-Based Positioning Navigation and Timing. GPS accuracy. 2017. http://www.gps.gov/systems/gps/performance/accuracy/ (accessed 14th July 2017).
17. Garmin. What is GPS? http://www8.garmin.com/aboutGPS/ (accessed 23rd June 2014).
18. Nyirenda T. Epidemiology of Tuberculosis in Malawi. The Epidemiology of Malawi. 2nd ed. Geubbels E, Bowie C 2009.
19. Anhui Zhifei Longcom Biologic Pharmacy Co. Phase III Clinical Study of Efficacy and Safety of Vaccae™ to Prevent Tuberculosis. 27th December 2016. https://clinicaltrials.gov/show/NCT01979900 (accessed 3rd January 2017).
20. Government of Malawi Ministry of Health. Malawi Population-based HIV Impact assessment: summary sheet. 2016. https://www.hiv.health.gov.mw/images/Documents/MALAWIFactsheet.pdf (accessed 21st September 2017).
21. Checchi F, Stewart BT, Palmer JJ, Grundy C. Validity and feasibility of a satellite imagery-based method for rapid estimation of displaced populations. *International Journal of Health Geographics* 2013; **12**: 4-.

CHAPTER 6: Discussion

Summary of Chapter 6

In this final thesis chapter, I provide a discussion of the research undertaken in this PhD and the implications of the results for the TB vaccine development community. The chapter summarises the research objectives and how they were met, the key findings of the research, the strengths and limitations, and the contribution of this thesis to advancing knowledge in the TB vaccine field. I place these findings into the context of existing research and the current TB vaccine pipeline. I provide recommendations for development of TB vaccines from the perspective of vaccine developers, clinical triallists, and country-level decision makers. Opportunities for future research are identified and discussed.

CHAPTER 6 Discussion

6.1 Summary of the PhD Research

New TB vaccines are urgently needed to help meet the WHO goal of tuberculosis (TB) elimination by 2050. Epidemiological evidence to inform TB vaccine development strategies and to assist clinical trial site selection and design was identified as a critical research need to accelerate vaccine development. Existing modelling research on this subject has failed to fully address the data needs for developing TB vaccine target product profiles (TPPs). Available data to inform clinical trial design often lacks granularity, and prospective data collection tends to be expensive and time consuming. Through the research presented in this thesis I aimed to inform the development of appropriate TB vaccines and target populations to maximise future population-level impact, and help selection of trial sites and recruitment populations to accelerate vaccine development. This was achieved, firstly, through the development of mathematical models to estimate the population-level impact of vaccines with varied characteristics and implementation strategies to inform design of TB vaccine target product profiles. Secondly, the development of a novel, low cost rapidly implementable tool for spatial mapping of TB cases in high TB burden urban settings to accelerate and de-risk clinical trials.

The research in this thesis sought to fulfil four research objectives. These objectives, and how they were fulfilled, are summarised below.

6.1.1 Objective 1

The first objective sought to a) summarise existing literature exploring the effect of vaccine characteristics and implementation on the potential epidemiological impact of new TB vaccines, and b) identify research needs with respect to mathematical modelling of possible vaccine characteristics and implementation strategies for new TB vaccines.

This was achieved by conducting a systematic review of the TB vaccine mathematical modelling literature exploring the epidemiological impact of new TB vaccines. A narrative summary of the literature was produced and research needs identified. The most urgent research needs for TB vaccine development were selected for exploring in the subsequent modelling research. These included a comprehensive exploration of the impact of vaccine characteristics, modelling the impact of new TB vaccines in China, and exploring the impact of vaccinating older adults with new TB vaccines.

6.1.2 Objective 2

The second objective of this research aimed to develop a mathematical model and calibrate to epidemiological and demographic temporal and age distribution trends in China, and simulate the introduction of new TB vaccines to investigate the population-level epidemiological impact of varying TB vaccine characteristics. As part of this work, I sought to identify characteristics that could help maximise population level impact in this setting to inform TPP 'ideal' vaccine characteristics. Additionally, to identify the combination of vaccine characteristics that would be most likely to deliver pre-specified minimum incidence rate reductions compared to the no new vaccine baseline in 2050. These results would be used to help inform minimum TPP characteristics.

This objective was achieved through the development of an age-structured and age-calibrated transmission model of TB epidemiology in China, followed by the incorporation of new TB vaccines to explore the epidemiological impact of different combinations of efficacy against infection and disease, pre- versus post-infection vaccines, and varied durations of protection and frequencies of mass vaccination campaigns. Population-level impact of these vaccines implemented over 2025-50 was estimated, and proposals for TPP vaccine characteristics to achieve a minimum population level impact were made.

6.1.3 Objective 3

Objective 1 identified a research need to explore the epidemiological impact of vaccinating older adults. Therefore, objective 3 aimed to investigate the impact of age-targeted programmes vaccinating adolescents versus older adults on population-level TB epidemiology in China.

This was achieved by adapting the age-structured model developed in objective 2 to explore age-targeted vaccination. A range of vaccine types were modelled to explore variation in the relative impact of new TB vaccines by age of vaccination.

6.1.4 Objective 4

To support the translation of TB vaccine TPPs into clinical trials, objective 4 sought to develop a low cost, rapidly implementable and easy to use tool for spatial mapping of TB in high burden urban settings without municipal address systems. Such a tool could be used to identify areas of high burden in populations potentially relevant for TB vaccine trial recruitment.

This was achieved through the development, implementation and evaluation of a novel app (ePAL) integrated in to the NTP in Blantyre, Malawi, for clinic-based identification of place of residence of patients registering for TB care. Using data collected with ePAL, spatial maps with examples of outcomes and participant characteristics relevant to TB vaccine trials were generated.

6.2 Summary of findings

6.2.1 Mathematical modelling to inform TPP design

The main findings with regards to informing TB vaccine TPPs are summarised by vaccine characteristic below.

Prevention of infection versus disease:

- In China, efficacy for prevention of disease was key to maximising impact on disease burden with new TB vaccines over the 2025-50 time horizon.
- Incidence rate reduction in 2050 was relatively insensitive to changes in efficacy for prevention of infection.

Pre- versus post-infection vaccines:

- With adolescent/adult vaccination (i.e. routine vaccination of 9 year olds plus 10-yearly mass campaigns of ≥10 year olds) and with older adult vaccination (i.e. routine vaccination of 60 year olds and a one-off mass campaign of 61-64 year olds), greatest epidemiological impact was achieved with vaccines effective post-infection or pre- and post-infection (P&PI).
- When effective P&PI or post-infection, overall efficacy was mostly dependent on efficacy against disease. For pre-infection vaccines, the epidemiological impact was dependent on both efficacy against infection and efficacy against disease.
- The relative impact of pre- versus post-infection vaccines varied by age targeted vaccination. Targeted vaccination of adolescents (15 year olds) achieved greater impact with pre-infection vaccines, whereas older adult vaccination was more effective with post-infection vaccines. However, when comparing identical vaccines, older adult vaccination provided consistently greater impact, so to maximise population-level impact, a post-infection vaccine delivered to older adults is required.

Duration of protection:

- With routine vaccination of 9 year olds and 10-yearly mass campaigns in China, a vaccine with at least 5 years protection would be required to achieve at least 20-29% incidence rate reduction in 2050 compared to the no new vaccine scenario. If at least 50-59% incidence rate reduction were required, a duration of at least 7 years would be needed with 10-yearly campaigns.

- Shorter durations of protection could be compensated for by increasing the frequency of mass campaigns (e.g. 5-yearly).
- Even high efficacy vaccines would have limited impact without sufficient duration of protection or frequent mass vaccination campaigns.

Vaccine efficacy:

- In China, to achieve 20-29% incidence rate reduction in 2050 with 5 years duration of protection and 10-yearly mass campaigns, the required median vaccine efficacies were: prevention of disease (VE-POD) 40% (range: 0-60%) *and* prevention of infection (VE-POI) 60% (range: 0-100%). To achieve 50-59% incidence rate reduction in 2050, the VE-POI would be unchanged, but the VE-POD would need to increase to 100% (100-100%). Ranges are wide because the POD and POI efficacies to provide a given level of impact are correlated, e.g. for 20-29% incidence rate reduction, if VE-POD=0%, a VE-POI of 90-100% would be required; or if VE-POI=0%, VE-POD of 40-60% would be required.
- When results of vaccine efficacy studies become available, heatmaps in Chapter 3 can provide an estimate of the potential impact of that vaccine.

Age-targeted vaccination:

- With equivalent vaccines, older adult vaccination delivered greater impact than adolescent vaccination, even if much lower coverage were achieved in older adult vaccination.
- Post-infection vaccines provided substantial impact when delivered to older adults, whereas impact was negligible when this vaccine was delivered to adolescents. Although pre-infection vaccines provided relatively similar levels of impact when vaccinating either age group, the impact was substantially lower than that achievable with vaccination of older adults with a post-infection vaccine.
- Recommendations for post-infection vaccines were robust to substantial reductions in efficacy and duration of protection in older adults, whereas for pre-

infection vaccines the preference for older adult vaccination could be diminished.

- In this setting, a vaccine efficacious post-infection or pre- and post-infection delivered to older adults will be critical to maximise population-level impact.

Maximum achievable impact:

- With the adolescent/adult delivery strategy, the maximum possible incidence rate reduction in 2050 was 79% (UR: 77%-81%) reducing incidence to 7 (UR: 6-8) cases per 100,000 population. This was achieved with 100% protection against infection and disease effective both pre- and post-infection and 10 years duration of protection. This averted 11.6 million TB cases (UR: 10.2-12.6) and 270,000 TB deaths (UR: 145,000-483,000) over 2025-2050.
- With the slightly more limited profiles explored in the age targeting modelling, the highest impact vaccine (80% VE, 20 years duration, no mass revaccination campaign, P&PI) averted 502,000 (UR: 431,000-591,000) cases when delivered to adolescents, or 3.0 million (UR: 2.5-3.5m) cases through older adult vaccination campaigns.
- The estimated impact of routine age-targeted vaccination was less than that achieved by broad mass vaccination. However, if resources are limited, targeted vaccination of older adults may help maximise the achievable impact.

6.2.2 Development of a spatial mapping tool to inform clinical trial design

The ePAL study developed, implemented and evaluated a new tool for spatial mapping of TB patients registering at TB clinics in Blantyre, Malawi. Data generated were used to explore notification rates and hotspots of trial-relevant populations.

- A low cost, easy to use app for real-time collection of high-resolution spatial data, was integrated in to routine TB registration.

- 1,899 TB patients were registered using ePAL in the 12-month study period, with high patient acceptance (98.7%).
- ePAL achieved clinic-based collection of patient location of residence accurate to a median of 84m (IQR: 35m-317m) in a high population density urban setting without a municipal address system.
- Data were used to identify areas with high TB burden potentially suitable for TB vaccine trials. To align with a potential trial population, HIV-negative Blantyre city residents with microbiologically confirmed pulmonary disease were mapped and compared to the full data set. Rates were substantially lower than the full data set, and some differences were observed in the spatial distribution of cases, demonstrating the importance of these data for informing clinical trial sample size estimations and targeted recruitment.

6.3 Strengths and limitations of this research

Strengths and limitations of the work are discussed in detail in each chapter. Below is a summary of some of the key strengths and weaknesses of the work in this thesis.

6.3.1 Strengths

6.3.1.1 Systematic review

The systematic review publication included in Chapter 2 is the first and only publication in the literature to bring together the available literature on mathematical modelling of the epidemiological impact of new TB vaccines. This research provides a comprehensive summary of the available literature, and is of significance for informing TB vaccine development as it provides summaries of the research available to inform several key aspects of TB vaccine TPPs. The article also highlights many important research gaps in the literature to help guide future modelling to meet the most urgent research needs.

6.3.1.2 Mathematical modelling

In the mathematical modelling research in this thesis, I developed a highly age-stratified and calibrated model with age-specific natural history parameters and social contact patterns. At time of development, this was the first TB vaccine model to account for heterogeneous social mixing patterns by age. All previous TB vaccine models had assumed contact rates between age groups were equivalent, whereas in this model I incorporated data-informed contact patterns from a study conducted in China. Since development of this model, one other TB vaccine modelling study has incorporated social contact patterns.[1] This study confirmed the importance of the inclusion of social contact patterns and age stratification in TB transmission models, and demonstrated this in the context of TB vaccines, which validates the decisions made in my research to create such a highly age-stratified model incorporating heterogeneous mixing. The Arregui publication is a top-level assessment of the impact of a single vaccine type at the regional level, and uses European mixing patterns scaled to regional demographics to inform social mixing.[1] My research takes this a step further, modelling a comprehensive range of vaccine characteristics at the country level, parameterising with a country-specific contact matrix. The use of country specific data was considered important, as this accounts not just for the impact of demographics on contact patterns, but also includes any social influences on contact patterns.

This research was the first to explore the potential impact of new TB vaccines targeted to older age groups, and to provide a comparison of this approach to the current strategic focus on adolescent vaccination. This required the age-specific model parameterisation and calibration mentioned above, and is the first China TB model to include calibration to age-stratified mortality and notification rates. The results of this work were consistent with previous literature, and strengthen the recommendation that vaccines effective post-infection will be needed in this setting. Previous modelling research has focussed on infant or adolescent vaccination, and mostly recommend vaccination in adolescents or adults. Whereas in this research we develop the argument for vaccinating adults further by demonstrating that in ageing

epidemics, most impact could be achieved by targeting older adults, potentially minimising the resources required per case averted.

This is the first study to provide an in-depth exploration of the impact of such a wide range of vaccine efficacies, durations of protection, and vaccination populations (pre- and post-infection). Such a comprehensive exploration of vaccine characteristics provides a detailed analysis of the impact individually and in tandem of the different characteristics, and also ensures that once trial results become available that modelling results are available for the relevant characteristics.

This modelling has been developed in the context of only two other China TB vaccine modelling studies, one of which was only published in 2017.[2,3] Neither of these existing publications accounted for age specificities in epidemiology or demographics. Further, in these studies, vaccination was either at birth or all-ages, and vaccine characteristics modelled were unclear or unrealistic (e.g. 100% vaccine efficacy). Therefore, this was the first study to employ a model calibrated to age stratified demographic and epidemiological data to explore the impact of TB vaccines and age targeted TB vaccination.

6.3.1.3 ePAL app

The ePAL app is the first of its kind to allow accurate clinic-based collection of place of residence as part of routine NTP care and without the need for home visits. The tool was developed in the context of a clear need for new, rapidly implementable methodology for assessing suitability of trial sites for TB vaccine clinical trials, and ensuring that local data relevant to the populations and outcomes of interest for the trial are used for design of the sample size and recruitment strategy. Other options for collection of such data either lack the appropriate granularity or come at substantial cost both financially and time.

Strengths of ePAL for collection of spatial data include its accuracy, ease of use, low cost, integration into routine NTP data collection with QAQC processes in place, real

time data accessible via a web interface, and ability to function offline with infrequent access to mobile data and electricity.

A key strength of the app was tailoring of content and the user experience through stakeholder involvement: for example, community engagement for identification of locally-relevant POIs in the initial development, and ongoing feedback and improvement of the app and data collection through regular meetings with the TB officers. This may have also contributed to the very high participation rate in the study. The app interface and training were designed for ease of use by individuals without experience of electronic data collection. Error rates were low for single-entry electronic data capture. ePAL-based data collection was successfully incorporated in to the National Tuberculosis Programme through this integration in to routine patient registration. ePAL is an easy to use tool, with NTP staff able to easily and effectively use the ePAL app after one day of training.

ePAL development and implementation was rapid, with POI collection for local adaptation of the app, training of data collectors, implementation, and a run-in period for refining the app and field usage achieved in less than five months. With lessons learned from this study, future sites implementing ePAL could achieve this even faster. ePAL implementation and evaluation were also achieved at low cost, thanks to use of relatively inexpensive technology, and the need for minimal full-time study staff. To employ the data to rapidly inform trial design, the generation of data in real time will be beneficial, and data can be monitored remotely in real-time via the web interface.

6.3.2 Limitations of this work

6.3.2.1 Systematic review

The main limitation of the systematic review was the single-reviewer approach to literature sifting. Risks from this approach were minimised as the sifting was repeated in order to update the review, therefore effectively double sifting by a single

reviewer. In addition, a low threshold was employed for checking sifting queries with a second reviewer.

Secondly, although an adapted quality scoring tool was developed specifically for this study, the tool was not well suited to assessing analytic models.

6.3.2.2 Mathematical modelling

Limitations in the TB vaccine models developed can be considered as structural, assumption, calibration or outcome limitations.

Structural limitations in the model included the method of modelling reinfection of recovered populations, and of case detection.

In this model, recovered populations experiencing reinfection that does not progress directly to disease transitioned to the latent state to become “slow progressors”. This was likely an underestimate of their risk of progression, and could be improved by retaining reinfected recovered populations in the recovered state. Such a change would not have an effect on the direction of impact or conclusions of this research. It would be anticipated to marginally increase the number and proportion of cases arising from the recovered population and would likely strengthen the conclusion that older adult vaccination with post-infection vaccines would provide greater impact than adolescent vaccination.

In terms of case detection, the model is structured such that a proportion of incident cases are detected, parameterised to align with WHO-reported case detection ratio data. Those who are not detected enter the prevalent pool and can only exit via TB death, death from other causes, or natural cure. A more realistic approach to modelling case detection would most likely be to reduce case detection of those entering the prevalent state, and allow care seeking with worsening symptoms in the prevalent state. Data are not readily available to parameterise the model according to those detected within 6 months of symptoms and those detected later. However,

if such data were available, the model could be re-structured to allow a certain proportion of case detection to occur within 6 months of onset, and the remainder to occur over time, with the rate of detection linked to disease progression. This would be achieved by developing a series of disease states representing progression of untreated disease, and each state would be associated with a case detection ratio. However, as mentioned above, data are not available to parameterise this model structure.

Assumption limitations included the assumptions around the future of existing and new control measures, natural history parameters and immunosenescence in elderly populations, unchanging contact patterns over time and feasibility of vaccination campaigns.

In the model, it was assumed that no new interventions would be developed over the time frame of this model. The development landscape for drugs and diagnostics is progressing, but was considered too uncertain to be able to include other new interventions in this research. The introduction of new diagnostics and treatment regimens could potentially alter the course of the epidemic. If this were to occur, the estimations of absolute impact of new TB vaccines could potentially be an overestimate. The relative impact of the different vaccine types would not be anticipated to be affected. Perhaps with the exception of a highly safe and effective treatment for clearance of latent infection, which if safe in the elderly could impact the relative impact of the different vaccine types. Due to high current treatment success and case detection estimates in China, these were assumed to plateau. If improvements in current care measures were achieved, the impact achieved by vaccines could potentially be reduced. However, this would not affect conclusions with respect to age targeting unless such an intervention varied in effectiveness by age.

There is a notable data gap with regards to elderly TB natural history parameters. We represented this uncertainty by sampling from elderly parameter priors spanning

HIV-negative and HIV-positive adult ranges in the calibration process. The extreme end of the HIV ranges may not be reflective of immunesenescence, but fitting methods allowed sampling from this range to allow for the potential impact of immunosenescence. Immunosenescent waning of vaccine efficacy was also an unknown, so was assumed 2% per year waning in the elderly population in the main analysis, and explored up to 5% in sensitivity analyses. Results were robust to immunosenecent sensitivity analyses.

Given population ageing, the assumption that the contact matrix remained constant throughout the study period may provide conservative impact estimates. Larger older adult and elderly populations in the future could increase the transmission from these age groups. This could be modelled in future sensitivity analyses, though could only account for demographic changes, changes caused by societal differences cannot be predicted.

The vaccination campaigns modelled were considered feasible based upon expert opinion and implementation of other vaccines in China and elsewhere. In reality, success of campaigns will be reliant on the vaccine developed, the prioritisation of TB as a public health problem, and the availability of facilities or funding to develop facilities for the large-scale manufacture of TB vaccines. At the global level, adolescent and adult vaccination with just 20% coverage has been estimated to require seven dedicated manufacturers,[4] therefore the 70% coverage mass campaigns modelled here may be feasible for a country with as much manufacturing capacity as China,[5] but may not be feasible from a global perspective. This research does not explore all possible routine vaccination ages as those modelled were considered appropriate given the anticipated future epidemiology and the existence of platforms for vaccine delivery. Once a vaccine is available and the characteristics are known, modelling could be used to further optimise the older adult vaccination age, and explore different mass campaign options.

Calibration limitations included the lack of data to separately calibrate latent and recovered populations, and the challenge of calibrating to the low TB mortality rates in China.

It was found in the calibration process that the rates of reactivation from latency or after recovery from disease were inversely associated and mutually compensatory. These two populations were the main contributors to development of disease; therefore, this inverse relationship was a function of calibration to incident disease data. Unfortunately, incidence data stratified by latent and recovered populations do not exist to allow independent calibration for these two populations. Fitting to new versus previously treated TB notification data was considered, but the data were considered not directly reflective of the modelled population, and have several inherent biases.

The empirically derived case fatality rate (CFR) was particularly low in China relative to historic evidence on CFRs in untreated populations,[6] therefore model calibration employed a calibration factor to scale down the case fatality rate to align with the data. As with other modelling exercises,[7] it was assumed that good access to TB care through CDC facilities led to low case fatality in China.

The main outcome limitation was the time horizon of the model, as this may influence the relative impact of vaccine types. Pre-infection vaccines, for example, tend to perform better over longer time horizons.[8] However, beyond 2050 there would be too much uncertainty in the model. Full benefit of vaccines delivered pre-2050 would extend beyond 2050, therefore relative vaccine impact may change over longer horizons.[8] Results presented from the models were specific to the 2025-50 time horizon.

The results presented in the thesis are representative of the Chinese epidemic, which is an epidemic driven by reactivation disease in an ageing population that has experienced high historical exposure. It should be noted that the relative impact of

different characteristics is likely to differ in other epidemiological scenarios, and will be investigated in the continuation of this work exploring vaccine impact in India and South Africa.

6.3.2.3 ePAL app

Study limitations included challenges associated with home visits for evaluation measurements, initial case capture, and the resultant missing data.

The main limitation was the difficulty of conducting home visits in high density urban areas. These were conducted by contacting the participant to arrange a home visit and following verbal descriptions given by the patient either to the study staff or TB officers. This was standard practice before ePAL but poses challenges for achieving follow up visits. Use of ePAL coordinates was avoided to prevent observer bias. Tracking down the correct households proved very time-consuming, leading to some long delays between registration and the home visit. While this highlights the need for alternatives such as ePAL, and analyses indicated no significant bias in ePAL accuracy with delayed follow up, it is possible that incomplete capture may have biased our estimates of ePAL accuracy if, for instance, deliberately incorrect information had been provided or patients living in more distant locations were harder to track down for follow up visits.

Fourteen percent of evaluation visits were >1km from the place of residence registered in ePAL. Analyses suggest this may be associated with factors relevant to job insecurity/mobility or concerns regarding TB/HIV stigma. Informal feedback suggested that some of the distant patients relocated between ePAL collection and follow up, due to staying with relatives during treatment or the severe flooding in early 2015. More systematic qualitative interviews could help understand the issues underlying large distance errors, and to identify potential solutions for minimisation of this issue. Shortening the time between registration and follow up could help reduce the number relocating, but given relocation or multiple addresses may be

pertinent information for mapping and for patient follow up, the option for identifying multiple residencies could be a helpful addition to the ePAL app.

Case capture was low early in the study due to insufficient resources at the central hospital, but dramatically improved after addition of a staff member and a second data collection tablet.

For the GPS coordinates measured at the participant's home, it should be noted that even with taking averaged waypoints an inherent inaccuracy of several metres still exists. In addition, poor weather conditions and satellite signal blocking by proximity to buildings in high-density areas are likely to have affected measurement accuracy in this setting. Therefore, several metres of unavoidable 'inaccuracy' will have been systematically introduced, even if the correct building were selected.

6.4 Contribution of the thesis to advancing knowledge in TB vaccine development

Epidemiological evidence was identified by the TB vaccine development community as a critical research need for advancing the development of new TB vaccines.[9] This encompassed two main areas of research: 1) the need for data-driven identification of vaccine characteristics and target populations for vaccination that could achieve the greatest reduction in disease burden, and 2) a better understanding of local epidemiology to inform the selection of trial sites, recruitment populations and design clinical trials.

The existing mathematical modelling evidence to inform TB vaccine development strategy had not previously been summarised. To identify available literature for informing the first research need, in the first section of this thesis I conducted the first ever systematic review of the available mathematical modelling literature exploring the epidemiological impact of new TB vaccines. This review provides TB

vaccine developers a comprehensive summary of the range of available epidemiological modelling data to inform vaccine development decisions. This review identified several substantial research gaps for informing TB vaccine development. These included the impact of prevention of infection and disease efficacy combinations, the division in the literature as to the impact of pre- versus post-infection vaccines, the impact of varying duration of protection, the dearth of TB vaccine modelling literature for China, and the impact of vaccinating older adults.

Understanding the relative impact of vaccine characteristics and implementation is central to data-driven decision making when designing TB vaccine target product profiles. The research presented in the second section of the thesis (Chapter 3 and 4) provides a comprehensive exploration of the impact of new TB vaccine characteristics and a comparison of adolescent versus older adult age targeting of new vaccines in China. Target product profiles currently under development lacked the modelling data needed to inform the ideal and minimum vaccine characteristics, thus the comprehensive exploration of vaccine characteristics in the research presented can inform these characteristics. In particular, the results highlighted the importance of prevention of disease vaccines effective in populations latently infected and recovered from disease, which has direct implications for clinical trial design. The results highlighted the need for duration of protection of at least 5 years with 10-yearly mass adult vaccination campaigns, suggesting longer durations of trial follow up, or for planning for more frequent mass adult vaccination campaigns, will be required.

Older adults and the elderly were previously excluded from most clinical trials, and the impact of older adult vaccination with new TB vaccines had never been explored in the literature. The age targeted vaccination modelling results informed the inclusion of the elderly population in the Gates TB vaccine strategy as a secondary population of interest.[10]

The results of this research can directly inform the characteristics detailed in target product profiles, clinical trial outcomes and recruitment populations for vaccines with planned use in China and epidemiologically similar settings, to ensure that vaccines are developed for and delivered to populations that could maximise population-level impact of new TB vaccines.

Modelling demonstrated the importance of disease outcomes in TB vaccine clinical trials. Poor availability of epidemiological data in clinical trial sites limits data-informed design of prevention of disease efficacy studies. It is hoped that data generated by the ePAL app will assist site selection and trial design for TB vaccine trials, and in particular allow identification of hotspots and estimation of disease rates in potential recruitment populations. ePAL is the first of its kind to provide a clinic-based electronic data-capture tool for identifying TB patient place of residence, developed using high resolution satellite maps developed through community engagement. For the research site in Blantyre it provides the necessary data to inform trial design, and could be rapidly and inexpensively translated to other settings.

Although further research is needed to completely fulfil the two epidemiological research needs identified in the TB vaccine blueprint,[9] the research in this thesis substantially moves forward the TB vaccine modelling literature through its attention to age-specificities in the modelling methodology, exploration of age targeting vaccination to older age groups, and comprehensive exploration of the impact of vaccine characteristics in a high burden, previously under-modelled setting. Availability of appropriate epidemiological data is a significant obstacle to TB vaccine trial site selection and design. The ePAL app provides a setting-appropriate tool for the collection of the spatial data required for new or existing clinical trial research sites to generate the data needed for site selection, sample size calculations and recruitment design. It is hoped that ePAL will provide the methodology to fulfil this second major epidemiological research need in a resource and time efficient manner in other sites.

6.5 Discussion of PhD research in the context of recommendations for TB vaccine development strategy

Decisions in TB vaccine development, in the form of target product profiles and clinical trial design, are taken by a variety of stakeholders, including cross-product bodies (e.g. Aeras, TBVI, WHO, BMGF) and vaccine developers (e.g. academia and industry). As an academic partner to these stakeholders, the reach of the research presented in this thesis is limited to the generation of evidence and tools to inform these decisions; however, it is hoped that the results of this research will inform vaccine development. Recommendations for development strategists (e.g. BMGF, WHO, cross-product bodies, developers), clinical triallists, and country level decision makers are provided below.

6.5.1 For development strategists and clinical triallists

Although there is some value in prevention of infection, efficacy for prevention of disease will be key to maximising impact on disease burden in China during 2025-2050. Higher incidence rate reductions by 2050 were only achievable with non-zero vaccine efficacy against disease. Therefore, knowledge of only the vaccine efficacy for prevention of infection would be insufficient to guarantee high incidence rate reductions. Therefore, vaccine development strategies for China and similar epidemics should aim to accelerate development of candidate vaccines with at least some anticipated protection against disease.

Stage gating between trials may be based upon the efficacy expected to deliver a pre-defined future population-level impact. If only one outcome (i.e. disease *or* infection) is measured in a clinical trial, to be conservative it must be assumed that there is no efficacy against the other clinical outcome, therefore the required vaccine efficacy to proceed with that candidate will be higher than if both outcomes were measured.

Vaccines effective post-infection were shown to be essential for maximising future impact in China. Ideally clinical trials should enrol both pre- and post-infection

populations of sufficient sample size to allow stratification of efficacy estimates by infection status. However, the required sample size would likely be infeasible, so at a minimum, trials should recruit post-infection populations to demonstrate safety, and potentially conduct stratified analyses as secondary outcomes.

To achieve a minimum incidence rate reduction of 20-29%, at least 5 years protection was required if mass campaigns were every 10 years. Therefore, clinical trials should be designed to provide data on duration of protection beyond the usual 2-3 years of a phase IIb trial. For example, phase II clinical trials could include an extended immunological or infection/disease outcome follow-up in a subset of participants.

Adolescent-targeted vaccines, the focus of current development plans, would have low impact in ageing, reactivation-driven TB epidemics like China. In these settings, an efficacious post-infection vaccine delivered to older adults will be critical to maximise population-level impact and would provide a crucial contribution towards achieving the WHO 2050 TB goals. Older adults should, therefore, be included in TB vaccine clinical development and implementation planning. To allow on-label use following vaccine registration, older age groups need to be included in clinical trials, ideally from phase IIB, but at a minimum by phase III, to ensure that sufficient participants have been exposed to the investigational product to demonstrate safety and immunogenicity to the regulators.

Maps generated using data from ePAL demonstrated the spatial heterogeneity of TB disease in Blantyre, Malawi, highlighting the importance of local-level data to inform trial recruitment strategies. The large number of patients resident outside of Blantyre but registering in urban clinics is consistent with other literature,[11] and highlights the importance of not directly using district level data for sample size calculations, as this would underpower the study. In addition, the rates and spatial distribution of disease may also change when considering trial-relevant populations (e.g. microbiological outcomes, HIV positive patients, age of recruitment), therefore ensuring sample sizes

are calculated based upon data relevant to the outcomes and population of interest are also important to ensure a sufficiently powered study,

6.5.2 For country-level decision makers

With longer durations of vaccine protection or more frequent revaccination strategies, large epidemiological impact could be achieved. Therefore, if duration of protection of new vaccines is found to be low, country-level planning for more frequent revaccination campaigns will be critical to maximise the population level impact.

The estimated impact of routine age-targeted vaccination was less than that achieved by broad mass vaccination. However, it is essential that older adult age groups are included in whichever vaccination campaign is conducted. If resources are limited, targeted vaccination of older adults may help maximise the achievable impact. Even with reduced coverage, older adult vaccination was the preferred vaccination age with post-infection or pre- and post-infection vaccines. Therefore, older adult platforms may need to be developed or strengthened for delivery of vaccines to older age groups.

6.6 Opportunities for future research

The research presented in this thesis is just the first step towards fulfilling the epidemiological and modelling research needs for TB vaccine development. Many additional questions were identified in the systematic review or have arisen as this work has developed. Planned and potential future research is identified and discussed.

6.6.1 Mathematical modelling of new TB vaccines

The modelling in this thesis aimed to provide a detailed exploration of TB vaccines in China, enabled by the development of models reflective of the epidemiological and demographic trends in this setting. Exploration of vaccines in China was important, as it is a high TB burden country contributing substantially to global TB burden, with a declining epidemiological trend due to successful control programmes but is insufficient to meet 2050 elimination goals, and an ageing population. In addition, given phase III results are imminent from a TB vaccine trial in China, if efficacy is observed, the results of this research will allow estimation of potential future impact and planning for implementation.

Older adult vaccination was modelled instead of elderly vaccination, due to concerns regarding safety and immunogenicity. However, zoster vaccines have recently been shown to be safe and efficacious in the elderly.[12,13] This is an important development for vaccination and longer term protection of these oldest age groups, and could potentially pave the way for how TB vaccines are developed moving forward. If the elderly were included in future trials, modelling direct vaccination of the elderly could be of interest using the China model.

The results from China are likely applicable to other high-burden countries with declining, reactivation-driven epidemics and ageing populations (e.g. Thailand, Russia). However, other high-burden settings with higher transmission rates (e.g. India) and substantial HIV co-epidemics (e.g. South Africa) may require slightly different tailoring of vaccine development strategies. I have received funding from the Bill and Melinda Gates Foundation to re-calibrate the TB vaccine model to the epidemics in India and South Africa. Modelling the South African epidemic will require the incorporation of an HIV stratum in to the model, and consideration of the potential relative efficacy of new TB vaccines in HIV-positive adults. In these settings, it may be of interest to compare infant, adolescent and older adult vaccination.

The systematic review identified modelling the impact of TB vaccines on multi-drug resistance (MDR) as a research gap. We have recently received funding from Aeras to model the epidemiological, health economic and budget impact of new TB vaccines on MDR-TB. The addition of health economic analyses would be a valuable addition to this work, and a valuable tool for advocating investment in new TB vaccines.

In the modelling research, I assumed that no game-changing new drugs or diagnostics would be introduced over the time frame of this model. However, if other new tools were developed, they could affect the impact of new TB vaccines. The MRC has funded a PhD student to join our research team to explore the potential impact of other new interventions on the future impact of new TB vaccine.

6.6.2 ePAL app

The development and evaluation of ePAL is hopefully just the first step in the use of this app for supporting development of new TB vaccines.

As described in Chapter 5, the use of ePAL in Blantyre has generated a large data set that has not yet been fully explored. In future work, I plan to conduct additional analyses exploring the spatial distribution of TB in Blantyre and associated risk factors, to help inform current public health measures, and in the hope of informing future clinical trials. These analyses will include exploring risk factor data, employing the new enumeration data set as a denominator, conducting observed versus expected analyses to identify statistically significant TB hotspots, exploring the distribution of TB disease in children in Blantyre.

Recent research has also indicated the potential importance of locations outside of the home for *M.tb* transmission. [14] As has been done for water sources in the Blantyre typhoid study, it would be of interest to use ePAL to explore other locations associated with *M.tb* transmission risk, such as churches, bars and work places attended. This would likely be too time consuming to request of the TB Officers

currently collecting data using ePAL, but could potentially be conducted for a subset of patients in Blantyre by study staff.

It is also hoped that ePAL will continue to be employed in other studies in Blantyre, and I am currently exploring the possibility of implementing ePAL in studies in South Africa and India.

An interesting future avenue to bring together the epidemiological and mathematical modelling research, would be to employ spatial data supplied by ePAL to parameterise a spatial TB model to explore impact of vaccination programmes at the community level. One study has explored TB vaccination in a context of spatial heterogeneity in India. However, this was based upon larger areas, therefore such a detailed local data set could allow for interesting local-level modelling. Collaborators at the London School of Hygiene and Tropical Medicine plan to employ the ePAL data set to inform mathematical modelling of spatially targeted interventions. The current research plans focus on existing interventions, but could in the future be adapted to explore the impact of spatially targeted new TB vaccines.

6.7 Conclusion

There is a clear and urgent public health need for new TB vaccines. The TB vaccine pipeline is the strongest to-date, yet progression of candidates through the pipeline remains slow. Epidemiological data to inform TB vaccine development strategy and to assist trial site selection and design was identified as a critical research need to accelerate development and maximise impact of new TB vaccines.

The research presented in this thesis informs the development of appropriate TB vaccines and target populations to maximise future population-level impact and accelerate TB vaccine development. A prevention of disease vaccine efficacious post-infection and delivered to older adults would contribute towards maximising population-level impact in China. Adolescent-targeted tuberculosis vaccines are likely to have low impact in ageing, reactivation-driven epidemics like China, which suggests a modification of the current strategic focus on adolescents. Clinical trials should assess disease endpoints, include *M.tb*-infected and older adult populations, and extend beyond the usual 2-3 years follow up. To support design of disease endpoint trials, ePAL may provide an accurate, easily implementable, low-cost tool for identification of areas of high TB burden in settings without addresses.

It is hoped that this research, and planned future research, will inform how TB vaccine developers, clinical triallists, and country-level public health decision makers prioritise and plan their approach to the development and implementation of new TB vaccines. Continued development of epidemiological models and data collection tools will be essential for the acceleration of TB vaccine development and to maximise future impact of new TB vaccines.

6.8 Chapter 6 References

1. Arregui S, Sanz J, Marinova D, et al. A data-driven model for the assessment of age-dependent patterns of Tuberculosis burden and impact evaluation of novel vaccines. *bioRxiv* 2017: Online first.
2. Dye C, Glaziou P, Floyd K, Raviglione M. Prospects for tuberculosis elimination. *Annu Rev Public Health* 2013; **34**: 271-86.
3. Liu S, Li Y, Bi Y, Huang Q. Mixed vaccination strategy for the control of tuberculosis: A case study in China. *Math Biosci Eng* 2017; **14**(3): 695-708.
4. Aeras and TBVI. TB Vaccine Research and Development: A Business Case for Investment. http://www.aeras.org/pdf/TB_RD_Business_Case_Draft_3.pdf (accessed 14th January 2017 2017).
5. Wang J. [Small vaccines contain big market] (In Chinese). *Guide of China Medicine* 2005; **12**: 98-9.
6. 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.
7. Houben RM, Menzies NA, Sumner T, et al. Feasibility of achieving the 2025 WHO global tuberculosis targets in South Africa, China, and India: a combined analysis of 11 mathematical models. *The Lancet Global health* 2016; **4**(11): e806-e15.
8. Harris RC, Sumner T, Knight GM, White RG. Systematic review of mathematical models exploring the epidemiological impact of future TB vaccines. *Hum Vaccin Immunother* 2016; **12**(11): 2813-32.
9. Brennan MJ, Thole J. Tuberculosis vaccines: a strategic blueprint for the next decade. *Tuberculosis (Edinburgh, Scotland)* 2012; **92 Suppl 1**: S6-13.
10. Hanekom W. Revision of the Bill and Melinda Gates Foundation TB vaccine Strategy – 2014. 2014. https://drive.google.com/file/d/0B2K5XWn1bjpORmpHLUdoOXJFTlk/view (accessed 4th September 2016).
11. Subbaraman R, Thomas BE, Sellappan S, et al. Tuberculosis patients in an Indian mega-city: Where do they live and where are they diagnosed? *PloS one* 2017; **12**(8): e0183240.
12. Cunningham AL, Lal H, Kovac M, et al. Efficacy of the Herpes Zoster Subunit Vaccine in Adults 70 Years of Age or Older. *New England Journal of Medicine* 2016; **375**(11): 1019-32.
13. Lal H, Cunningham AL, Godeaux O, et al. Efficacy of an adjuvanted herpes zoster subunit vaccine in older adults. *N Engl J Med* 2015; **372**(22): 2087-96.
14. 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 infectious diseases* 2016; **16**(1): 71.

APPENDICES

Appendix A: Additional systematic review materials

Protocol: Systematic review of models exploring the epidemiological impact of novel TB vaccines

The following protocol was registered at on PROSPERO.[1]

Prepared by: Rebecca Harris

Document date: 15th December 2015

Version: 1

Literature search strategy

1.0 RESEARCH OBJECTIVES

Background

With a strong pipeline of 15 novel tuberculosis (TB) vaccine candidates in clinical trials,[2] there is a growing modelling literature exploring the potential impact of novel TB vaccines. Such models can be used to inform decision making with regards to the characteristics and target population for novel vaccines, to allow limited funding to be appropriately prioritised to ensure the greatest public health impact would be achieved when new vaccines come to market. To-date, the body of literature on this topic has not been summarised. The objective of this review is to narratively summarise the methodology and outcomes from modelling research exploring the epidemiological impact of pipeline or theoretical TB vaccines.

1.2 Research Question

What are the methods and estimated epidemiological impact of pipeline or theoretical human tuberculosis vaccines as estimated by mathematical modelling?

The PICOT framework for this research question is as follows:

Table 1: PICOT question

Limit	Definition
Population	Humans, of any age in any country or globally
Intervention	Novel/theoretical/pipeline TB vaccines *Not* BCG-only
Comparator	No intervention, currently available interventions (at current or scaled-up levels), or other theoretical interventions.
Outcome	Tuberculosis epidemiological impact (e.g. incidence, prevalence, mortality, number needed to vaccinate, cost effectiveness) Not Mycobacterium bovis
Time	No limit
Study Design	Epidemiological mathematical models *Not* within-host impact models *Not* reviews/commentaries

2.0 METHODS

2.1 Search Strategy

Three databases (Pubmed, Embase and WHO Global Index Medicus (GIM)) providing access to seven databases/indexes (Pubmed/Medline, Embase, African Index Medicus, LILACS, SEARO Index Medicus, WPRO Index Medicus, EMRO Index Medicus) will be searched. Searches will use free text and Mesh/Emtree/DeCS terms tailored by database for groups of terms covering tuberculosis, modelling and vaccines (table 2). Search terms are combined with Boolean OR within groups, and by Boolean AND

between search term groups. The search in Pubmed will be run with filter “human”, and the WHO GIM search will be limited to regional databases to avoid duplication.

Table 2: Search terms by database

Database	Search Term Group		
	Modelling	Tuberculosis	Vaccine
Pubmed	"Models, Theoretical"[Mesh]) OR "mathematical model*"	TB OR tuberculosis OR "Tuberculosis"[Mesh]	vaccin* OR immuniz* OR immunis* OR "Tuberculosis Vaccines"[Mesh]
Embase	("mathematical model$".mp. OR mathematical model.mp. or mathematical model/)	(tuberculosis control/ or exp tuberculosis/ or Mycobacterium tuberculosis/ or tb.mp. or tuberculosis.mp.)	(exp vaccine/ or (vaccin$ or immunis$ or immuniz$).mp.)
WHO Global Index Medicus (regional databases)	("computer models" OR "epidemiologic models" OR "mathematical models")	TB OR tuberculosis OR "tuberculosis"	Not required as very few hits with first two search terms

Table 3: Application of PICOT through search limits and manual search criteria

Limit	Definition	Limit management	
Population		Humans, of any age in any country or globally	“Human” search limit (pubmed only) and manual search criterion
Intervention		Novel/theoretical/pipeline TB vaccines *Not* BCG-only	Vaccine search terms and manual search criteria
Comparator		No intervention, currently available interventions (at current or scaled-up levels), or other theoretical interventions.	Manual search criteria

Outcome		Tuberculosis epidemiological impact (e.g. incidence, prevalence, mortality, number needed to vaccinate, cost effectiveness) Not Mycobacterium bovis	TB search terms and manual search criteria
Time		No limit	No limit applied
Methodology		Epidemiological mathematical models *Not* within-host impact models *Not* reviews/commentaries	Modelling search terms manual search criteria

2.2 Selection of studies and data extraction

The database searching, sifting and data extraction will be conducted in by a single reviewer (RCH). A three-stage sifting process will be employed to screen publications first at title, abstract, then at full text level for eligibility for inclusion using the below-details inclusion and exclusion criteria. Any uncertainties in inclusion will be decided through discussion with a second reviewer (RW). Reference lists of included studies will be hand searched for studies meeting the inclusion criteria. Onward citation searching will be conducted for all included articles.

Inclusion Criteria

- Mathematical model
- Systematic review of models of novel TB vaccines, or commentary adding to the analyses/interpretation of models reported elsewhere
- Intervention is novel/future/hypothetical vaccine against tuberculosis or of an unspecified novel TB intervention with characteristics in-line with a vaccine

- Reported outcomes are of epidemiological impact of vaccine (e.g. incidence, prevalence, mortality, number needed to vaccinate, cost effectiveness)

Exclusion Criteria

- Within-host/immunological vaccine impact models
- Review or commentary not adding to existing body of knowledge
- TB epidemiological models not reporting impact of vaccine
- TB epidemiological models reporting only interventions other than vaccines
- Model only reporting on impact of BCG with known/fixed efficacy
- Disease or infection caused by *Mycobacterium bovis* or other non-*Mycobacterium tuberculosis* strain.

Data will be extracted from those papers selected for final inclusion using a standardised Microsoft Excel® database. Extracted data will include study objectives, model methods (e.g. model structure, mixing patterns, model assumptions, parameter data sources), intervention characteristics (e.g. vaccine efficacy, duration of protection, waning of protection, vaccine target population and schedule), epidemiological outcomes, sensitivity analyses, and model limitations.

3.0 QUALITY ASSESSMENT AND DATA SYNTHESIS

3.1 Quality Assessment

Few validated tools exist for the assessment of quality of epidemiological modelling studies. Fone *et al.* (2003) have proposed and tested an adapted version of the Weightman *et al.* (2000) tool for the critical appraisal of the quality of modelling

studies in health care.[3,4] And more recently Caro *et al*. have detailed the criteria for good research practice in modelling research.[5] In this review, the Fone *et al.* tool will be used as a basis for quality assessment, but updated based upon the Caro *et al.* report to ensure clarity and coverage of certain criteria.[3,5] Based upon extracted data, each included article will be critically appraised for quality using a piloted version of this extraction form.

3.2 Data Synthesis

Extracted data will be synthesised using a narrative approach and will focus on the modelling methods used, estimates of epidemiological impact, and identifying evidence gaps or limitations.

3.3 Dissemination

The review manuscript will be submitted for publication in a peer reviewed journal.

4.0 **References**

1. Harris RC, White RG, Sumner T, Knight GM. Systematic review of models exploring the epidemiological impact of novel TB vaccines. PROSPERO 2016:CRD42016033266 2016. http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42016033266 (accessed March 7th 2016).
2. World Health Organization. Global Tuberculosis Report 2015, 2015.
3. Fone D, Hollinghurst S, Temple M, et al. Systematic review of the use and value of computer simulation modelling in population health and health care delivery. *Journal of public health medicine* 2003; **25**(4): 325-35.
4. Weightman A, Barker J, Lancaster J. Health Evidence Bulletins Wales Project Methodology 3. . Cardiff: Department of Information Services: UWCM, 2000.
5. Caro JJ, Briggs AH, Siebert U, Kuntz KM. Modeling good research practices--overview: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force--1. *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research* 2012; **15**(6): 796-803.
6. Abu-Raddad LJ, Sabatelli L, Achterberg JT, et al. Epidemiological benefits of more-effective tuberculosis vaccines, drugs, and diagnostics. *Proceedings of the National Academy of Sciences of the United States of America* 2009; **106**(33): 13980-5.

7. Bhunu CP, Garira W, Mukandavire Z, Magombedze G. Modelling the effects of pre-exposure and post-exposure vaccines in tuberculosis control. *Journal of theoretical biology* 2008; **254**(3): 633-49.
8. Lietman T, Blower SM. Potential impact of tuberculosis vaccines as epidemic control agents. *Clin Infect Dis* 2000; **30 Suppl 3**: S316-22.
9. Ziv E, Daley CL, Blower S. Potential public health impact of new tuberculosis vaccines. *Emerg Infect Dis* 2004; **10**(9): 1529-35.
10. Gomes MG, Franco AO, Gomes MC, Medley GF. The reinfection threshold promotes variability in tuberculosis epidemiology and vaccine efficacy. *Proceedings Biological sciences / The Royal Society* 2004; **271**(1539): 617-23.
11. Murray CJ, Salomon JA. Modeling the impact of global tuberculosis control strategies. *Proceedings of the National Academy of Sciences of the United States of America* 1998; **95**(23): 13881-6.
12. Dye C, Williams BG. Eliminating human tuberculosis in the twenty-first century. *Journal of the Royal Society, Interface / the Royal Society* 2008; **5**(23): 653-62.
13. Dye C, Glaziou P, Floyd K, Raviglione M. Prospects for tuberculosis elimination. *Annu Rev Public Health* 2013; **34**: 271-86.
14. Knight GM, Griffiths UK, Sumner T, et al. Impact and cost-effectiveness of new tuberculosis vaccines in low- and middle-income countries. *Proc Natl Acad Sci U S A* 2014; **111**(43): 15520-5.
15. Dye C. Tuberculosis 2000-2010: control, but not elimination [The Comstock Lecture]. *The International Journal of Tuberculosis and Lung Disease* 2000; **4**(12): S146-S52.
16. ReVelle CS, Lynn WR, Feldmann F. Mathematical models for the economic allocation of tuberculosis control activities in developing nations. *American review of respiratory disease* 1967; **96**(5): 893-909.
17. Cohen T, Colijn C, Murray M. Modeling the effects of strain diversity and mechanisms of strain competition on the potential performance of new tuberculosis vaccines. *Proc Natl Acad Sci U S A* 2008; **105**(42): 16302-7.
18. Tseng CL, Oxlade O, Menzies D, Aspler A, Schwartzman K. Cost-effectiveness of novel vaccines for tuberculosis control: a decision analysis study. *BMC Public Health* 2011; **11**: 55.
19. Ditkowsky JB, Schwartzman K. Potential cost-effectiveness of a new infant tuberculosis vaccine in South Africa--implications for clinical trials: a decision analysis. *PloS one* 2014; **9**(1): e83526.
20. Pienaar E, Fluitt AM, Whitney SE, Freifeld AG, Viljoen HJ. A model of tuberculosis transmission and intervention strategies in an urban residential area. *Comput Biol Chem* 2010; **34**(2): 86-96.
21. Channing L, Sinanovic E. Modelling the cost-effectiveness of a new infant vaccine to prevent tuberculosis disease in children in South Africa. *Cost Effectiveness and Resource Allocation* 2014; **12**(1): 1-9.
22. Dye C. Making wider use of the world's most widely used vaccine: Bacille Calmette-Guerin revaccination reconsidered. *Journal of the Royal Society, Interface / the Royal Society* 2013; **10**(87): 20130365.

23. Castillo-Chavez C, Feng Z. Global stability of an age-structure model for TB and its applications to optimal vaccination strategies. *Mathematical biosciences* 1998; **151**(2): 135-54.
24. Gabriela MGM, Rodrigues P, Hilker FM, et al. Implications of partial immunity on the prospects for tuberculosis control by post-exposure interventions. *Journal of theoretical biology* 2007; **248**(4): 608-17.
25. Hawn TR, Day TA, Scriba TJ, et al. Tuberculosis vaccines and prevention of infection. *Microbiology and molecular biology reviews : MMBR* 2014; **78**(4): 650-71.
26. Rahman M, Sekimoto M, Takamatsu I, et al. Economic evaluation of universal BCG vaccination of Japanese infants. *International journal of epidemiology* 2001; **30**(2): 380-5.
27. Rodrigues P, Margheri A, Rebelo C, Gomes MG. Heterogeneity in susceptibility to infection can explain high reinfection rates. *Journal of theoretical biology* 2009; **259**(2): 280-90

Risk of bias tool for assessment of epidemiological modelling studies

	Criterion (adapted from Fone *et al.* and Caro *et al.)*	**Considerations (adapted from Fone et al. and Caro et al.)**	**Score considerations (0, poor to 2, good)**	
1	Are the aims and objectives clear?	Are the research questions and modelling objectives clearly defined?	0 Not stated 1 Stated but vague 2 Stated and focussed	Definitions: max 8 points
2	Is the setting and population clearly defined?	Does the paper clearly state the setting (e.g. geographical location, high/low TB burden)?	0 Not stated 1 Stated but vague or details missing 2 Stated and focussed	
		In health economics models, has the perspective been stated?		
		Does the paper clearly state the modelled population? (e.g. patient or population group characteristics)		
		Have sub-populations necessary for the research question and setting been modelled?		
3	Are the intervention and comparators adequately defined?	Does the paper clearly state the population(s) targeted for vaccination?	0 Not stated or very unclear 1 Stated but details missing 2 Stated and all necessary details stated	
		Does the paper clearly define the vaccine characteristics (e.g. vaccine efficacy, duration of protection, number of doses, waning, timing)?		
		If there is a comparator (no vaccine, baseline or alternative intervention scenario), is it clearly defined?		

4	Are the outcome measures defined and answer the research question?	Does the paper clearly define the outcomes of interest? Do the outcomes correspond to the research question?	0 Not stated, very unclear or not suited to research question 1 Stated but details missing or not directly aligned with research question 2 Stated, all necessary details stated, and aligned with research question	
5	Are the model structure and time horizon clearly described and appropriate for the research question?	Is the model structure clearly reported and appropriate for the research question? Does the model reflect current knowledge of disease natural history? Is the time horizon and time step of the model clearly stated and appropriate to the research question (i.e. is it long enough to capture health effects)?	0 Not appropriate model structure, or poor/no description of model 1 Incomplete description, and/or appropriate in part for research question 2 Complete and reproducible, appropriate structure and time horizon	Model methods: max 4 points
6	Are the modelling methods appropriate for the research question and adequately described?	Were the modelling methods clearly described, and suited to the research question?	0 Not appropriate model structure, or poor/no description of methods 1 Incomplete description, and/or appropriate in part for research question 2 Complete and reproducible, appropriate method	
7	Are the parameters, ranges and data sources specified?	Are all parameters and their ranges reported? Are the data sources for parameters reported?	0 Poorly reported 1 Some information missing	Model inputs: max 6 points

			2 Complete reporting of parameters, ranges and data sources	
8	Are any assumptions explicit and justified?	Are all assumptions explicit and justified?	0 Not reported 1 Explicit 2 Explicit and justified	
9	Is the quality of data considered and is uncertainty explored through uncertainty and/or sensitivity analyses?	Are data limitations discussed? Are any of the sources known to the reviewer to be inappropriate (e.g. do not match the parameter, are outdated, or known to be poor quality)?	0 No sources or uncertainty 1 Partially addressed, and/or some data inappropriate 2 Fully addressed	
		Is uncertainty in model structure, parameters and/or assumptions explored through uncertainty and/or sensitivity analyses?		
10	Is the method of fitting described and suitable?	Is the method of fitting/calibrating the model clearly described?	0 Not done, unsuitable method or poor/no description	Fitting/ validation: max 4 points
		Is the method of model fitting/calibration suitable?	1 Incomplete description or method not optimal 2 Complete description and suitable methods	
11	Has the model been validated?	Has an assessment of validity of the results been made by comparing across one or more different model structures, or against a validation data set?	0 Not considered 1 States criteria for validation 2 Validation undertaken	
12	Have the results been clearly and completely	Have the outcome values and their uncertainty ranges for each	0 Not reported, very unclear or not suited to research question	Results: max 4 points

	presented, with a range of uncertainty?	intervention/scenario been reported? Do the results match the objectives? Are sensitivity analyses clearly reported?	1 Stated, but ranges or planned sensitivity analyses missing and/or not directly aligned with research question 2 Values and ranges and planned sensitivity analyses reported and aligned with research question.	
13	Are the results appropriately interpreted and discussed in context?	Does the discussion reflect a fair and balanced interpretation of the results? Are the results of the study discussed in context and is generalisability considered? Are possible biases and limitations discussed?	0 No/poor discussion 1 Some discussion but key points, limitations or context missed 2 Full discussion of key points in context, generalisability considered, limitations discussed	
14	Are the funding source and conflicts of interest reported?	Is the funding and the role of the funder clearly stated? Is there a conflict of interest statement?	0 No statement of funding or conflicts 1 Funding or conflicts reported 2 Funding and conflict statement	Conflicts: Max 2 points
Overall Scoring: Max 28 points				
Very high	>22			
High	19-22			
Medium	14-18			
Low	<14			

Additional results: Model structures and fitting methodology

Model structures are tailored to suit a given research question, but have also evolved based on the growing body of knowledge about TB natural history and epidemiology. The majority of models include the susceptible (S), latent (L), active disease (I) and recovered (R) natural history states. Some models include latent fast and slow progressing states,[6-9] though many models bypass the 'latent fast' state to transition directly to either stable latent infection or active disease.[10,11] Other features included in some models are: treatment status,[6,11] variable infectiousness of active disease or disease type,[12-16] different infective strains,[17] vaccine waning,[8] and stratification for age, HIV,[11,14,18,19] community structure,[20] and vaccination. Markov models identified also included other health states such as re-infected,[21] miliary TB,[21] TB meningitis[21], diagnosis,[18] treatment status,[18,19] and drug resistance status[18,19].

Fitting methods reported included Sobol sequence sampling and approximate Bayesian computation,[14] Downhill simplex method,[6] maximum likelihood,[22] or adjustment of parameters to achieve required equilibrium values using Berkley Madonna "curve fit" routine,[12,13] generalised reduced gradient non linear engine of solver in Excel or manually.[11-13] Fitting methods were unclear or not reported in six of the models.[7-9,15-17] Illustrative examples, Markov or analytical models did not require fitting.[10,18-21,23-27]

Additional results: summary of quality scoring

Study aims and outcome measures tended to be well defined, but there was often missing information with regards to the modelled population, vaccine characteristics, time horizon, on occasion model structure and parameter ranges and references. The model structure and methods were generally suited to the research question, but some models had out-dated natural history structures or inappropriate time horizons. Overall, the scoring for fitting methods was relatively low, as several studies lacked or were unclear in their reporting, plus this is not a relevant criterion for

theoretical papers. Model assumptions were generally stated and justified. Studies were incredibly variable with respect to the completeness of conduct and reporting of uncertainty ranges and sensitivity analyses. Points were lost in reporting of results due to poor reporting of uncertainty and sensitivity analyses, but overall results were appropriately interpreted and discussed in context. Only one model had been validated, and most papers were missing a conflict of interest and/or funding statement.

Appendix B: Supplementary modelling information

B1. Additional results from Chapter 3

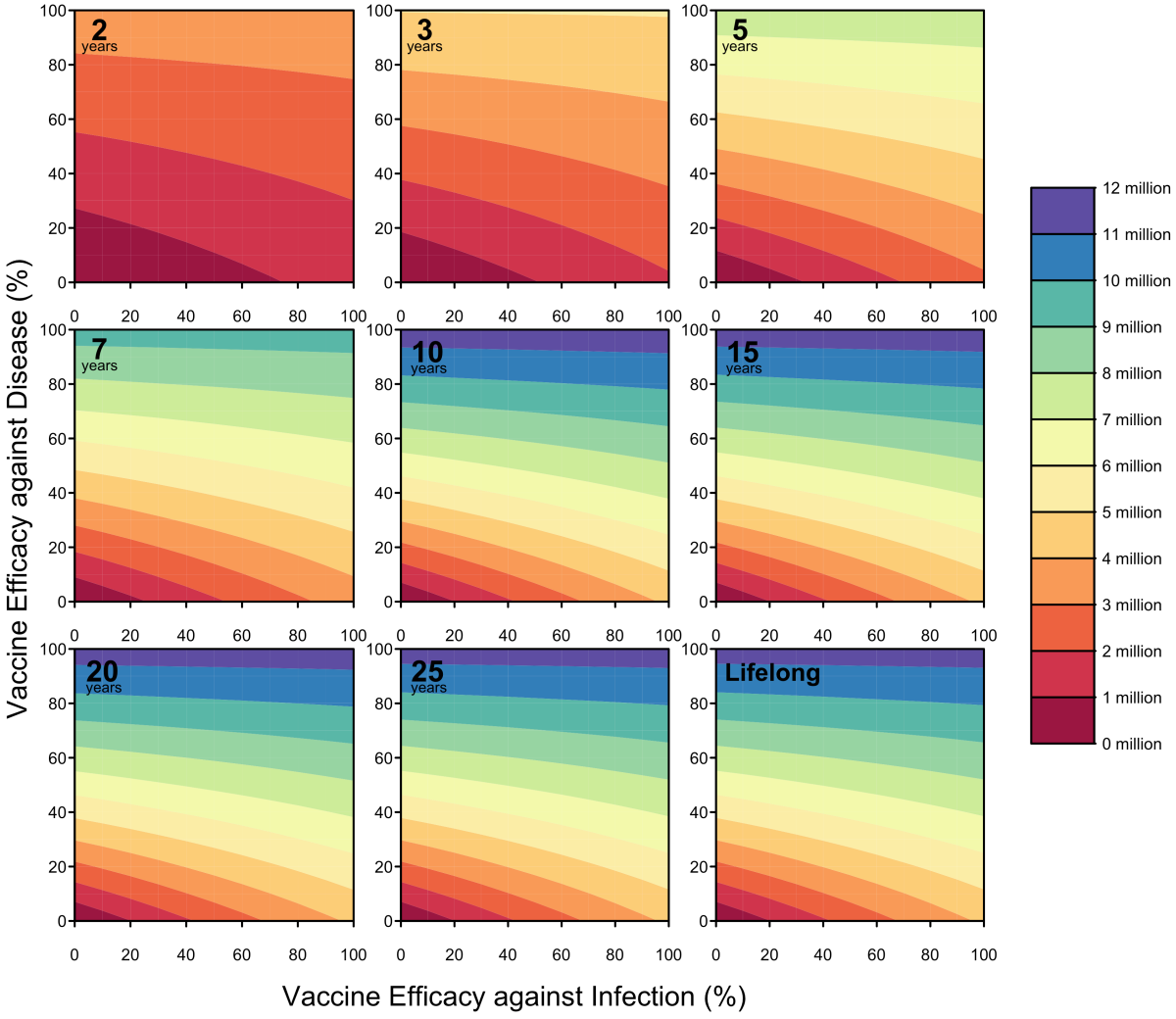


Figure B1: Median cumulative number of cases averted for the period 2025-2050 for pre- and post-infection vaccine compared to no new vaccine baseline, by percentage efficacy against infection (x-axis), percentage efficacy against disease (y-axis), and duration of protection (as indicated in top left corner of panel), all with 10-yearly mass vaccination campaigns.

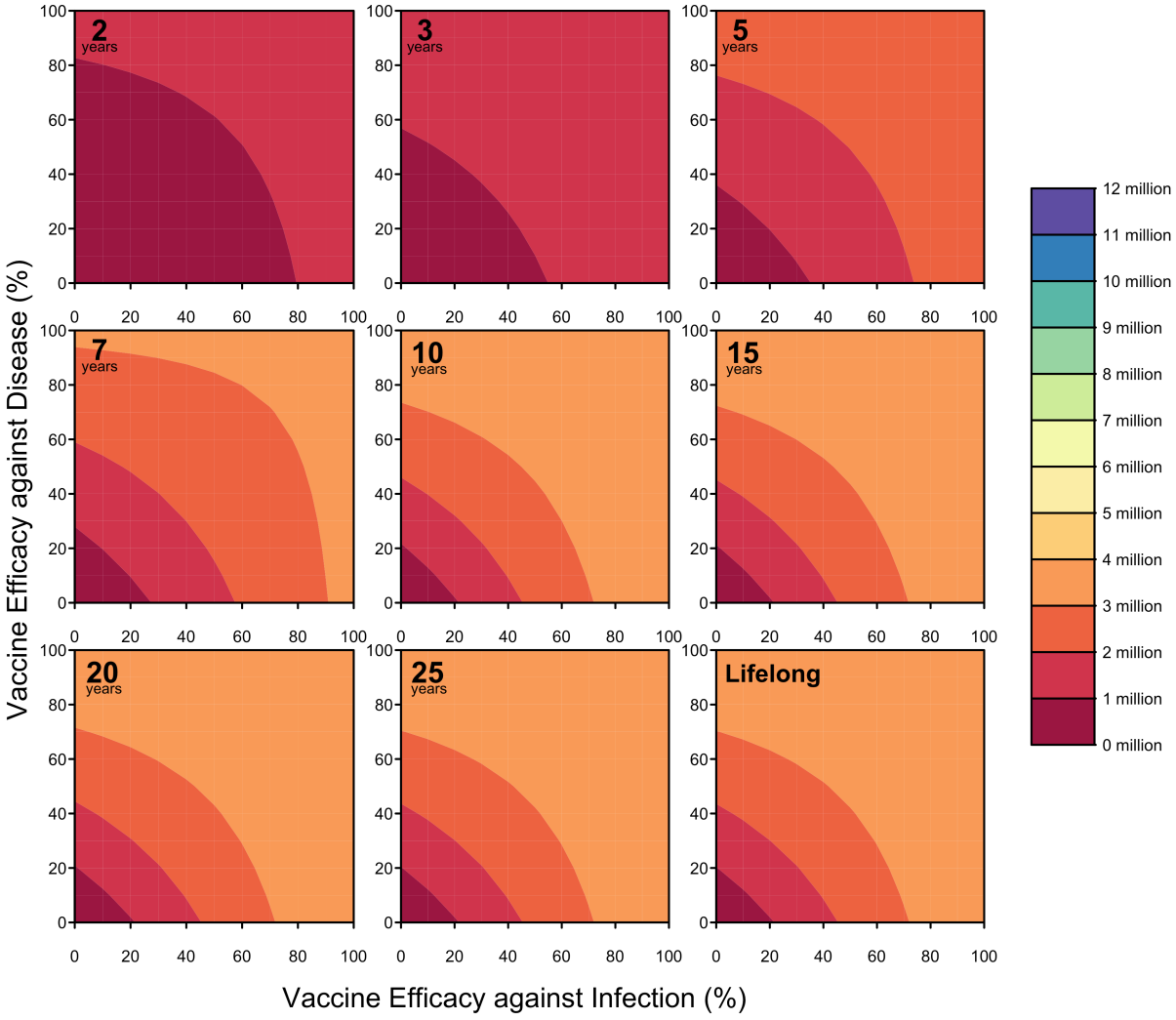


Figure B2: Median cumulative number of cases averted for the period 2025-2050 for pre-infection vaccine compared to no new vaccine baseline, by percentage efficacy against infection (x-axis), percentage efficacy against disease (y-axis), and duration of protection (as indicated in top left corner of panel), all with 10-yearly mass vaccination campaigns.

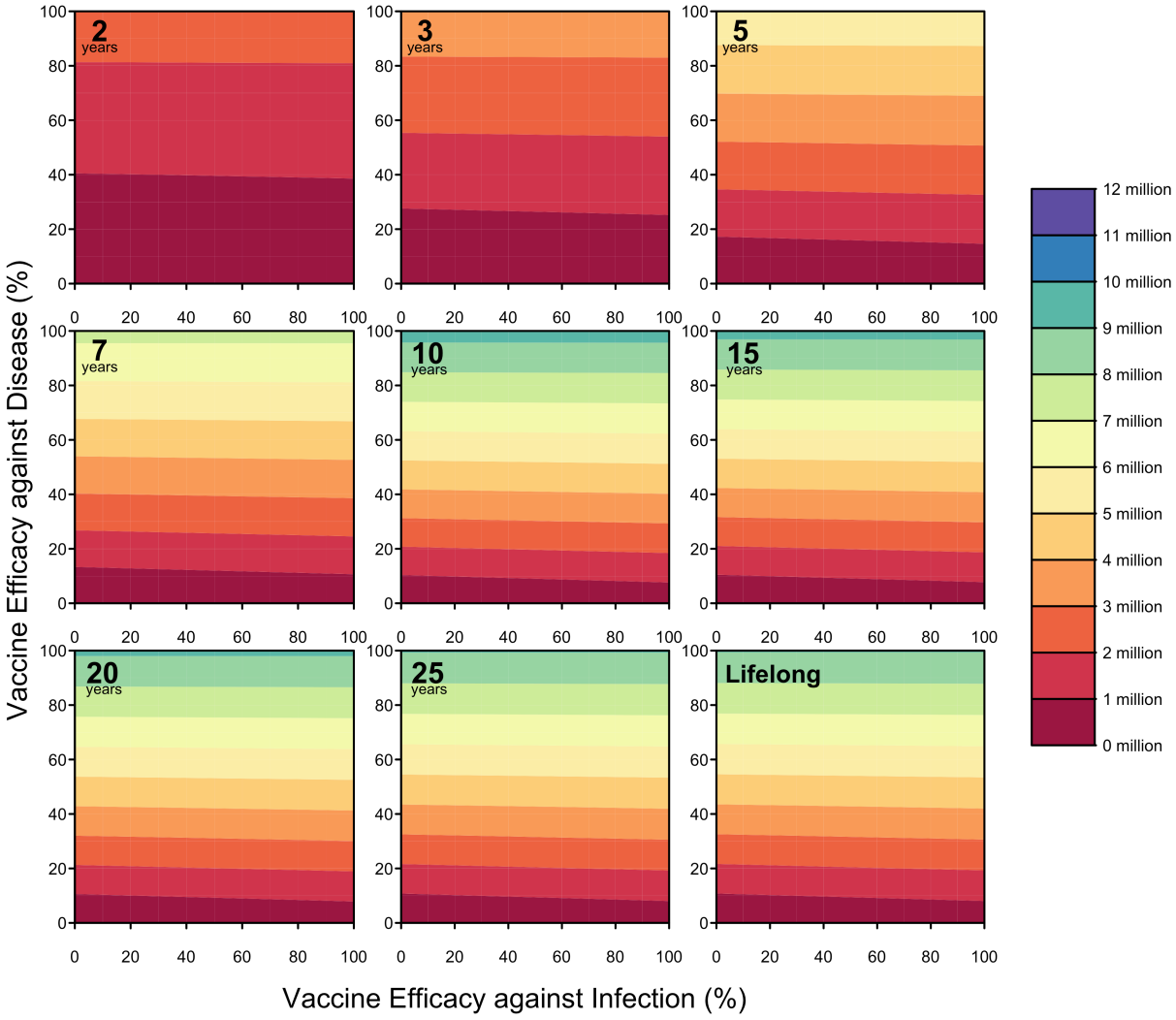


Figure B3: Median cumulative number of cases averted for the period 2025-2050 for post-infection vaccines compared to no new vaccine baseline, by percentage efficacy against infection (x-axis), percentage efficacy against disease (y-axis), and duration of protection (as indicated in top left corner of panel), all with 10-yearly mass vaccination campaigns.

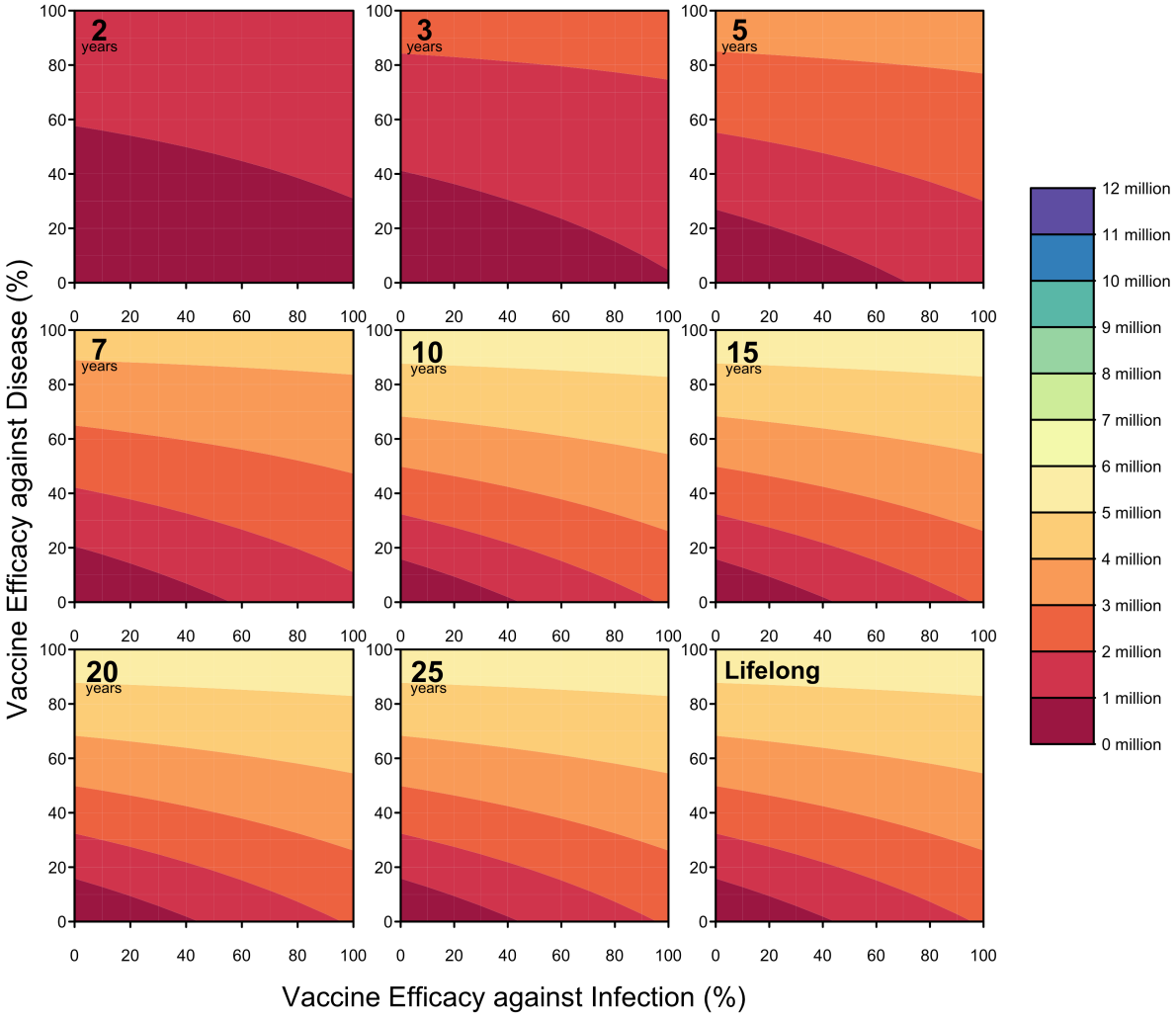
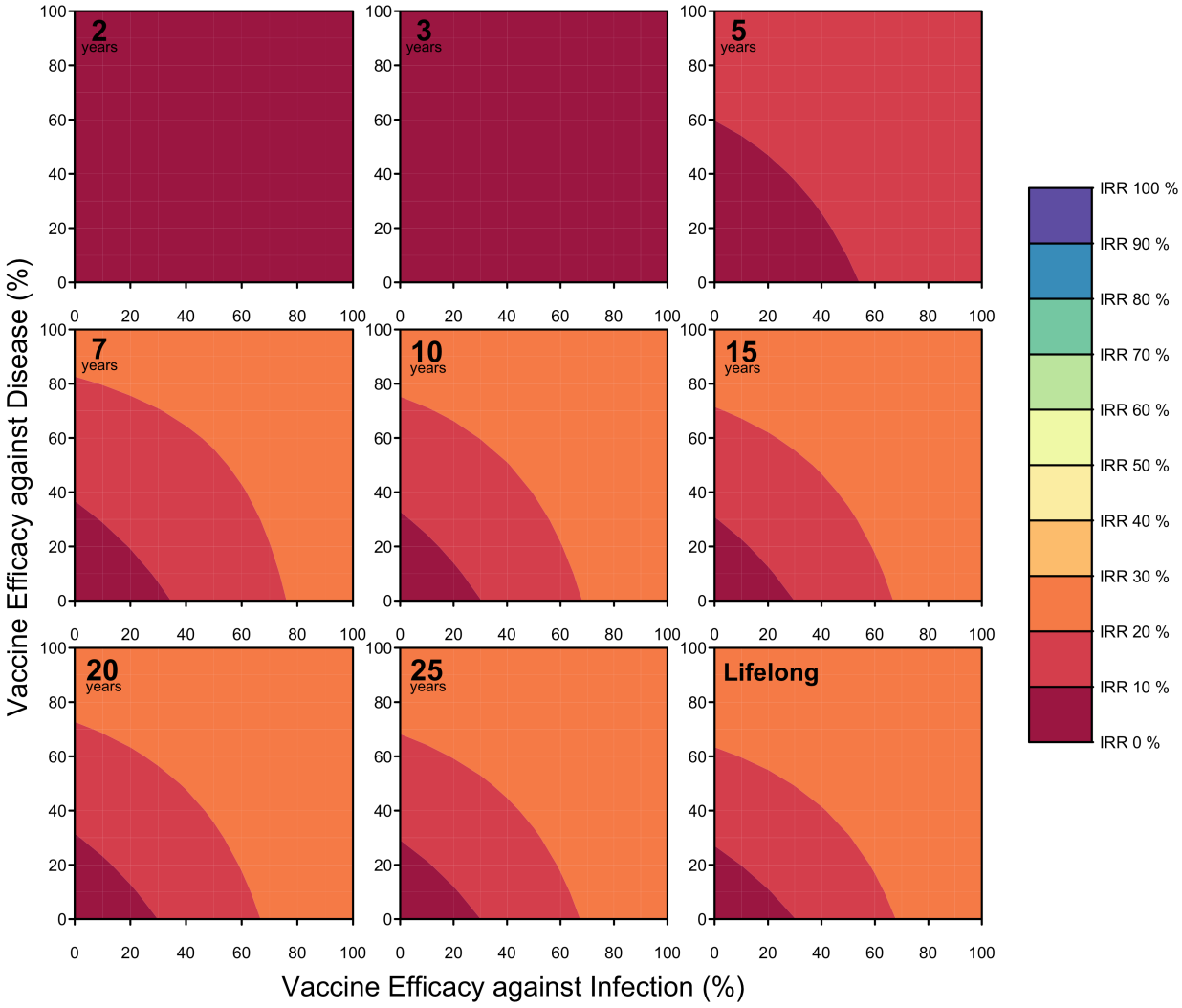
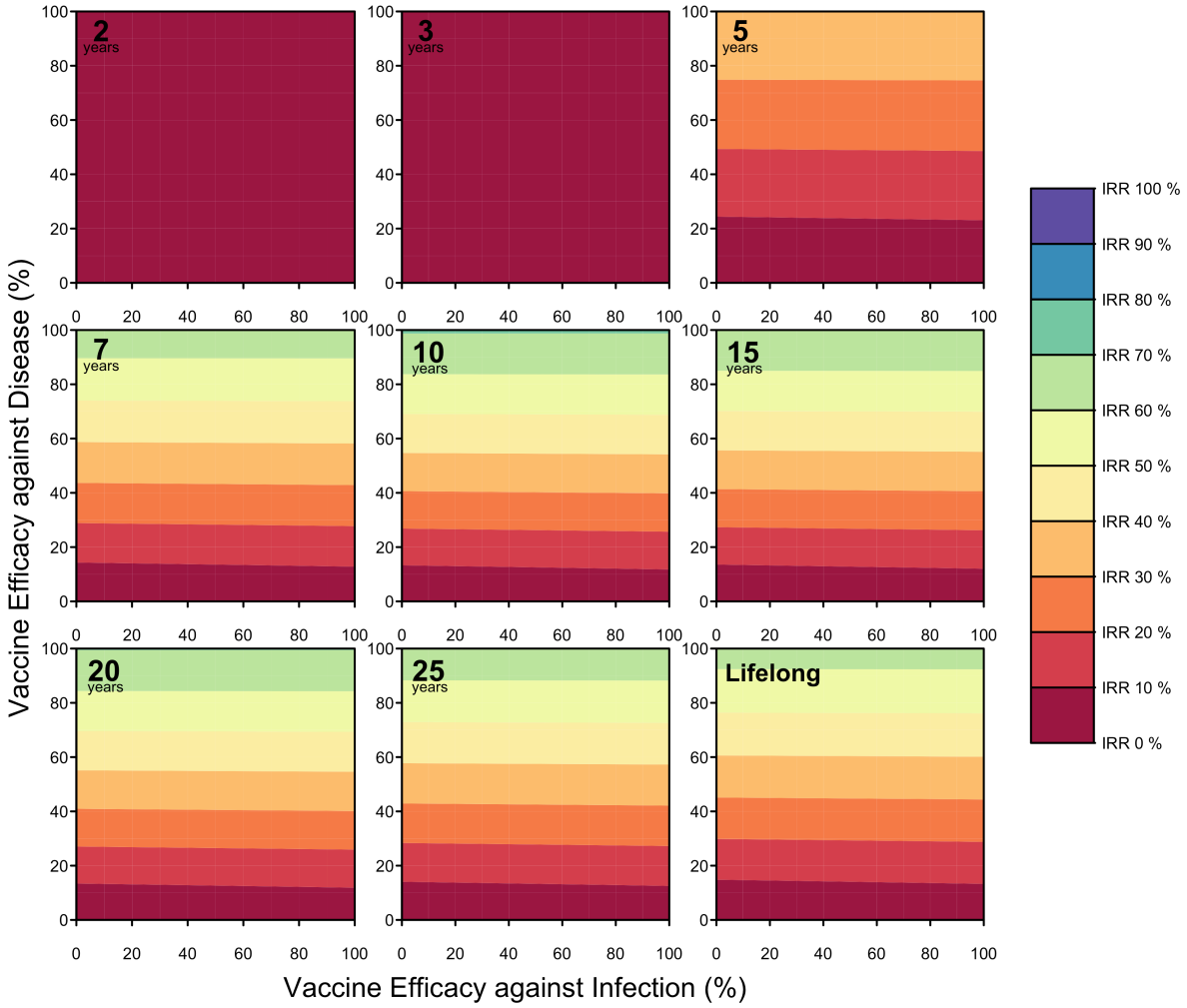


Figure B4: Median cumulative number of cases averted for the period 2025-2035 for pre- and post-infection vaccine compared to no new vaccine baseline, by percentage efficacy against infection (x-axis), percentage efficacy against disease (y-axis), and duration of protection (as indicated in top left corner of panel), all with 10-yearly mass vaccination campaigns.



*Incidence rate reduction (IRR) of 90% is representative of an incidence rate (IR) of 3 (UR: 3-4) per 100,000 population. IRR 50% = IR 17 (UR: 14-19) per 100,000 population. IRR 0% = IR 34 (UR: 28-38) per 100,000 population.

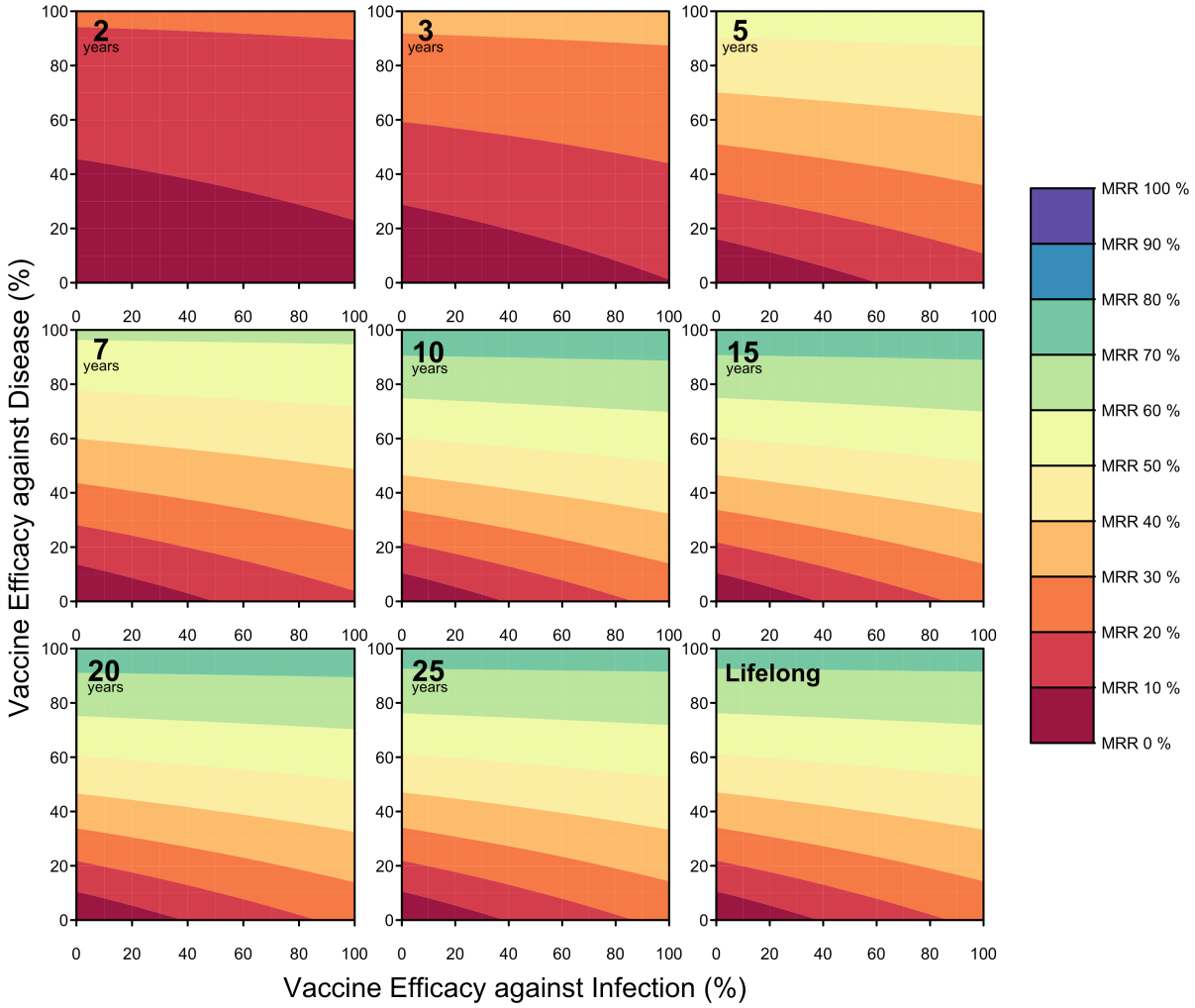
Figure B5: Median incidence rate reduction in 2050 for pre- infection vaccines compared to no new vaccine baseline, by percentage efficacy against infection (x-axis), percentage efficacy against disease (y-axis), and duration of protection (as indicated in top left corner of panel), all with 10-yearly mass vaccination campaigns.



*Incidence rate reduction (IRR) of 90% is representative of an incidence rate (IR) of 3 (UR: 3-4) per 100,000 population. IRR 50% = IR 17 (UR: 14-19) per 100,000 population. IRR 0% = IR 34 (UR: 28-38) per 100,000 population.

Figure B6: Median incidence rate reduction in 2050 for post-infection vaccines compared to no new vaccine baseline, by percentage efficacy against infection (x-axis), percentage efficacy against

disease (y-axis), and duration of protection (as indicated in top left corner of panel), all with 10-yearly mass vaccination campaigns



*Mortality rate reduction (MRR) of 90% is representative of a mortality rate (MR) of 0.11 (UR: 0.06-0.2) per 100,000 population. MRR 50% = MR 17 (UR: 0.29-1.02) per 100,000 population. MRR 0% = MR 1.1 (UR: 0.57-2.03) per 100,000 population.

Figure B7: Median mortality rate reduction (%) for pre- and post-infection vaccine compared to no new vaccine baseline in 2050, by percentage efficacy against infection (x-axis), percentage efficacy against disease (y-axis), and duration of protection (as indicated in top left corner of panel), all with 10-yearly mass vaccination campaigns.

B2. Additional results Chapter 4

Table B1: Cumulative number needed to vaccinate per case averted 2025-2050

Vaccine Mechanism	Age targeting	Coverage (%)	Vaccine Efficacy (%)	Duration	Individuals Vaccinated 2025-2050, Thousands (Uncertainty range)	Cases averted 2025-2050, Thousands (Uncertainty range)	Cumulative NNV per case averted 2025-2050 (Uncertainty range)
PRI	Adolescent	30	40	10	135,917 (135,838-135,980)	75 (64-88)	1822 (1551-2110)
				20	135,917 (135,838-135,980)	113 (97-134)	1206 (1015-1407)
			60	10	135,917 (135,838-135,980)	111 (96-130)	1227 (1044-1422)
				20	135,917 (135,838-135,980)	167 (143-198)	815 (686-950)
			80	10	135,917 (135,838-135,980)	146 (126-172)	930 (791-1077)
				20	135,917 (135,838-135,980)	220 (188-261)	619 (521-722)
		70	40	10	317,139 (316,955-317,287)	169 (146-199)	1873 (1593-2170)
				20	317,139 (316,955-317,287)	254 (218-301)	1249 (1052-1457)
			60	10	317,139 (316,955-317,287)	248 (214-292)	1278 (1087-1481)
				20	317,139 (316,955-317,287)	370 (317-439)	858 (722-1001)
			80	10	317,139 (316,956-317,288)	323 (279-380)	982 (834-1137)
				20	317,139 (316,956-317,288)	479 (410-568)	663 (558-773)

Vaccine Mechanism	Age targeting	Coverage (%)	Vaccine Efficacy (%)	Duration	Individuals Vaccinated 2025-2050, Thousands (Uncertainty range)	Cases averted 2025-2050, Thousands (Uncertainty range)	NNV per case averted (Uncertainty range)
PRI	Older Adult	30	40	10	162,162 (161,417-162,745)	110 (85-151)	1,473 (1,077-1,907)
				20	162,162 (161,417-162,745)	148 (114-207)	1,093 (786-1,425)
			60	10	162,162 (161,417-162,745)	164 (127-224)	990 (725-1,281)
				20	162,162 (161,417-162,745)	220 (169-306)	736 (530-959)
			80	10	162,162 (161,417-162,745)	217 (168-296)	748 (549-967)
				20	162,162 (161,417-162,745)	291 (224-403)	558 (402-725)
		70	40	10	378,378 (376,639-379,739)	251 (195-343)	1,505 (1,104-1,944)
				20	378,378 (376,639-379,739)	337 (260-467)	1,123 (811-1,459)
			60	10	378,378 (376,640-379,740)	370 (287-504)	1,022 (752-1,318)
				20	378,378 (376,640-379,740)	494 (381-681)	766 (556-993)
			80	10	378,379 (376,640-379,740)	484 (377-658)	781 (576-1,005)
				20	378,379 (376,640-379,740)	644 (499-884)	587 (428-760)

Vaccine Mechanism	Age targeting	Coverage (%)	Vaccine Efficacy (%)	Duration	Individuals Vaccinated 2025-2050, Thousands (Uncertainty range)	Cases averted 2025-2050, Thousands (Uncertainty range)	Cumulative NNV per case averted 2025-2050 (Uncertainty range)
PSI-L	Adolescent	30	40	10	135,916 (135,838-135,980)	2 (2-3)	57,905 (42,648-75,818)
				20	135,916 (135,838-135,980)	4 (3-5)	35,941 (26,488-46,864)
			60	10	135,916 (135,838-135,980)	3 (3-5)	38,894 (28,645-50,932)
				20	135,916 (135,838-135,980)	6 (4-8)	24,170 (17,812-31,518)
			80	10	135,916 (135,838-135,980)	5 (4-6)	29,389 (21,644-38,490)
				20	135,916 (135,838-135,980)	7 (6-10)	18,285 (13,475-23,846)
		70	40	10	317,139 (316,955-317,287)	5 (4-7)	59,070 (43,503-77,369)
				20	317,139 (316,955-317,287)	9 (7-12)	36,781 (27,104-47,971)
			60	10	317,139 (316,955-317,287)	8 (6-11)	40,065 (29,505-52,492)
				20	317,139 (316,955-317,287)	13 (10-17)	25,019 (18,434-32,636)
			80	10	317,139 (316,955-317,287)	10 (8-14)	30,568 (22,508-40,058)
				20	317,139 (316,955-317,287)	17 (13-22)	19,142 (14,101-24,974)

Vaccine Mechanism	Age targeting	Coverage (%)	Vaccine Efficacy (%)	Duration	Individuals Vaccinated 2025-2050, Thousands (Uncertainty range)	Cases averted 2025-2050, Thousands (Uncertainty range)	Cumulative NNV per case averted 2025-2050 (Uncertainty range)
PSI-L	Older Adult	30	40	10	162,162 (161,417-162,745)	193 (39-317)	838 (511-4,212)
				20	162,162 (161,417-162,745)	280 (52-466)	579 (348-3,138)
			60	10	162,162 (161,417-162,745)	288 (57-473)	562 (342-2,824)
				20	162,162 (161,417-162,745)	418 (77-694)	389 (233-2,106)
			80	10	162,162 (161,417-162,745)	382 (76-628)	424 (258-2,130)
				20	162,162 (161,417-162,745)	553 (102-920)	293 (176-1,589)
		70	40	10	378,378 (376,641-379,740)	445 (89-730)	851 (518-4,275)
				20	378,378 (376,641-379,740)	643 (119-1,068)	589 (354-3,192)
			60	10	378,378 (376,642-379,740)	658 (131-1,081)	574 (350-2,886)
				20	378,379 (376,642-379,740)	950 (175-1,579)	399 (239-2,160)
			80	10	378,379 (376,643-379,741)	867 (173-1,424)	436 (266-2,192)
				20	378,379 (376,644-379,741)	1247 (230-2075)	304 (182-1,644)

Vaccine Mechanism	Age targeting	Coverage (%)	Vaccine Efficacy (%)	Duration	Individuals Vaccinated 2025-2050, Thousands (Uncertainty range)	Cases averted 2025-2050, Thousands (Uncertainty range)	Cumulative NNV per case averted 2025-2050 (Uncertainty range)
PSI-LR	Adolescent	30	40	10	135,916 (135,838-135,980)	4 (3-5)	38,771 (29,534-50,337)
				20	135,916 (135,838-135,980)	6 (4-7)	23,774 (18,156-30,782)
			60	10	135,917 (135,838-135,980)	5 (4-7)	26,042 (19,838-33,813)
				20	135,917 (135,838-135,980)	9 (7-11)	15,987 (12,210-20,701)
			80	10	135,917 (135,838-135,980)	7 (5-9)	19,679 (14,990-25,551)
				20	135,917 (135,838-135,980)	11 (9-15)	12,094 (9,237-15,661)
		70	40	10	317,139 (316,955-317,287)	8 (6-11)	39,554 (30,129-51,359)
				20	317,139 (316,955-317,287)	13 (10-17)	24,327 (18,581-31,503)
			60	10	317,139 (316,955-317,287)	12 (9-16)	26,831 (20,437-34,840)
				20	317,139 (316,955-317,287)	19 (15-25)	16,547 (12,638-21,429)
			80	10	317,139 (316,955-317,287)	15 (12-20)	20,471 (15,592-26,585)
				20	317,139 (316,955-317,287)	25 (19-33)	12,660 (9,669-16,396)

Vaccine Mechanism	Age targeting	Coverage (%)	Vaccine Efficacy (%)	Duration	Individuals Vaccinated 2025-2050, Thousands (Uncertainty range)	Cases averted 2025-2050, Thousands (Uncertainty range)	Cumulative NNV per case averted 2025-2050 (Uncertainty range)
PSI-LR	Older Adult	30	40	10	162,162 (161,417-162,746)	380 (305-431)	426 (375-532)
				20	162,162 (161,417-162,746)	547 (432-625)	297 (259-376)
			60	10	162,162 (161,417-162,746)	567 (455-643)	286 (252-357)
				20	162,162 (161,418-162,746)	815 (643-931)	199 (174-252)
			80	10	162,163 (161,418-162,746)	752 (603-853)	216 (190-269)
				20	162,163 (161,418-162,746)	1,079 (852-1,233)	150 (131-190)
		70	40	10	378,380 (376,642-379,741)	874 (700-992)	433 (381-541)
				20	378,380 (376,642-379,741)	1,254 (990-1,432)	302 (264-383)
			60	10	378,381 (376,644-379,743)	1,295 (1,037-1,469)	292 (257-365)
				20	378,381 (376,644-379,743)	1,852 (1,461-2,117)	204 (179-259)
			80	10	378,382 (376,645-379,744)	1,704 (1,364-1,935)	222 (195-278)
				20	378,383 (376,646-379,744)	2,432 (1,919-2,781)	156 (136-197)

Vaccine Mechanism	Age targeting	Coverage (%)	Vaccine Efficacy (%)	Duration	Individuals Vaccinated 2025-2050, Thousands (Uncertainty range)	Cases averted 2025-2050, Thousands (Uncertainty range)	Cumulative NNV per case averted 2025-2050 (uncertainty range)
PPI	Adolescent	30	40	10	135,917 (135,838-135,980)	78 (68-92)	1,743 (1,485-2,013)
				20	135,917 (135,838-135,980)	118 (102-139)	1,149 (976-1,339)
			60	10	135,917 (135,838-135,980)	116 (100-136)	1,174 (1,001-1,356)
				20	135,917 (135,838-135,980)	175 (150-206)	776 (659-905)
			80	10	135,917 (135,838-135,980)	153 (132-179)	890 (758-1,028)
				20	135,917 (135,838-135,980)	230 (198-271)	590 (501-688)
		70	40	10	317,139 (316,955-317,287)	177 (153-208)	1,792 (1,527-2,070)
				20	317,139 (316,955-317,287)	266 (229-314)	1,190 (1,011-1,388)
			60	10	317,139 (316,955-317,287)	259 (224-304)	1,223 (1,043-1,414)
				20	317,139 (316,955-317,287)	388 (333-456)	818 (695-954)
			80	10	317,139 (316,956-317,288)	338 (292-396)	939 (801-1,086)
				20	317,139 (316,956-317,288)	502 (431-591)	632 (537-737)

Vaccine Mechanism	Age targeting	Coverage (%)	Vaccine Efficacy (%)	Duration	Individuals Vaccinated 2025-2050, Thousands (Uncertainty range)	Cases averted 2025-2050, Thousands (Uncertainty range)	Cumulative NNV per case averted 2025-2050 (uncertainty range)
PPI	Older Adult	30	40	10	162,162 (161,417-162,746)	489 (418-566)	332 (286-387)
				20	162,162 (161,417-162,746)	693 (585-807)	234 (200-277)
			60	10	162,163 (161,418-162,746)	727 (622-841)	223 (192-261)
				20	162,163 (161,418-162,746)	1,028 (868-1,196)	158 (135-187)
			80	10	162,163 (161,418-162,746)	962 (822-1,111)	169 (146-197)
				20	162,163 (161,418-162,746)	1,355 (1,145-1,576)	120 (103-141)
		70	40	10	378,381 (376,643-379,742)	1,116 (954-1,289)	339 (293-396)
				20	378,381 (376,643-379,742)	1,571 (1,327-1,824)	241 (207-285)
			60	10	378,382 (376,645-379,744)	1,643 (1,403-1,893)	230 (199-269)
				20	378,383 (376,645-379,744)	2,300 (1,944-2,663)	164 (142-194)
			80	10	378,384 (376,647-379,745)	2,150 (1,837-2,473)	176 (153-206)
				20	378,384 (376,648-379,746)	2,995 (2,532-3,459)	126 (109-149)

Table B2: Cumulative number needed to vaccinate per death averted 2025-2050

Vaccine Mechanism	Age targeting	Coverage (%)	Vaccine Efficacy (%)	Duration	Individuals Vaccinated 2025-2050, Thousands (Uncertainty range)	Death averted 2025-2050, Thousands (Uncertainty range)	NNV per death averted (Uncertainty range)
PRI	Adolescent	30	40	10	135,917 (135,838-135,980)	0.85 (0.45-1.40)	140,995 (108,261-201,462)
				20	135,917 (135,838-135,980)	1.18 (0.62-1.97)	102,019 (77,815-145,567)
			60	10	135,917 (135,838-135,980)	1.26 (0.67-2.08)	95,446 (73,282-136,383)
				20	135,917 (135,838-135,980)	1.74 (0.92-2.91)	69,254 (52,823-98,818)
			80	10	135,917 (135,838-135,980)	1.66 (0.89-2.74)	72,682 (55,800-103,857)
				20	135,917 (135,838-135,980)	2.29 (1.21-3.83)	52,880 (40,334-75,456)
		70	40	10	317,139 (316,955-317,287)	1.92 (1.03-3.17)	146,838 (112,726-209,823)
				20	317,139 (316,955-317,287)	2.65 (1.39-4.43)	107,023 (81,630-152,716)
			60	10	317,139 (316,955-317,287)	2.81 (1.51-4.64)	101,379 (77,813-144,867)
				20	317,139 (316,955-317,287)	3.85 (2.03-6.44)	74,340 (56,700-106,081)
			80	10	317,139 (316,956-317,288)	3.65 (1.96-6.04)	78,703 (60,397-112,465)
				20	317,139 (316,956-317,288)	4.98 (2.62-8.34)	58,048 (44,272-82,833)

Vaccine Mechanism	Age targeting	Coverage (%)	Vaccine Efficacy (%)	Duration	Individuals Vaccinated 2025-2050, Thousands (Uncertainty range)	Deaths averted 2025-2050, Thousands (Uncertainty range)	Cumulative NNV per death averted 2025-2050 (Uncertainty range)
PRI	Older Adult	30	40	10	162,162 (161,417-162,745)	2.79 (1.40-6.41)	64,045 (52,494-74,077)
				20	162,162 (161,417-162,745)	3.66 (1.76-8.71)	50,007 (40,335-56,999)
			60	10	162,162 (161,417-162,745)	4.14 (2.08-9.51)	43,078 (35,326-49,839)
				20	162,162 (161,417-162,745)	5.43 (2.61-12.90)	33,721 (27,189-38,416)
			80	10	162,162 (161,417-162,745)	5.47 (2.75-12.54)	32,596 (26,742-37,721)
				20	162,162 (161,417-162,745)	7.15 (3.44-16.98)	25,580 (20,617-29,126)
		70	40	10	378,378 (376,639-379,739)	6.34 (3.19-14.53)	65,575 (53,816-75,898)
				20	378,378 (376,639-379,739)	8.29 (3.98-19.64)	51,546 (41,536-58,671)
			60	10	378,378 (376,640-379,740)	9.31 (4.69-21.27)	44,613 (36,654-51,666)
				20	378,378 (376,640-379,740)	12.13 (5.83-28.61)	35,270 (28,397-40,097)
			80	10	378,379 (376,640-379,740)	12.15 (6.13-27.69)	34,136 (28,076-39,554)
				20	378,379 (376,640-379,740)	15.78 (7.59-37.07)	27,132 (21,832-30,816)

Vaccine Mechanism	Age targeting	Coverage (%)	Vaccine Efficacy (%)	Duration	Individuals Vaccinated 2025-2050, Thousands (Uncertainty range)	Deaths averted 2025-2050, Thousands (Uncertainty range)	NNV per deaths averted (Uncertainty range)
PSI-L	Adolescent	30	40	10	135,916 (135,838-135,980)	0.03 (0.01-0.06)	3,706,285 (2,958,178-6,096,566)
				20	135,916 (135,838-135,980)	0.04 (0.02-0.08)	2,462,677 (1,965,549-4,056,801)
			60	10	135,916 (135,838-135,980)	0.04 (0.02-0.09)	2,500,691 (1,996,083-4,113,801)
				20	135,916 (135,838-135,980)	0.06 (0.03-0.13)	1,663,455 (1,327,628-2,740,425)
			80	10	135,916 (135,838-135,980)	0.05 (0.03-0.11)	1,898,161 (1,515,252-3,122,865)
				20	135,916 (135,838-135,980)	0.08 (0.04-0.17)	1,264,068 (1,008,846-2,082,613)
		70	40	10	317,139 (316,955-317,287)	0.06 (0.03-0.13)	3,826,806 (3,054,996-6,296,245)
				20	317,139 (316,955-317,287)	0.09 (0.05-0.19)	2,550,356 (2,035,391-4,202,038)
			60	10	317,139 (316,955-317,287)	0.09 (0.05-0.20)	2,623,571 (2,094,819-4,317,445)
				20	317,139 (316,955-317,287)	0.13 (0.07-0.28)	1,753,138 (1,399,062-2,889,019)
			80	10	317,139 (316,955-317,287)	0.12 (0.06-0.26)	2,023,385 (1,615,897-3,330,456)
				20	317,139 (316,955-317,287)	0.17 (0.09-0.37)	1,355,760 (1,081,876-2,234,576)

Vaccine Mechanism	Age targeting	Coverage (%)	Vaccine Efficacy (%)	Duration	Individuals Vaccinated 2025-2050, Thousands (Uncertainty range)	Deaths averted 2025-2050, Thousands (Uncertainty range)	Cumulative NNV per death averted 2025-2050 (Uncertainty range)
PSI-L	Older Adult	30	40	10	162,162 (161,417-162,745)	4.60 (0.65-13.31)	45,342 (37,100-79,890)
				20	162,162 (161,417-162,745)	6.34 (0.85-19.17)	32,516 (27,191-59,563)
			60	10	162,162 (161,417-162,745)	6.85 (0.97-19.83)	30,446 (24,908-53,646)
				20	162,162 (161,417-162,745)	9.44 (1.27-28.51)	21,857 (18,274-40,037)
			80	10	162,162 (161,417-162,745)	9.08 (1.28-26.25)	22,999 (18,812-40,524)
				20	162,162 (161,417-162,745)	12.48 (1.68-37.71)	16,528 (13,815-30,275)
		70	40	10	378,378 (376,641-379,740)	10.54 (1.49-30.48)	46,217 (37,799-81,436)
				20	378,378 (376,641-379,740)	14.49 (1.95-43.76)	33,237 (27,778-60,881)
			60	10	378,378 (376,642-379,740)	15.55 (2.19-44.98)	31,324 (25,609-55,197)
				20	378,379 (376,642-379,740)	21.33 (2.87-64.41)	22,582 (18,864-41,362)
			80	10	378,379 (376,643-379,741)	20.41 (2.88-59.01)	23,879 (19,515-42,079)
				20	378,379 (376,644-379,741)	27.92 (3.76-84.29)	17,257 (14,408-31,607)

Vaccine Mechanism	Age targeting	Coverage (%)	Vaccine Efficacy (%)	Duration	Individuals Vaccinated 2025-2050, Thousands (Uncertainty range)	Deaths averted 2025-2050, Thousands (Uncertainty range)	Cumulative NNV per death averted 2025-2050 (Uncertainty range)
PSI-LR	Adolescent	30	40	10	135,916 (135,838-135,980)	0.04 (0.02-0.09)	1,953,107 (1,814,214-3,607,241)
				20	135,916 (135,838-135,980)	0.06 (0.03-0.13)	1,264,352 (1,173,728-2,346,808)
			60	10	135,917 (135,838-135,980)	0.06 (0.03-0.13)	1,317,972 (1,224,238-2,434,448)
				20	135,917 (135,838-135,980)	0.09 (0.05-0.19)	854,077 (792,859-1,585,394)
			80	10	135,917 (135,838-135,980)	0.08 (0.04-0.17)	1,000,553 (929,387-1,848,334)
				20	135,917 (135,838-135,980)	0.12 (0.06-0.25)	649,058 (602,536-1,204,912)
		70	40	10	317,139 (316,955-317,287)	0.09 (0.05-0.20)	2,017,371 (1,873,873-3,726,978)
				20	317,139 (316,955-317,287)	0.14 (0.07-0.29)	1,309,585 (1,215,717-2,431,232)
			60	10	317,139 (316,955-317,287)	0.14 (0.07-0.29)	1,383,557 (1,285,122-2,556,696)
				20	317,139 (316,955-317,287)	0.20 (0.10-0.42)	900,378 (835,840-1,671,837)
			80	10	317,139 (316,955-317,287)	0.18 (0.09-0.376)	1,067,459 (991,497-1,973,094)
				20	317,139 (316,955-317,287)	0.26 (0.13-0.55)	696,435 (646,515-1,293,391)

Vaccine Mechanism	Age targeting	Coverage (%)	Vaccine Efficacy (%)	Duration	Individuals Vaccinated 2025-2050, Thousands (Uncertainty range)	Deaths averted 2025-2050, Thousands (Uncertainty range)	Cumulative NNV per death averted 2025-2050 (Uncertainty range)
PSI-LR	Older Adult	30	40	10	162,162 (161,417-162,746)	9.83 (4.83-19.37)	19,036 (14,653-21,122)
				20	162,162 (161,417-162,746)	13.70 (6.59-27.68)	13,582 (10,382-15,247)
			60	10	162,162 (161,417-162,746)	14.64 (7.20-28.85)	12,785 (9,841-14,183)
				20	162,162 (161,418-162,746)	20.38 (9.80-41.18)	9,131 (6,980-10,249)
			80	10	162,163 (161,418-162,746)	19.38 (9.53-38.19)	9,659 (7,435-10,713)
				20	162,163 (161,418-162,746)	26.95 (12.96-54.45)	6,906 (5,279-7,750)
		70	40	10	378,380 (376,642-379,741)	22.50 (11.06-44.34)	19,413 (14,943-21,528)
				20	378,380 (376,642-379,741)	31.27 (15.03-63.19)	13,890 (10,617-15,585)
			60	10	378,381 (376,644-379,743)	33.21 (16.32-65.42)	13,164 (10,132-14,590)
				20	378,381 (376,644-379,743)	46.03 (22.11-93.00)	9,441 (7,216-10,589)
			80	10	378,382 (376,645-379,744)	43.57 (21.40-85.83)	10,040 (7,727-11,122)
				20	378,383 (376,646-379,744)	60.22 (28.93-121.71)	7,218 (5,517-8,092)

Vaccine Mechanism	Age targeting	Coverage (%)	Vaccine Efficacy (%)	Duration	Individuals Vaccinated 2025-2050, Thousands (Uncertainty range)	Deaths averted 2025-2050, Thousands (Uncertainty range)	Cumulative NNV per death averted 2025-2050 (Uncertainty range)
PPI	Adolescent	30	40	10	135,917 (135,838-135,980)	0.89 (0.48-1.47)	131,098 (102,232-190,927)
				20	135,917 (135,838-135,980)	1.23 (0.65-2.05)	94,158 (73,049-137,194)
			60	10	135,917 (135,838-135,980)	1.31 (0.71-2.18)	88,762 (69,211-129,270)
				20	135,917 (135,838-135,980)	1.83 (0.97-3.04)	63,931 (49,596-93,150)
			80	10	135,917 (135,838-135,980)	1.73 (0.934-2.88)	67,604 (52,708-98,455)
				20	135,917 (135,838-135,980)	2.40 (1.27-3.99)	48,826 (37,877-71,140)
		70	40	10	317,139 (316,955-317,287)	2.00 (1.08-3.33)	136,595 (106,491-198,928)
				20	317,139 (316,955-317,287)	2.78 (1.47-4.62)	98,833 (76,667-143,998)
			60	10	317,139 (316,955-317,287)	2.93 (1.58-4.87)	94,346 (73,535-137,392)
				20	317,139 (316,955-317,287)	4.04 (2.14-6.72)	68,685 (53,275-100,065)
			80	10	317,139 (316,956-317,288)	3.81 (2.06-6.34)	73,274 (57,096-106,698)
				20	317,139 (316,956-317,288)	5.22 (2.76-8.69)	53,659 (41,615-78,168)

Vaccine Mechanism	Age targeting	Coverage (%)	Vaccine Efficacy (%)	Duration	Individuals Vaccinated 2025-2050, Thousands (Uncertainty range)	Deaths averted 2025-2050, Thousands (Uncertainty range)	Cumulative NNV per death averted 2025-2050 (Uncertainty range)
PPI	Older Adult	30	40	10	162,162 (161,417-162,746)	12.57 (6.56-24.13)	14,698 (11,502-16,502)
				20	162,162 (161,417-162,746)	17.17 (8.74-34.06)	10,671 (8,303-12,097)
			60	10	162,163 (161,418-162,746)	18.68 (9.75-35.85)	9,897 (7,743-11,107)
				20	162,163 (161,418-162,746)	25.47 (12.97-50.51)	7,199 (5,600-8,159)
			80	10	162,163 (161,418-162,746)	24.66 (12.88-47.35)	7,496 (5,864-8,410)
				20	162,163 (161,418-162,746)	33.58 (17.10-66.59)	5,464 (4,249-6,190)
		70	40	10	378,381 (376,643-379,742)	28.59 (14.93-54.90)	15,091 (11,804-16,927)
				20	378,381 (376,643-379,742)	38.89 (19.81-77.10)	11,013 (8,564-12,474)
			60	10	378,382 (376,645-379,744)	41.94 (21.92-80.58)	10,292 (8,048-11,534)
				20	378,383 (376,645-379,744)	56.80 (28.98-112.67)	7,544 (5,863-8,538)
			80	10	378,384 (376,647-379,745)	54.71 (28.63-105.18)	7,893 (6,170-8,838)
				20	378,384 (376,648-379,746)	73.73 (37.70-146.41)	5,810 (4,514-6,571)

B3. Additional methods chapter 4

Immunosenescence and vaccination

As discussed briefly in the Chapter 4, elderly populations are at increased risk of infection and disease. This is thought to be likely due to a combination of immunosenescence and comorbid conditions reducing the immune system's ability to manage infections.[1] Age-associated immunosenescence has been shown to affect both the innate (e.g. reduced phagocytic activity) and the adaptive (e.g. diminished antibody response to new antigens and reduced T-cell output) immune response.[1] Therefore, vaccines may be less efficacious in this age group, or need to be specifically designed or dosed to overcome known reduced responsiveness in this population.[2]

Live attenuated vaccines are generally thought to induce more effective immune responses, but in some cases have been associated with higher risk of adverse events, which is of concern for older age groups.[1] For example, the live attenuated yellow fever vaccine has been associated with up to 10 times the risk of severe adverse events in the elderly compared to younger populations.[1]

Few vaccines have been rigorously assessed for efficacy in elderly populations. Influenza and varicella vaccines are two vaccines currently delivered to older age groups. As discussed in the main text, one of the varicella vaccines (Zostavax®) was found to have declining efficacy by age in the elderly,[3] suggestive of an impact of immunosenescence on vaccine efficacy in the 70-79 and ≥80 years age groups. In the elderly, a dose approximately 14 times the infant does was delivered. However, efficacy of a different zoster vaccine was found not to decline with age in the elderly.[4] Therefore, effects of immunosenescence can potentially be counteracted if the vaccine is appropriately designed or dosed.[1]

Another consideration in the elderly is the existence of co-morbidities, which may affect vaccine efficacy, and for some comorbidities may lead to contraindication due to safety concerns.[1] Therefore comorbidities could affect coverage and efficacy.

Vaccination of the elderly clearly comes with safety and efficacy challenges. Some of which can potentially be overcome with careful design of the vaccine and dosing, but for some candidates the effects of immunosenescence may be unavoidable. In this model I implemented immunosenescence in the elderly, as data from other vaccines are suggestive of a likely reduction in protective efficacy in older age groups. I assumed 2% immunosenescent waning per year from the age of 65 years based upon expert opinion. However, as there are few data to inform this parameter, I also conducted a sensitivity analysis varying this waning from 0% to 5%. The 0% scenario was in line with data from the Zoster vaccine that demonstrated invariate efficacy by age.[4] Data from the Zostavax® vaccine showing declining efficacy by age was used to inform the upper end of the sensitivity analysis (5%). This vaccine demonstrated 64% efficacy in 60-69 year olds, 41% efficacy 70-79 years, and 18% efficacy in ≥80 year olds.[3] As a back-of-the-envelope estimation, efficacy would need to start at 68% in 60-64 year olds, and decline by approximately 5% per year from the age of 65 years to produce these results. This is a very approximate estimation, but in the absence of other data sources to inform this analysis, provides a quantitative basis for the upper range of the sensitivity analysis for the immunosenescence parameter.

These considerations should be accounted for in the development of new TB vaccines – older adults and the elderly need to be included in clinical trials for evaluation of safety and efficacy to inform vaccine use in this age group, appropriate modelling assumptions, and consideration of whether alterations to the vaccine are needed to improve efficacy in this age group.

B4. Additional results for Chapter 4

Reduction in 2050 TB Incidence rate

Although not the primary aim of this research, some possible trade-offs between vaccine characteristics were identified. Results suggest that with a 40% VE, 10 year protection and 30% coverage vaccine as the starting point, roughly double the impact could be achieved by increasing duration of protection from 10 to 20 years, or vaccine efficacy from 40% to 80%, or coverage from 30% to 70%.

Many vaccines (excepting PSI-LR and some 20 year duration vaccines) demonstrated similar median estimates and overlapping URs when comparing low efficacy vaccines with high coverage (40% VE, 70% coverage) to high efficacy vaccines with low coverage (80% VE, 30% coverage), suggesting that shortfalls in vaccine efficacy could often be compensated for by increased coverage, and vice versa.

Therefore, even if an ideal candidate is not developed, investment in achieving high coverage rates and ensuring the highest burden populations are vaccinated, could potentially overcome shortfalls in intrinsic vaccine characteristics.

Incidence rate reduction ranged 0.01% to 32% with feasible vaccine profiles and targeted vaccination strategies explored. Although the overall impact could be improved by inclusion of broader mass campaigns, with these relatively feasible vaccination scenarios the impact difference between the most and least effective vaccines and delivery to older adults versus adolescents was substantial, highlighting the importance of ensuring that an appropriately designed vaccine is developed and delivered to a suitable population to maximise impact.

Low vaccine efficacy (40% instead of 80%) could be compensated for by increasing coverage from 30% to 70%.

B5: References

1. Amanna IJ. Balancing the Efficacy and Safety of Vaccines in the Elderly. *Open Longevity Science* 2012; **6**(2012): 64-72.
2. Aspinall R, Del Giudice G, Effros RB, Grubeck-Loebenstein B, Sambhara S. Challenges for vaccination in the elderly. *Immunity & ageing : I & A* 2007; **4**: 9.
3. Merck. Zostavax Package Insert. Whitehouse Station: Merck; 2009. p. 9.
4. Cunningham AL, Lal H, Kovac M, et al. Efficacy of the Herpes Zoster Subunit Vaccine in Adults 70 Years of Age or Older. *New England Journal of Medicine* 2016; **375**(11): 1019-32.

Appendix C: ePAL Protocol

Clinic-based electronic capture of the spatial distribution of TB case place of residence in Blantyre, Malawi (The TB ePAL Study)

Principal Investigator:

London School of Hygiene and Tropical Medicine (LSHTM)

Rebecca Harris

Host Institute:

Malawi-Liverpool-Wellcome Clinical Research Programme

Local collaborators:

MLW

Professor Elizabeth Corbett

Augustine Choko

Rodrick Sambakunsi

David Matiya

Alfred Chimala

Professor Rob Heyderman

District Health Office:

Dr Owen Malemiya

Ministry of Health, National TB Programme & HIV unit:

NTP Manager: Dr James Mpunga

TB Research: Mr Andrew Dimba

International Collaborators:

Dr Richard White **LSHTM, UK**

Dr Peter MacPherson **LSTM, UK**

Dr Emily Webb **LSHTM, UK**

Version 2.0 last amended 2/6/2014 by RCH

Abbreviations

ACF	Active case-finding for TB
ARV	Anti-retroviral
ART	Anti-retroviral therapy
CEA	Census enumeration area
eCRF	electronic Case Report Form
EDC	Electronic Data Capture
ePAL	electronic PArticipant Locator
GPS	Global Positioning System
LSHTM	London School of Hygiene and Tropical Medicine
MLW	Malawi-Liverpool-Wellcome Trust
HSA	Health Surveillance Assistants (Community Health Workers)
PI	Principal Investigator
PII	Personally Identifiable Information
POI	Point/Place(s) Of Interest
SES	Socio-Economic Status

Glossary

Adult	**Person aged ≥18 years old at time of TB treatment initiation visit**
ePAL (electronic Participant Locator)	Software loaded onto electronic tablets for use at the health centres to collect, save and transfer participant information and place of residence to a central database. Contains data capture forms and annotated maps of Blantyre City.
Evaluation period	Approximately 6 month period following the run-in period (below) during which approximately 196 consenting participants will receive follow-up visits at their place of residence to collect GPS coordinates for evaluation of the ePAL tool.
Participant	Person eligible, consented and enrolled into the study.
TB patient	Person receiving care under the diagnosis of tuberculosis (regardless of method of diagnosis).
Place of Residence	The place where a person usually eats meals and sleeps
Patient initiating TB treatment or TB treatment initiator	TB patient presenting at health centre registering for initiation of treatment for TB.
Run-in period	Period at the beginning of the study when the tool is rolled out and data collected from evaluation visits is used to inform the improvement of the tool and its implementation.
Treatment initiation visit	Visit to the health centre as part of routine care at which people are registered as TB patients and commence treatment for TB disease.
Follow up visit	Study-specific visit to the participant's place of residence after the treatment initiation visit. Only conducted for participants selected and consenting to participate in the tool evaluation.

SUMMARY

Background

The burden of tuberculosis (TB) varies geographically, but this heterogeneity is often poorly understood at the local level. In Malawi (and elsewhere) TB patients' addresses or description of residence location are recorded at TB registration to allow patient follow up, but these data are rarely used for spatial analysis. Moreover, in areas without official address systems, the accuracy of descriptions is often low, which is a significant obstacle for patient follow-up. The development of a method to systematically collect high-resolution TB patient location of residence could allow better retention in TB treatment programmes and could be used to maximise efficiency of local TB programmes through targeted interventions to high burden areas.

In this study we aim to develop and evaluate a novel electronic mapping tool (electronic PArticipant Locator - ePAL) to map the geographical distribution of the place of residence of people registering for TB treatment in Blantyre City, Malawi. If successful, it is envisaged that this tool could have significant future utility for similar mapping of other diseases of public health importance, and for patient follow up for TB and other diseases for which patients are managed in the community in Blantyre, Malawi, and beyond.

Type of study

A prospective, single-site, epidemiology study using the ePAL tool on electronic tablets at TB centres in Blantyre City for clinic-based collection of participant information including demographics, health status and the physical location of place of residence of patients aged ≥18 years initiating TB treatment.

Development of the ePAL tool will require annotation of Blantyre map images with places of interest (POIs) to help participants orientate themselves in the map. This will be achieved by community engagement through Health Surveillance Assistants (HSAs) to identify important landmarks such as markets, places of worship, streets or shops,

which will then be mapped by GPS location capture by study field staff and uploaded to the maps on the tablets.

The evaluation period of the ePAL tool will enrol approximately 196 participants to evaluate the agreement between the ePAL-recorded co-ordinates taken at the health centre and a measurement taken using GPS at a follow up visit at the participant's place of residence. This group will also be asked some additional questions during the treatment initiation visit pertaining to exposure to TB risk factors, care seeking behaviour and socio-economic status (SES). The evaluation period will be preceded by a run-in period, during which the collected evaluation data will be used for refinement of the tool and its implementation.

Objectives and endpoints:

Primary objectives

- To map the spatial distribution of the place of residence of all adults initiating treatment for pulmonary tuberculosis.
- To estimate the census enumeration area (CEA)-wise case notification rates of all adults initiating treatment for pulmonary tuberculosis.

Secondary objectives

- To map the spatial distribution and calculate CEA-wise case notification rates of the place of residence of all adults initiating treatment for tuberculosis.
- To map the spatial distribution and calculate CEA-wise case notification rates of the place of residence of microbiologically-confirmed adult cases initiating treatment for pulmonary tuberculosis.
- To map the spatial distribution of adult pulmonary tuberculosis notifications, stratified by age, gender and HIV status.
- To describe the spatial distribution of adult pulmonary tuberculosis notifications according to socioeconomic status, care seeking behaviour and risk factors for tuberculosis.

Evaluation objective

- To evaluate the spatial accuracy of ePAL for identifying place of residence of adult TB cases resident in Chilomoni or Ndirande and presenting to the Ndirande, Chilomoni or Queen's Hospital TB clinics.

Study population: Adult residents (aged 18 years or above) of Blantyre City registering for TB treatment.

Methods and Procedures: A new software programme consisting of case report forms and maps for data collection will be developed and installed on to tablets for this study. The high resolution satellite photograph maps used in this software for identifying place of residence will require the addition of places of interest (POIs) to allow the users to orientate themselves in the map. These POIs will be identified through community engagement via health surveillance assistants (HSAs), who will assist study field workers to collect GPS coordinates of the POIs so that they can be added to the maps on the tablet.

All TB patients registering for treatment in participating health centres in Blantyre City will be invited to participate by the TB officer or nurse registering them for treatment. TB patients will be provided with an information sheet, and verbal consent obtained. Participant information such as demographics and health status will be recorded in the ePAL electronic case report forms (eCRFs). The participant will be guided by the TB officer through maps on the tablet to locate and select their place of residence. The Hit TB Hard study (COMREC P.011/10/1020) is an on-going active case finding study, for which sputum is collected for monitoring and evaluation purposes throughout Blantyre City. The sputum barcode from the Hit TB Hard study will be recorded to link the data collected in this study to the participant's culture results. Data will be stored on the tablet, and uploaded daily to the MLW server via 3G connection.

Participants will be invited to participate in the evaluation group if resident in Chilomoni or Ndirande and presenting to the Ndirande, Chilomoni or Queen's Hospital

clinics. All such patients presenting during the run-in period will be enrolled (up to a limit of as the number of home visits possible per day by field staff), and approximately 196 patients will be enrolled during the evaluation period. For those willing to participate in the evaluation, written consent (or witnessed thumbprint) will be obtained. The evaluation includes additional collection of data related to SES, care seeking behaviour and TB risk factor exposure during the TB treatment initiation visit, plus a study follow-up visit to the participant's place of residence, where a GPS reading of the place of residence will be taken.

Data collection and management: Data capture for this study is mostly electronic. Data collected at the treatment initiation visit will be collected directly on the electronic tablet. For the evaluation group participants, coordinates collected during evaluation follow-up visits at the patient's place of residence will be measured using a GPS recorder and logged on a paper record sheet. Data will be managed by the study team in MLW and LSHTM in the MLW data office and LSHTM offices. The clinic-based tablets will be password protected and automatically lock after 5 minutes of inactivity to protect the participant data stored on the tablet. The electronic database on the server at MLW (and the daily-synced LSHTM database) will be password protected and maintained on a secure server. Data will be backed up weekly on password-protected laptops, and secure off-site servers.

Statistical considerations and analysis: The study will attempt to enrol all adults registering for TB treatment at participating health centres in Blantyre City, estimated at 2800-3000 patients in a 12-month study period. During the 2-month run-in period at the start of the study, data collected will be used for refinement of the tool and its implementation. The evaluation period of the study will enrol approximately 196 participants to assess the reliability of ePAL versus GPS measurements of place of residence. Given the likely inaccuracy of locating to the household level, but taking into account the high density of points of interest to assist orientation in the map, the expected level of agreement of ePAL to within 20m of GPS measurements is approximately 85%, and the minimum acceptable level of agreement for the tool to be of value is 80%. This sample size of 196 participants is estimated to be sufficient to

provide precision of 5% when estimating proportion agreement with point estimate of 85% (i.e. lower limit of 95% confidence interval of 80%).

Primary analyses will map the spatial distribution of pulmonary TB cases in Blantyre city, and estimate the case notification rates per 100,000 population per year for each CEA. Secondary analyses will estimate the spatial distribution and case notification rates with all TB cases and microbiologically-confirmed TB cases as the numerator. Mapping, and case notification rates where possible, will be conducted for data stratified by age, gender and HIV status. Descriptive analyses of SES, care seeking and TB risk factor data will also be conducted with data from participants enrolled into the evaluation group.

Ethical considerations: This study is likely to benefit the community through better understanding of the distribution of TB cases in Blantyre city and potential for improved patient follow up. Application is being made to the Institutional Review Boards of College of Medicine, Blantyre Malawi (COMREC), and London School of Hygiene And Tropical Medicine, UK. Risks from study participation are minimal, and relate to linkage of personally identifiable information (PII) to sensitive health data, and to the risk of revealing health information. All possible steps will be taken to minimise these risks, such as use of a numerical unique identifier instead of names in the database, and conducting the home visit in a manner that minimises risk of revealing health status, by visiting at a time agreed by the participant, not bearing any labels indicating the field worker is part of a TB or HIV study, and ensuring the health information about the participant is not communicated to others unless express permission is given by the participant.

Data collected in this study is the information recommended to be collected as part of patient registration in the national TB system (place of residence, demographics, health status) and the potential for harm is minimal, therefore verbal consent is considered appropriate for participation in this study. Permission to obtain verbal consent is requested from the ethics committee. Written consent (or witnessed thumbprint for participants who are illiterate) will be obtained for those participating in the evaluation group, as some information collected is not routinely collected.

Expected findings and their dissemination: It is anticipated that there may be large differences in numbers of cases and TB rates between some census enumeration areas, these may be linked to known risk factors for TB. Weekly reports containing interim summaries of the data collected in the main study and tool evaluation, by health centre and for the study as a whole, will be sent to the TB officers and nurses involved in the study via 3G to the tablets at the health centres. Dissemination of final results will be through local and international conferences, community feedback meetings, NTP meetings, and peer-reviewed publications.

BACKGROUND

Variation in the burden of tuberculosis (TB) is mostly well recognised at national and regional levels as a result of government surveillance systems. However, such heterogeneity is often poorly understood at the local level, such as between areas in a city. A small number of studies in Sub-Saharan Africa have demonstrated such local variations exist for TB disease. For example, in Cape Town, South Africa, 39 enumerator districts (across two communities) with an average population of 900 inhabitants reported adult TB notification rates ranging between 0 and 2,847/100,000 adults per year.[1] Similar variations were reported in a settlement-level study in the Gambia, where rates in settlements of 4,000-68,000 people ranged from 48 to 239/100,000 population/year.[2] The potential for such large local variations in TB disease burden demonstrates that regional or even city-level data may not be sufficient for designing targeted control programmes, and highlights the need for local-level burden of disease information.

Efficient targeting of control programmes based on geographical variation in burden of disease is well documented for many diseases. Malaria, for example, is heterogeneously distributed in certain settings, where it has been demonstrated that using local epidemiology for rational targeting of interventions is essential for effective and efficient disease control.[3] Spatial mapping and an analysis of the associated risk factors for TB could play an important role in understanding transmission dynamics, targeting interventions for disease treatment and risk

reduction, and identifying MDR TB disease clusters.[4] Such local data could be essential for cost-effective targeting of health interventions to the localities experiencing the greatest burden of TB disease, but for tuberculosis these data are rarely available.

Existing spatial studies often use enumerated DHS populations or prospectively capture place of residence using GPS, which is resource-intensive.[2,5,6] Therefore, due to costs or logistics, such detailed prospective TB case mapping in non-enumerated populations is uncommon. In Sub-Saharan Africa, spatial and temporal analyses of household-level TB data have been conducted in communities in only a few countries, namely South Africa, Ghana and Ethiopia.[2,5,6] In Malawi (and elsewhere) TB patients' addresses are recorded at TB registration to allow patient follow up, but these data are rarely used for spatial analysis. Moreover, in areas without official address systems, the accuracy of such descriptions is usually low and patients often seek care outside of the TB clinic catchment area of their place of residence, creating significant obstacles for patient follow-up. The development of a low-cost method to systematically collect high-resolution location of residence when presenting at TB clinics could allow better retention in TB treatment programmes and could be used to maximise efficiency of local TB programmes through targeted interventions to high burden areas.

The mapbook system developed in Blantyre, Malawi is a validated paper based tool demonstrated to accurately and rapidly identify the location of place of residence of anti-retroviral (ARV) therapy initiators at health facilities.[7] Such a tool allows mapping of non-enumerated populations, as are found in urban Blantyre's informal settlements, to residential clusters of approximately 1300 people. Although this tool proved valuable for identifying participants from study clusters presenting at health centres, paper-based tools are widely being phased out in field studies in favour of electronic data capture (EDC). Therefore, to improve its utility by increased throughput and automation, and to future-proof such a tool, an electronic system would be the logical evolution of the paper mapbook. In addition, it is possible that an electronic system could be used to capture cases down to the household level as opposed to clusters as used in the paper mapbook study, which would provide greater

resolution in disease maps and could be developed further as a tool for patient follow up in subsequent studies. In the future, such an electronic capture system could be an appropriate tool for much broader data collection in Blantyre and Malawi as part of routine case data capture to understand local-level tuberculosis epidemiology, and target limited available resources more effectively.

Although Malawi has experienced a steady decline in TB incidence since 1998, progress has been slow and incidence of disease remains high fuelled by high HIV prevalence (National TB incidence 163/100,000 population/year).[8] Burden of both HIV and TB disease is highest in Southern Malawi, accounting for approximately 60% of reported TB cases.[9] Therefore there remains a significant need for research to assess and reduce TB disease burden. Two major TB and HIV studies constituting the Hit TB Hard research programme are on-going in Blantyre. This includes one demonstration project assessing the impact of active case finding (ACF) in a population of approximately 108,000 adults in North West Blantyre, and a cluster randomised trial of home-based HIV testing in 33,600 adults residing in 28 clusters nested within this area. As monitoring and evaluation for these studies, TB culture testing is available to patients presenting at TB centres throughout Blantyre City. At time of sputum collection for culture testing, the paper mapbook is used to identify whether patients are resident in one of the home-based HIV testing intervention clusters. An electronic tool for locating place of residence would be a highly valuable addition to the Hit TB Hard research programme, both to assist patient follow up, and to permit spatial mapping of TB to inform implementation of public health interventions, such as targeted contact tracing and active case finding, and community based TB and HIV diagnosis and treatment in hotspots.

In South Africa, spatial mapping studies have been used to understand the risk factors associated with local disease clusters.[3] In this study, risk factor data will be collected from evaluation participants, if TB hotspots are identified as overlapping with specific risk factors these data could inform selection of appropriate interventions in addition to directing where they should be implemented.

There is a clear need for tools to better understand the local-level distribution of TB cases in Malawi. In this study we aim to develop and evaluate a novel electronic mapping tool (electronic PArticipant Locator - ePAL) to map the geographical distribution of the place of residence of people registering for TB treatment in Blantyre City, Malawi. If successful, it is envisaged that this tool could have significant future utility for similar mapping of other diseases of public health importance, and for patient follow up for TB and other diseases for which patients are managed in the community in Blantyre, in Malawi, and beyond.

RATIONALE, STUDY DESIGN AND OBJECTIVES

Rationale

Understanding local variation in burden of disease and associated risk factors can permit more efficient selection and targeting of limited resources for TB disease control. However, studies to generate these data usually require existing enumeration of the population of interest or large budgets to conduct this ahead of the study. Such studies are not appropriate where resources are constrained or for national tuberculosis programmes where surveillance covers large populations, therefore a low-cost methodology for identifying location of TB cases in the context of the national TB programme would be desirable to permit collection of such local-level data.

Accurate local TB case distribution data are currently not available in Blantyre, Malawi. This is primarily due to the lack of municipal address systems, but is also impeded by the inaccuracy of verbal descriptions of place of residence and the large amount of resources that would be required to collect such descriptions from the clinic paper registries. Furthermore, the inaccuracy of verbal descriptions means that TB staff are often unable to locate patients lost to treatment follow up. Thus, there is a need to digitise and improve the accuracy of such data collection to allow efficient and effective collection of place of residence, to improve both patient-level follow up and decision making by the TB control programme. This study will aim to meet this need through the development and evaluation of an electronic tablet tool to collect place of residence of TB patients when registering at TB centres in Blantyre, Malawi. Such a

tool has the potential for broader use within control programmes and research studies throughout Malawi and beyond for patient follow up and to generate data to inform TB control decision making.

Study design

This study consists of two components: A) The prospective collection of demographic data and home location of TB treatment initiators using the ePAL tool; and B) an evaluation of the accuracy of data collected using ePAL by comparison to GPS co-ordinates collected at the patient's place of residence.

Main Study

This prospective epidemiology study consists of clinic-based collection of the place of residence of patients ≥18 years initiating TB treatment (TB notifications) using the ePAL tool at TB centres in Blantyre City.

This study is an epidemiological study, as such there will be no study-directed administration of pharmaceutical interventions nor additional collection of specimens beyond what is already part of the routine national system and existing Hit TB Hard study (COMREC P.11/10/1020). There is no randomisation to a medical intervention or care provided, and all participating clinics will receive the ePAL tool for data collection.

The study will be initiated in the Queen Elizabeth Central Hospital TB clinic, where any refinements required for the study logistics, staff training or the ePAL tool can be made. The ePAL tool will then be rolled out across all participating TB registration centres in Blantyre City (population approximately 661,256)[10]. Enrolment will continue until the end of the ePAL evaluation phase, and if demonstrated to be accurate will be continued until the end of the monitoring and evaluation element of the Hit TB Hard study, expected 3-4Q 2015.

All patients ≥18 years presenting to participating TB centres for TB treatment initiation will be invited to participate in the study; those verbally consenting to participation

will be enrolled. Upon enrolment, routine demographics and disease and treatment history will be collected. As part of routine care, HIV testing is offered to all patients registering for TB treatment; therefore HIV status can be ascertained from the results of the test at TB treatment initiation, or alternatively patient reporting or patient records are also a suitable information source. Patients may refuse for HIV status to be recorded. During the same visit physical location of the treatment initiator's place of residence will be recorded by the patient using ePAL, assisted by the centre's TB officer. An evaluation of the ePAL tool is nested within the study. Approximately 196 patients will be invited to participate in the evaluation.

Tool Evaluation

An assessment of the agreement between the clinic-based capture of place of residence using the ePAL tool and the gold standard of GPS co-ordinates captured at the place of residence is required to evaluate the accuracy of the tool.

Participants will be invited to participate in the evaluation group if resident in Chilomoni or Ndirande and presenting to the Ndirande, Chilomoni or Queen's Hospital clinics during the run-in or evaluation periods. All such patients presenting during the run-in period will be enrolled (up to a limit of as the number of home visits possible per day by field staff), and all such patients up to a limit of approximately 196 patients will be enrolled during the evaluation period (See 7.1 for sample size calculations). If the enrolment limit is reached before the end of the evaluation period, all participants will be automatically assigned to the main study. During the 2-month run-in period at the start of the study, data collected will be used for refinement of the tool and its implementation. The data collected in the evaluation period will be used for the formal evaluation of the accuracy of the ePAL tool for identifying place of residence.

For those participating in the evaluation, written consent (or witnessed thumbprint) will be obtained in place of verbal consent. The evaluation will include the same data collection as detailed in the main study, but includes additional collection of data related to SES and TB risk factor exposure during the TB treatment initiation visit, plus there will be a follow up study visit to the participant's place of residence. At this visit,

GPS co-ordinates will be collected at the participant's place of residence by a member of the study team using a mobile GPS device.

Objectives and related endpoints

Primary objectives

- To map the spatial distribution of the place of residence of all adults initiating treatment for pulmonary tuberculosis.
- To estimate the CEA-wise case notification rates of all adults initiating treatment for pulmonary tuberculosis.

Primary endpoints

- Occurrence of any adult pulmonary TB case initiating TB treatment
- Criteria/definition: Any new or retreatment pulmonary TB patient aged ≥18 years starting a new course of TB treatment.
- Geospatial characteristics of TB case place of residence
- Criteria/definition: Geospatial reference of place of residence as determined by the patient using the ePAL system.

Secondary objectives

- The secondary objectives aim to map the distribution of cases based on alternative TB case definitions, and to stratify by various factors of interest.
- To map the spatial distribution and calculate CEA-wise case notification rates of the place of residence of all adults initiating treatment for tuberculosis.
- To map the spatial distribution and calculate CEA-wise case notification rates of the place of residence of microbiologically-confirmed adult cases initiating treatment for pulmonary tuberculosis.
- To map the spatial distribution of adult pulmonary tuberculosis notifications, stratified by age.
- To map the spatial distribution of adult pulmonary tuberculosis notifications, stratified by gender.
- To map the spatial distribution of adult pulmonary tuberculosis notifications, stratified by HIV status.

- To describe the spatial distribution of adult pulmonary tuberculosis notifications according to socio-economic status, care seeking behaviour and risk factors for tuberculosis.

Secondary endpoints are as follows:

- Occurrence of any adult TB case initiating TB treatment
- Criteria/definition: Any new or retreatment pulmonary or extra-pulmonary TB patient aged ≥18 years starting a new course of TB treatment.
- Occurrence of adult microbiologically-confirmed pulmonary *Mycobacterium tuberculosis* disease initiating TB treatment
- Criteria/definition: New or retreatment case aged ≥18 years starting a new course of TB treatment, determined by sputum culture (liquid or solid), smear microscopy, or Genexpert.
- Demographic characteristics and TB risk factors
- Criteria/definition: Age, gender, socio-economic status, TB history, care seeking behaviours.
- Occurrence of HIV infection
- Criteria/definition: HIV infection determined by patient reporting, medical record of HIV status, or results of new test conducted during the visit as part of routine care.

Evaluation objectives

- To evaluate the spatial accuracy of ePAL for identifying place of residence of TB cases resident in Chilomoni or Ndirande and presenting to the Ndirande, Chilomoni or Queen's Hospital TB clinics.

Evaluation Endpoints:

- Geospatial characteristics of place of residence
- Criteria/definition 1: Geospatial reference of place of residence as determined by the patient using ePAL.

- Criteria/definition 2: Geospatial reference of place of residence as determined by GPS measurement at the place of residence by field researcher.

Estimation of population size

Incidence rates will be calculated for the Census Enumeration Areas (CEAs) as defined in the Malawi 2008 census.[10] Denominator data for these regions within Blantyre city will be available from the 2008 census. Current population will be estimated by applying annual predicted population growth rates for Blantyre Urban available from the National Statistics Office to the 2008 census population estimates.[11] It is hoped that denominators stratified by age group and gender will also be available from this database, but if unavailable, the known age and gender structure for Blantyre city will be applied to the CEA populations. HIV positive and negative population denominators will be calculated for the CEAs using HIV prevalence rates from the Malawi Demographic and Health Survey applied to the census population estimates for the CEAs.[12]

Verbal descriptions of place of residence have proven inaccurate for identifying place of residence.[7] In Blantyre, a paper mapbook used at point of care has been demonstrated to allow accurate and rapid identification of whether ARV therapy initiators are residents within study clusters.[6] Such a tool allows cluster-level mapping of non-enumerated populations, such as urban Blantyre's informal settlements. Although valuable, paper-based tools are widely being phased out in field studies in favour of electronic data capture (EDC). Therefore, the ePAL tool to collect place of residence electronically at the clinic will increase the resolution to which this information can be collected and improve its utility by increased throughput and automation. It will provide greater resolution in disease maps and could be developed further as a tool for patient follow up in subsequent studies.

As this is a new tool, an evaluation component has been built into this study to assess the resolution to which ePAL can accurately record location of residence in Blantyre.

Choice of objectives and endpoints

The selection of pulmonary TB case notification as the primary objective is reflective of the relative contribution of pulmonary disease to burden of disease (71% of all notified TB in Malawi in 2003),[9] its importance in disease transmission dynamics, and availability of TB-specific diagnostics providing relative reliability of diagnosis. TB case notifications have been selected as a proxy for TB incidence due to practicality in terms of study design, and due to the increased case detection in Blantyre City provided by the extended monitoring and evaluation element of the Hit TB Hard study (COMREC P.11/10/1020).

Although resource constraints usually limit the availability of culture diagnosis, given the citywide sputum culture testing on-going as part of the Hit TB Hard study, microbiologically confirmed pulmonary TB is a viable outcome of interest for this study. It is thus included as a secondary outcome of interest.

Extra-pulmonary TB is not analysed in isolation as it is anticipated that the small number of cases may yield personally identifiable information (PII). Therefore, a composite endpoint of all TB cases (pulmonary and extra-pulmonary) is included in the secondary objectives to consider extra-pulmonary cases.

Not only is it important to understand where the hotspots of TB are occurring to tailor public health interventions appropriately, it is also essential to understand the demographic and risk factors associated with TB hotspots. This will be explored through mapping the residence of TB cases stratified by demographic and risk factors. The use of routinely collected data as planned in this study is vulnerable to changes in the health system, such as introduction of more sensitive tests or increased screening rates. However, as this study does not aim to estimate temporal trends, such improvements could only affect study objectives if such improvements were implemented unevenly across Blantyre city.

Choice of evaluation objectives and gold standard

The tool evaluation will be an accuracy study comparing the ePAL-measured location of place of residence with a gold standard measurement.

In the MacPherson et al. study, two independent research assistants were unable to identify the cluster residency based on verbal descriptions in 31.8% and 59.7% of cases respectively;[7] therefore when considering resolution down to the household level, verbal descriptions would likely be highly inaccurate. Therefore, verbal descriptions will not be used as the gold standard in this study. Instead GPS coordinates, accurate to a resolution of a few metres, will be used as the gold standard for place of residence in this study.

The place of residence coordinates recorded using the ePAL tool will be the comparator evaluated against the gold standard of GPS coordinates measured at the place of residence.

Duration of study

Study start is anticipated in 3Q 2014, with phased roll out across health centres, starting at Queen Elizabeth Central Hospital. The evaluation element will consist of approximately 2-month run-in period and approximately 6-month evaluation period (Figure 1). Data collection using ePAL will continue until at least the end of the evaluation period (anticipated 2Q 2015), and if the evaluation demonstrates good accuracy, data collection will continue until the end of the extended monitoring and evaluation of the Hit TB Hard study, which is anticipated to finish in 3-4Q 2015.

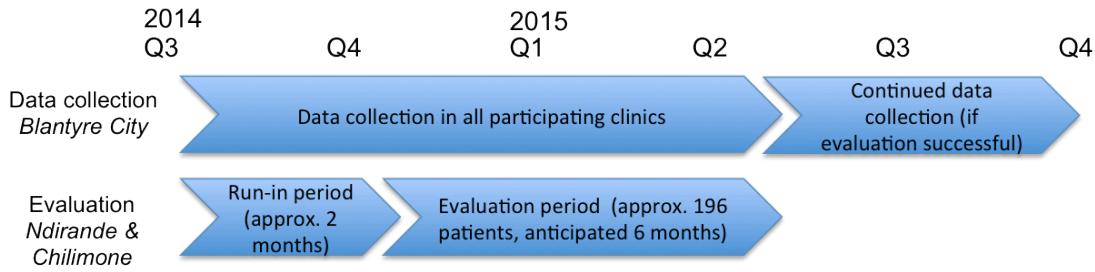


Figure 1: Anticipated timelines and location of the data collection, run-in and evaluation periods of the TB ePAL study

STUDY POPULATION

Only patients ≥18 years presenting to participating health centres for TB treatment initiation will be invited to participate in the study.

TB services and registrations

The study will be conducted in health centres providing TB registration and care in Blantyre City, encompassing 396 CEAs and a population of approximately 661,256.[10] Approximately 344,500 residents are ≥18 years, therefore this study will be implemented in the services providing care to this population of study-eligible age. In Malawi, TB treatment is only available through government facilities, therefore TB treatment registration through these clinics should provide relatively comprehensive coverage of the city's total treatment initiators. The health centres in Blantyre City anticipated to participate are listed in Appendix A.

In 2012, 3,347 TB cases were registered in urban Blantyre, of which 3,019 were adults aged 15 years or above (Hit TB Hard study, unpublished). Therefore, in the planned study period we anticipate to enrol approximately 2,800-3,000 adults aged 18 or above.

Participant selection, inclusion and exclusion criteria

- Participants will be enrolled from participating TB clinics in Blantyre (Appendix A) at registration for TB treatment.
- The inclusion criteria for participation in the study are:
- age 18 years or above
- new registration for TB treatment (either new case or retreatment case)
- verbal consent to participate in the study (or written/witnessed thumbprint consent in evaluation population)

The exclusion criteria for participation in the study are:

- age less than 18 years

- patients on continuing TB treatment
- declines to participate

All participants can decide to withdraw from the study at any time. A record will be made of any participants withdrawing their participation after enrolment.

Participant selection for tool evaluation

Participants will be invited to participate in the evaluation group if resident in Chilomoni or Ndirande and presenting to the Ndirande, Chilomoni or Queen's Hospital clinics during the run-in or evaluation periods. All such patients presenting during the run-in period will be enrolled (up to a limit of as the number of home visits possible per day by field staff), and all such patients up to approximately 196 patients will be enrolled during the evaluation period. If the enrolment limit is reached before the end of the evaluation period, remaining participants will be enrolled in the main study. Participants willing to participate in the study, but declining to participate in tool evaluation will be included in the main study and evaluation refusal recorded.

METHODS AND PROCEDURES

Tool development

Software development

New software will be developed for the purpose of this study. The software will run from electronic tablets based in the health centres. The tool will consist of an electronic data collection form to input basic demographic and health status data as routinely collected in the TB registers. The TB officer/nurse will then help the participant navigate through a map of Blantyre to locate and select the patient's place of residence, automatically recording the location coordinates upon selection.

Map source and annotation

High-resolution satellite photographs of Blantyre city are available from Google Earth. The current images were collected on 17/09/2013, therefore should accurately depict the location of residences during the study period.

As found by early testing of the paper mapbook, existing Google map points of interest (POI) proved ineffective for patient orientation in the map (Augustine Choko, Personal Communication). POIs identified through community engagement proved much more effective.

In this study, to navigate to the household level it is expected that a much higher density of POIs may be required. A small pilot of POI densities involving 2-5 resident Community Health Workers (Health Surveillance Assistants: HSAs) will be required to gauge the density needed (Figure 2). A field worker will record the GPS location of the POIs and maps labelled with POIs will be tested in each area by, for example, inviting people in the street or on market day to participate. At least 20 participants will take part in the pilot. Those willing to participate will be verbally consented, and asked to locate their place of residence on the map, followed by a GPS reading of their place of residence by the field worker. Participants in the pilot will be compensated 1000K for their participation. The results from the pilot will be used to identify the density of POIs relative to population density that will be collected across the city (i.e. in densely populated areas the number of POIs may need to be adjusted to maintain a similar POI:population density).

For the POIs collected across Blantyre City, resident HSAs will assist with the identification of POIs in their respective catchment areas. An introductory meeting will be held with HSAs to introduce the study and explain the methodology to follow for POI collection. HSAs will be compensated 2000K for attending the meeting. Following HSA engagement with the local community to identify POIs, study staff will collaborate with HSAs to collect GPS coordinates of each POI using mobile GPS handsets. HSAs will be compensated 3000K upon completion of POI GPS collection.
These coordinates and POI name will be uploaded to the tool maps. The software will include the functionality to add or update POIs during the study to ensure map annotations are up to date.

Training and implementation

TB officers will be provided training at the beginning of the study in the use of the ePAL tool. Roll out of the ePAL tool will be phased, beginning with Queen Elizabeth Central Hospital, where trouble shooting and tool refinement will take place. ePAL will subsequently be rolled out in the participating centres across Blantyre City (Figure 2). Monthly meetings are held with TB Officers as part of the Hit TB Hard study, so this study will be added to the agendas of these monthly meetings to allow for minor re-training and to resolve any problems in the study. If major problems occur or if a more substantial re-training is deemed necessary during the course of the study, a re-training meeting will be scheduled.

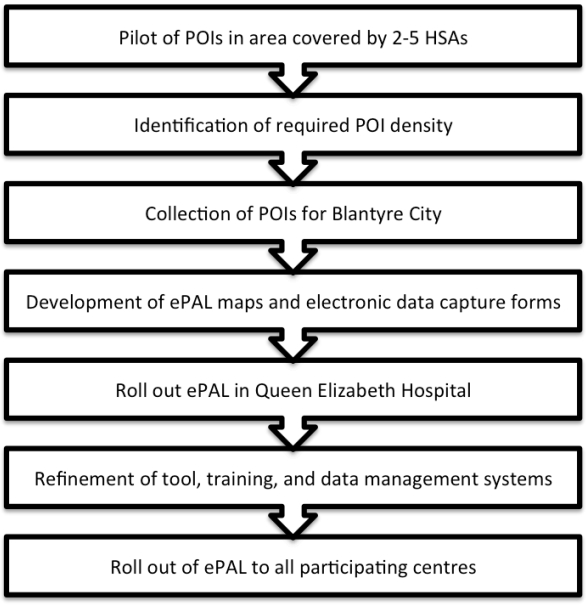


Figure 2: Collection of POIs and roll out of ePAL tool

Study procedures

Study procedures table

The study procedures for participants are outlined in Table 1.

Table 1: Study procedures table

Study Procedure	Main Study	Tool Evaluation
Study Visit 1 – *At clinic TB initiation visit*		
Verbal consent and information sheet provided at clinic	x	
Written consent and information sheet provided at clinic		x
Record barcode of sputum sample (collected as part of Hit TB Hard)	x	x
ePAL collection of demographic and health status data (e.g. age, HIV status, receipt of ART)	x	x
ePAL collection of SES and risk factors for TB (e.g. type of house, number of people sharing residence)		x
Place of residence collected using ePAL maps	x	x
Patient data saved to device	x	x
Study Visit 2 – *At participant's place of residence*		
Collect GPS coordinates at place of residence		x

ePAL field usage

This study will consist of the implementation and evaluation of an electronic participant location capture tool (ePAL) in participating health centres in Blantyre City.

TB Officers and nurses registering TB patients in participating TB clinics will be responsible for data collection using the ePAL tool. Ahead of ePAL implementation, all relevant staff will be trained in collection and storage of patient data, and push/pull of data and updates from the server. During the study, re-training and trouble shooting will be added to the monthly meetings held with the TB officers.

Following ePAL launch, all patients meeting the inclusion/exclusion criteria registering for TB treatment will be invited to participate. Following the usual treatment

registration processes (Figure 3, grey), patients willing to participate will be allocated to the main study, except for during the run-in or evaluation phases where if the patient is resident in Chilomoni or Ndirande, they will be allocated to the evaluation group (Figure 3).

Participants eligible and willing to participate in the main study will be verbally consented (For information sheet and verbal consent forms in English and Chichewa, see Appendix B). Refusals to participate in the study will be documented. Where consent is given, each participant will be given a unique participant identifier. Using this identifier, the participant record will be linked to the sputum sample results from the Hit TB Hard monitoring and evaluation study. Details of participant demographics and relevant medical information will be recorded by the TB officer or nurse in the ePAL electronic data collection forms. The TB officer/nurse will then guide the participant through the ePAL maps based on participant description of area of residence, assisting the patient to locate and select their place of residence on the tablet screen. The data are then saved on the tablet and uploaded daily via 3G to the MLW server. (Figure 3, white, left branch)

A weekly back up of the data will be collected directly from the tablet by a field worker and stored in an independent location from the main data set.

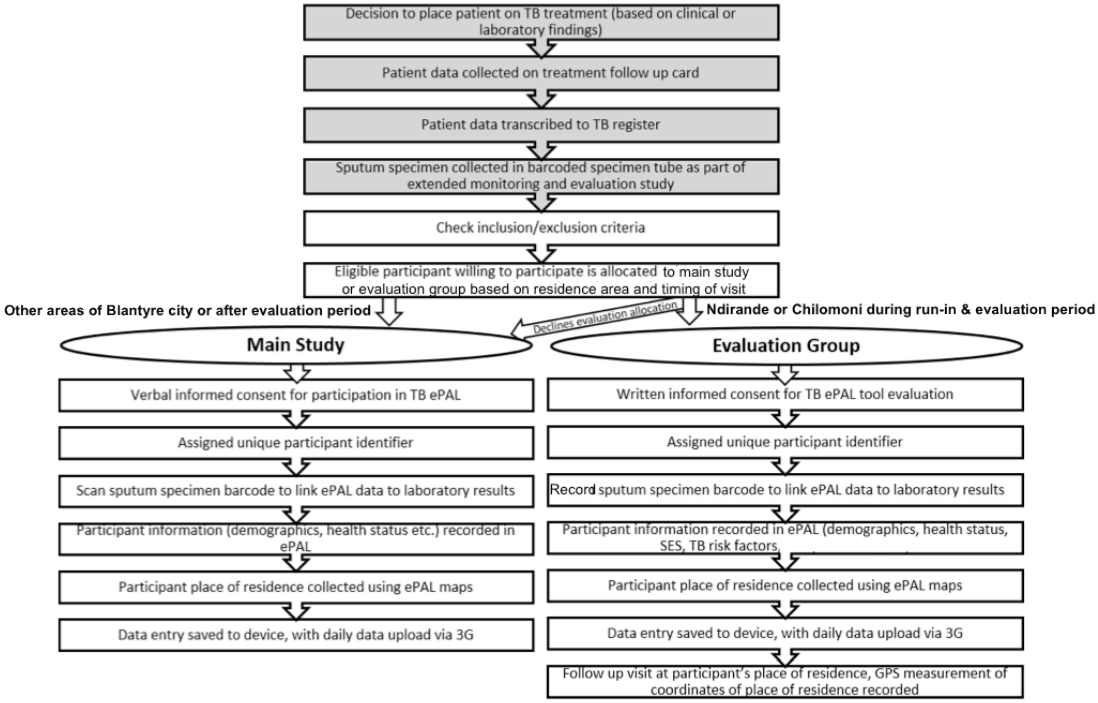


Figure 3: Participant flow through existing (grey) and study (white) processes

For patients eligible and willing to participate in the tool evaluation group, written consent will be requested (For information sheet and written consent forms in English and Chichewa, see Appendix C). In this group of participants, data collection will be the same as detailed above, plus the ePAL data collection forms will include some additional questions pertaining to SES, care seeking behaviour and risk factors for TB (Figure 3, white, right branch). These data will be used to investigate associations between hotspots and TB risk factors/SES, but are only collected for the participants in the evaluation group since written consent is required to collect this information. Refusal to participate in the evaluation will be recorded to allow an estimation of any bias potentially caused by refusals.

For those participating in tool evaluation, the treatment initiation visit will be followed by a visit to the place of residence. At this visit, GPS coordinates of the place of residence will be recorded. GPS data will be linked to the ePAL database based on the participant ID. Participants in the evaluation group will be compensated 1000K for their time. Continuation of care beyond the registration visit will not be impacted by this study.

Informed consent

This protocol requests to waive the requirement of signed or witnessed consent forms for the main study, as: - TB patients are defined and investigated in line with national policy. In this study, place of residence location is recorded using an electronic map tool as opposed to verbal description, but the target information is the same regardless of measurement method; The potential for harm is minimal

Verbal consent will be obtained from all eligible patients willing to participate and an information sheet provided specifying that this is a research study (see Appendix B) and providing other essential participant information. The expression of verbal consent and the name of the person taking the consent will be documented.

For participants participating in the evaluation component, there will be questions included pertaining to SES, care seeking behaviour and risk factors for TB and a follow up visit to the patient's place of residence, which is collection of information beyond standard practice. Written consent will therefore be obtained and an information sheet provided (see Appendix C) for patients agreeing to participate in the evaluation component.

Refusal to participate in either the main study or evaluation group will be recorded, and numbers refusing to enter the study or evaluation group documented.

Quality assurance and monitoring

Project managers will be responsible for day-to-day monitoring. A fortnightly meeting will be held with all field workers involved in this study to discuss any implementation issues. Monthly meetings are already in place for the Hit TB Hard study, so will be extended to incorporate trouble-shooting of any reported problems, small training updates, and to provide feedback to the officers on study progress and results. TB officers will be compensated for their time for participation in these meetings.

For quality assurance purposes, spot visits to the clinics carrying out ePAL data collection will be carried out to check the correct usage of the tool and to identify any implementation issues. This will be at highest frequency (at least once/week) for the first month of the study when it is anticipated that some implementation issues will

need to be resolved. For the reminder of the study monthly visits will be continued. Spot checks will consist of checking that new TB registrations are being recorded in the ePAL tool by comparing the number in ePAL to the number in the TB register, observing how ePAL is being implemented, and discussion with the TB officers about any implementation issues. Once weekly a member of the study team will check the tablets are well maintained and collect a backup of the data directly from the tablet. Spot checks will also be conducted for approximately 5% of field worker visits to place of residence to ensure GPS readings are taken correctly. These will be weighted so that around a third of the checks occur in the first month, and the remainder of the spot checks are distributed throughout the remainder of the run-in and evaluation period.

The RSC and MLW Research Governance Officer or the COMREC compliance officer may carry out external audit of this study, but otherwise, no external audit will be carried out.

DATA COLLECTION AND MANAGEMENT

Data management

Data will be managed at the project offices in Malawi-Liverpool-Wellcome Trust, with a daily live link to LSHTM, UK for additional data management. All data will be uploaded daily from the tablets in the clinics to the database on the MLW server via the 3G network. To protect against data loss due to network issues or tablet loss/damage, a separate back up database will be created from data collected directly from the tablets once per week. This back up database can be used to fill any gaps in the main database if required.

Some data checks will be encoded into the data collection system. Data checks and cleaning will be conducted on a rolling basis throughout the duration of the study. Any inconsistencies will be investigated and resolved where possible. Data analysis will be carried out in-house and at LSHTM.

Access to the final data set will be limited to the data manager (David Matiya), PI (Rebecca Harris) and statistician (Emily Webb). Sensitive information (including HIV results) will not be linked to the participants name in the final data set.

Data capture forms

Data will be captured on electronic tablets in the following data capture forms:
eCRF to collect demographics, health status and symptom information.
Additional eCRF to collect data on socioeconomic status, care seeking behaviour and TB risk factors for participants in the tool evaluation group. High-resolution map photographs for capture of place of residence.

Adverse events and confidentiality

No serious adverse events (SAEs) are anticipated as there are no medical interventions outside of routine standard of care. Any breaches of confidentiality following TB or HIV diagnosis will be systematically recorded and reported back to the respective IRBs. All possible steps will be taken to minimise the risk of confidentiality breach. From a data management perspective, a numerical unique identifier will be used instead of names in the database, file names that do not overtly identify the file content (e.g. data1) will be used for sensitive data, and where possible sensitive health information and location information will be stored in separate files linked only by the unique patient identifier. The home visit will be conducted in a manner that minimises risk of revealing health status, by visiting at a time agreed by the participant, not bearing any labels indicating the field worker is part of a TB or HIV study, and ensuring the health information about the participant is not communicated to others unless express permission is given by the participant.

Data security

All tablets holding patient data will be password protected, with passwords only known by study staff and the TB officers and nurses trained to use the tablets at the clinics. The tablets will automatically lock after 5 minutes of inactivity to protect the patient data stored on the tablet. Unnecessary functionality on the tablet will be

disabled as much as possible to reduce the risk of non-study related use of the tablets, thus minimising the risk of damage, loss or unauthorised transfer of participant data. The database stored on the MLW server and daily live linked to LSHTM, UK will be password protected. All data will be backed up daily by the MLW Data Office, with offsite back up once weekly.

Quality assurance

Data will checked for internal inconsistencies or missing fields during data entry, and on a rolling basis during the study.

STATISTICAL CONSIDERATIONS

Sample size justifications

In 2012, 3,347 TB cases were registered in urban Blantyre, of which 3,019 were adults aged 15 years or above (Hit TB Hard study, unpublished). Therefore, in the planned study period we anticipate to enrol approximately 2,800-3,000 adults aged 18 and above.

The evaluation of the reliability of ePAL compared to the gold standard of a GPS reading at the place of residence will be conducted for a sub-set of patients registering for TB treatment.

Given that good map literacy is not expected in the study population, very high agreement between the GPS and ePAL readings in the evaluation component would not be expected. However, experience from the mapbook project has demonstrated that with the assistance of a local TB officer with training in map navigation, patients were able to locate their residence to within and outside clusters with 100% and 95% agreement, respectively. Clusters encompassed populations of approximately 1300, therefore greater inaccuracy is expected when attempting to identify at the residence level. To assist with orientation in the maps, a higher density of POI annotations will be included in the ePAL maps and updates will be possible to tailor the map POIs to the evolving needs of the community. A few metres of measurement inaccuracy are expected from GPS readings, and although the level of inaccuracy is currently

unknown for the ePAL location capture it is anticipated to be a slightly larger inaccuracy than the GPS. In addition, point measurements are being taken for houses that are an area not a point, therefore a limited range of coordinates would be correct. Therefore, approximately 10-15m would be a reasonable expectation of the upper limit of measurement inaccuracy given the inherent inaccuracies in the two capture methods and the area of the house. Taking this into account, the sample size is calculated for measuring accuracy of the ePAL method to within 20m of the GPS measurement. Given the likely elevated inaccuracy of locating to the household level, but taking into account that improvements will be made to the tool POI annotations to assist orientation in the map, the expected proportion of participants for whom ePAL is accurate to within 20m is approximately 85%. The minimum acceptable level of agreement for the tool to be of value is 80%.

Using methodology for calculating sample sizes required to estimate a proportion with a pre-specified precision,[13] it is estimated that approximately 196 participants will be required in the evaluation group. This will be sufficient to provide a point estimate of the proportion for whom ePAL is accurate to within 20m of 85%, with a lower limit of the 95% confidence intervals for this proportion of 80% (i.e. a precision of 5%) (Table 2).

Table 2: Sample size calculations for a range of scenarios

	Point estimate of proportion for whom ePAL is accurate to within 20m (%)					
Lower limit of 95% confidence interval (%)		75	80	85	90	95
	65	73	28	13	6	3
	70	289	62	22	9	3
	75		246	49	16	5
	80			196	35	9
	85				139	19
	90					73

Although TB officers will be given training in use of the tool and orientation in the maps, it is anticipated that TB officers will improve at guiding patients through the maps during the run-in phase of the study. Data from evaluation visits will be valuable during this phase to identify underperforming areas to target additional training or map POI annotation improvements. Therefore, to allow for this period of implementation improvement, evaluation visits during this period are not counted within the sample size.

In the evaluation period, approximately 196 participants will be invited to participate in tool evaluation. During this period the tool is expected to be a consistently implemented part of TB registration visits, therefore demonstrating the tool's capability in routine programme settings. The data from these participants will be used for the formal evaluation of the accuracy of the ePAL tool for identifying place of residence.

Planned analyses

The analysis of primary and secondary outcomes for patients registering for TB treatment will be conducted on data captured in the cleaned electronic TB database to which routine quality checks will have been applied. Interim analyses will be carried out on a rolling basis to inform tool development and for regular feedback of results

to the TB clinics. Final analyses will be conducted at evaluation period conclusion and at study conclusion.

Primary objectives

TB case notifications for pulmonary TB will be overlaid on a map of Blantyre to display the spatial distribution of pulmonary TB cases in Blantyre city.

Case notification rate will be calculated as a proxy for incidence rate in each of the census enumeration areas. The location of place of residence, as recorded using ePAL, will be used to identify census enumeration area of residence of participating residents with pulmonary TB registering for TB treatment, which will form the numerator of case notification rate calculations. The population denominator for each of the CEAs will be calculated as described in Section 0. These numbers will be used to calculate case notification rate per 100,000 population per year for each census enumeration area. Exact 95% confidence intervals for the case-notification rates will be calculated. Maps and graphs showing the annual incidence rates by census enumeration area will be created.

Secondary objectives

For the secondary endpoint of the distribution of all TB patients and microbiologically-confirmed pulmonary TB, spatial distribution and case notification rates will be calculated as described for the primary objectives, using the alternative definitions of all and microbiologically confirmed TB (see section 3.3.2 for endpoint definitions) to define the numerator populations.

Stratified TB case notifications will be mapped in Blantyre, as was described in the primary objectives, but stratified by age, gender and HIV status. Age will be stratified in 10 year and 5-year intervals (though with first interval as 18-19yrs), gender as male versus female, and HIV status as positive, negative or unknown. Where possible, case notification rates will also be calculated stratified by age and gender using the same denominator as described in section 3.3.4 stratified by the relevant characteristics. If possible, stratification by HIV status will also be reported, using HIV prevalence rates from the Malawi Demographic and Health Survey applied to the census population

estimates as a denominator.[12] If stratification by HIV status leads to so few patients per area that it risks revealing the HIV status of the participant, either areas will be aggregated together or the information not reported.

Descriptive epidemiology will be used to explore areas of low SES or high levels of TB risk factors. These will be compared to maps of TB burden to observe whether areas of high risk factors are also those with TB hotspots.

Evaluation objectives

Data collected in the run-in period will be used to monitor tool implementation and trigger tool improvements or additional staff training where required. The data collected from approximately 196 participants in the evaluation period will be used for the formal evaluation of the accuracy of the ePAL tool for identifying place of residence as described below.

An evaluation of the accuracy of the tool compared to GPS readings taken at the place of residence is planned. The proportion of participants for whom ePAL is accurate to within 20m of the GPS-recorded location will be estimated, along with a 95% confidence interval. Further analyses of these data will assess the range of the horizontal distance between the ePAL-given location and the GPS-recorded location. The distance root mean square error (defined as the square root of the average of the squared horizontal position errors) will be calculated.

LABORATORY METHODS

No laboratory tests will be conducted as part of this study. Results from HIV tests conducted as part of routine care will be collected wherever the patient is willing for such data to be recorded. Patients consenting to participate in the enhanced monitoring and evaluation element of the Hit TB Hard study will also have sputum culture results available that will be linked to the data collected in this study based on the anonymous unique patient identifier.

ETHICAL CONSIDERATIONS

This study is likely to benefit the study community through better understanding of the distribution of TB cases in Blantyre, which has value for governmental TB programmes in designing control strategies. If the evaluation proves the tool is accurate, the tool has potential for continued use for creating on-going electronic databases for TB case notification data and improved retention of patients, who would otherwise be lost to follow up, through more accurate records of place of residence.
Application is being made to the Institutional Review Boards of College of Medicine, Blantyre, Malawi (COMREC), and London School of Hygiene And Tropical Medicine, UK.

There is no randomisation to a control versus an intervention, so issues of equipoise are not relevant in this study.

On submission for ethics review, permission for verbal informed consent will be requested for participants taking part in the main study using ePAL to identify their place of residence. The grounds for requesting verbal consent are that: TB patients are defined and investigated in line with national policy. In this study, place of residence location is recorded using an electronic map tool as opposed to verbal description, but the target information is the same regardless of measurement method; The potential for harm is minimal.

The main potential for risk of harm could arise if patient identity were revealed through the linking of sensitive health information with personally identifiable information (the location of residence) in the study database. As the objective of the study is to develop a tool to accurately identify the place of residence when at a TB clinic, collection of location coordinates and patient information is unavoidable. However, risk of patient identification will be minimised by only recording the unique patient identifier and not the name in the final electronic database. Only the principal investigator, Rebecca Harris, will hold a master list containing both the name and unique identifier of the participants.

Information leaflets will be provided to all participants, specifying that this is a research study (see Appendix B) and providing other essential participant information. For participants in the tool evaluation group, written (or witnessed thumbprint) informed consent will be obtained, as some of the data collected is beyond what is collected under the national TB programme. The potential for harm is the same as described above for the main study, plus there is an additional risk for the participant's TB disease status to be revealed through the visit to the place of residence. In addition, due to the association between TB and HIV, revealing a patient's TB status could lead also to assumptions in the community about their HIV status. These risks will be minimised by conducting the home visit in a manner that minimises risk of revealing health status, such as visiting at a time agreed by the participant, not bearing any labels indicating the field worker is part of a TB or HIV study, and ensuring the health information about the participant is not communicated to others unless express permission is given by the participant. For those participating in the evaluation component, a different information leaflet will be provided (see Appendix C), specifying that this is a research study and providing other essential participant information.

To protect the identity of study participants, maps prepared for publication will not reveal the location of the place of residence of the participants. When overlaying cases on to maps of Blantyre city, care will be taken to ensure the size of case markers and resolution of the map are insufficient to be able to identify the place of residence of the TB cases.

DISSEMINATION

Policy for sharing data

Data will be primarily shared with scientific collaborators, but we will as far as possible facilitate data sharing with any group requesting access to individual patient records by using anonymised data. Where appropriate, and with the proper safeguards, data will be made freely available through the Malawi-Liverpool-Wellcome Trust (MLW) Programme website.

Alongside original data, anonymised databases will be created and stored with all relevant data to facilitate any data transfer requests that are made subsequently.
Ethical clearance will be sought before data are transferred to other groups for secondary analysis. Priority will be given to local investigators and publicly funded international investigators with data management and sharing policies in line with those of the MLW Programme and the College of Medicine.

Strategy for current and future communication with user communities
Results will be disseminated to health centres through weekly summary reports sent to the health centre tablets. The study results will also be shared with NTP and District Health Office collaborators. The results will also be presented at one of the regular community feedback meetings organised by MLW's research dissemination office, and the team will approach local radio to propose a session for presentation of results with a community phone-in session. Results of the tool evaluation and final data analysis will be presented at national research meetings, such as the annual research day held by College Of Medicine, and presented at international research meetings. Results will be prepared for publication in international peer reviewed scientific journals and published in compliance with the Open Access policy of the Wellcome Trust.

REFERENCES

1. 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. Archives of disease in childhood. 1999 May;80(5):433-7
2. Touray K, Adetifa IM, Jallow A, Rigby J, Jeffries D, Cheung YB, et al. Spatial analysis of tuberculosis in an urban west African setting: is there evidence of clustering? Tropical medicine & international health : TM & IH. 2010 Jun;15(6):664-72
3. Carter R, Mendis KN, Roberts D. Spatial targeting of interventions against malaria. Bulletin of the World Health Organization. 2000;78(12):1401-11

4. Tanser FC, Le Sueur D. The application of geographical information systems to important public health problems in Africa. International journal of health geographics. 2002 Dec 9;1(1):4

5. Tadesse T, Demissie M, Berhane Y, Kebede Y, Abebe M. The clustering of smear-positive tuberculosis in Dabat, Ethiopia: a population based cross sectional study. PloS one. 2013;8(5):e65022

6. Munch Z, Van Lill SW, Booysen CN, Zietsman HL, Enarson DA, Beyers N. Tuberculosis transmission patterns in a high-incidence area: a spatial analysis. The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease. 2003 Mar;7(3):271-7

7. MacPherson P, Choko AT, Webb EL, Thindwa D, Squire SB, Sambakunsi R, et al. Development and validation of a global positioning system-based "map book" system for categorizing cluster residency status of community members living in high-density urban slums in Blantyre, Malawi. American journal of epidemiology. 2013 May 15;177(10):1143-7

8. WHO country Profile, Malawi: World Health Organization; [cited 2014 26th February]. Available from: https://extranet.who.int/sree/Reports?op=Replet&name=%2FWHO_HQ_Reports%2FG2%2FPROD%2FEXT%2FTBCountryProfile&ISO2=MW&LAN=EN&outtype=html.

9. Nyirenda T. Epidemiology of Tuberculosis in Malawi. The Epidemiology of Malawi. 2nd ed. Geubbels E, Bowie C 2009.

10. 2008 Population and Housing Census: National Statistics Office of Malawi. Available from: http://www.nsomalawi.mw/index.php/2008-population-and-housing-census.html.

11. Population Projections for Malawi: National Statistical Office of Malawi; 1998. Available from: http://www.nsomalawi.mw/index.php/publications/134-population-projections-for-malawi.html.

12. Malawi Demographic an Health Survey, 2010. In: Office MNS, editor. 2011.

13. Kirkwood B, Sterne J. Essential Medical Statistics. Second ed: Blackwell Publishing; 2003.

APPENDICES

Appendix A – Health centres anticipated to participate

Health centres anticipated to participate are:

BTBAH - Blantyre Adventist

BTBG - Bangwe

BTCH - Chilomoni

BTCW - Chitawira

BTMB - Mulambe

BTMW - Mwaiwathu

BTND - Ndirande

BTQE - Queen Elizabeth Central Hospital

BTSL - South Lunzu

BTZG – Zingwangwa

Appendix B – Verbal consent form and information sheet

Informed Consent Form (Verbal)

The TB-ePAL Study: Clinic-based electronic capture of TB case residence
Rebecca Harris

TBP01: Verbal consent for ePAL collection of place of residence data

This is a research project looking at the spread of TB in Blantyre by asking patients to identify where they live on a map.

The Government of Malawi records the address of all TB patients as part of providing care, but spoken addresses are not very accurate. We have developed a new electronic tool that we hope will more accurately record your location of residence. We are interested to use this tool to understand the spread of TB cases in Blantyre.

Your participation will help us to collect this data on TB cases in Blantyre. You do not have to participate if you do not want to, we will only record this information for those people who agree to it. Your information will be kept confidential. We will use a study number and not your name to store the results in a computer file.

There are no costs and no individual benefits or harms for taking part in this study.

If you agree to take part I will ask you some questions about yourself, such as age, gender, and HIV status. The TB officer will then help you find where you live on a map of Blantyre that can be shown on the screen of the computer.

If you want more information, please read the information leaflet supplied or ask one of the team to read it to you.

Do you have any questions about the research study?

Do you agree to provide details about yourself and your place of residence for this research?

Malawi-Liverpool-Wellcome Trust
Clinical Research Programme
Queen Elizabeth Central Hospital
Blantyre

Version 1.0; 26 Feb 2014

Tel: 01876444
Fax: 01875774
www.mlw.medcol.mw

Chilolezo cha kutengambali pa kafukufuku, chopelekedwapolankhula

Kafukufukuwa TB-ePAL: Katengedwekamakonokakokhalakwaodwala TB, kochitikirakuchipatala.
Rebecca Harris
TBP01: Chilolezochotengelamalookhala, chopelekedwapolankhula

Tikuchita kafukufuku amene akuwunika kufala kwa matenda a TB mu Blantyre. Tikuchita izi pakuwafunsa odwala matenda wa kuti atiwuze kumene amakhala pogwilitsa ntchito mapu.

Boma la Malawi limalemba malo okhala anthu onse omwe akudwala matenda a TB, ngati gawo la ndondomeko yosamalira odwala, koma kupeleka malo okhala pogwilitsa ntchito kulankhula sikukhala kolondola kwenikweni. Tapanga chipangizo chatsopano chamakono chimene tikukhulupilira kuti chithandiza kulemba malo anu okhala molondola kwambiri.Tikufuna kuti tigwilitse ntchito chipangizo chimenechi pofuna kudziwa kafalidwe kamatenda a TB mu Blantyre.

Kutenga mbali kwanu pa kafukufukuyu kutithandiza ketenga nkhani zokhudzamatenda a TB mu Blantyre. Simuli owumilizidwa kutengapo mbali pa kafukufukuyu. Nkhani zamatenda a TB tizitenga kwa anthu okhawo omwe avomela kutenga mbali pa kafukufukuyu. Nkhani zokhudza inu zomwe zitengedwe kudzela mu kafukufukuyu zikasungidwa mwa chinsinsi. Sitikagwilitsa ntchito dzina lanu posunga zotsatila zakafukufukuyu mu makina a computer. M'malo mwa dzina lanu, tikagwilitsa ntchito nambala yakafukufuku posunga zotsatilazo.

Simupeleka kanthu kalikonse kuti mutenge nawo mbali pa kafukufukuyu. Palibe phindu la inu nokha lomwe mupeze chifukwa chotenga mbali pa kafukufukuyu komanso kutenga nawo mbali pa kafukufukuyu sikungakuyikeni pa ziwopsyezo.

Ngati mulore kutenga mbali pa kafukufukuyu, ndikufunsani mafunso okhudza inuyo. Mafunsowa akhala ofuna kudziwa zaka zanu, ngati muli wamkazi kapena wamwamuna komanso ngati muli ndi kachilombo ka HIV kapena ayi. Tikatha izi, ogwila ntchito za TB (TB Officer) akuthandizani kuloza komwe mumakhala mu Mzinda wa Blantyre, pogwilitsa ntchito mapu omwe atha kuwonetsedwa pa galasi la computer.

Ngatimukufuna kudziwa zambiri za kafukufukuyu chonde welengani chikalata chomwe chapelekedwa chomwe chili ndi mbiri yakafukufukuyu. Muthakupempham'modzi mwa anthu opanga kafukufukuyu kuti akuwelengeleni.

Muli ndi mafunso aliwonse okhudza kafukufukuyu?

Kodi mukuvomela kutifotokozera zinthu zokhudza inu komanso kutidziwitsa kumene mumakhala, pothandizila kafukufukuyu?

Patient Information Sheet

The TB-ePAL Study: Clinic-based electronic capture of TB case residence

Rebecca Harris

TBP01: Verbal consent for ePAL collection of place of residence data

Why are we doing this study?
This is a research project looking at the spread of TB in Blantyre by asking patients to identify where they live on a map. The Government of Malawi records the address of all TB patients as part of providing care, but spoken addresses are not very accurate. We have developed a new electronic tool that we hope will more accurately record your location of residence. We are interested to use this tool to understand the spread of TB cases in Blantyre.

Why are we asking you to take part in this study?
Your participation will help us to collect this data on TB cases in Blantyre.

What will happen if I decide to take part in the study?
If you agree to take part I will ask you some questions about yourself, such as age, gender, and HIV status. The TB officer will then help you find where you live on a map of Blantyre that can be shown on the screen of the computer.

Who are we asking to participate?
All patients aged 18 years and above initiating TB treatment in the participating health clinics in Blantyre.

Confidentiality
All information obtained during the study will be held securely and stored on paper and computer files. The PI, Rebecca Harris, will take responsibility for keeping your information confidential. The computer files will use a code instead of your name to ensure that no one can identify you.

What are the risks and benefits of the study?
Taking part in the study will not cost you anything. There are no direct risks from taking part in this study. There are no direct individual benefits to taking part in this study; however, you will be contributing to the development of better tools for follow up and support of people with TB in Blantyre.

Voluntary participation and withdrawal
Your participation is voluntary. You may withdraw from the study at any time without giving a reason and without affecting your usual health care.

Where do we come from?
We work at the Malawi-Liverpool-Wellcome Trust (MLW) Clinical Research programme and the London School of Hygiene and Tropical Medicine. The MLW programme conducts research on diseases of local importance to Malawi and the region.

The Ethics Committees that have approved the study are:
The College of Medicine Research Ethics Committee (COMREC), Blantyre, Malawi, and the London School of Hygiene and Tropical Medicine, London, UK.

Questions
If you have any questions concerning participation in this study, please feel free to ask me. If you think of any questions after you have left, you can contact: Rebecca Harris (PI) or the COMREC secretariat, telephone: 01876444 or 01871911.

Chikalata cha zomwe odwala ayenela kudziwa

Kafukufuku wa eTB-PAL:Katengedwe ka makono ka kokhala kwa odwala TB, kochitikira kuchipatala.
Rebecca Harris
TBP01: Chilolezo chotengela malo okhala, chopelekedwa po lankhula

Chifukwa chiyani tikuchita kafukufukuyu?

Tikuchita kafukufuku ameneyu pofuna kuwunika kufala kwa matenda a TB mu Blantyre. Tikuchita izi pakuwafunsa odwala matendawa kuti atiwuze kumene amakhala pogwilitsa ntchito mapu. Boma la Malawi limalemba malo okhala anthu onse omwe akudwala matenda a TB, ngati gawo la ndondomeko yosamalira odwala, koma kupeleka malo okhala pogwilitsa ntchito kulankhula sikukhala kolondola kwenikweni. Tapanga chipangizo chatsopano chamakono chimene tikukhulupilira kuti chithandiza kulemba malo anu okhala molondola kwambiri. Tikufuna kuti tigwilitse ntchito chipangizo chimenechi pofuna kudziwa kafalidwe ka matenda a TB mu Blantyre.

Chifukwa chiyani tikukupemphani kuti mutengeko mbali pa kafukufukuameneyu?

Kutenga mbali kwanu pa kafukufukuyu kutithandiza ketenga nkhani zokhudza matenda a TB mu Blantyre.

Chichitike ndi chiyani ngati nditasankha kuti nditengepo mbali pa kafukufukuyu?

Ngati mulore kutenga mbali pa kafukufukuyu, ndikufunsani mafunso okhudza inuyo. Mafunsowa akhala ofuna kudziwa zaka zanu, ngati muli wa mkazi kapena wa mwamuna komanso ngati muli ndika chilombo ka HIV kapena ayi. Tikatha izi, ogwila ntchito za TB (TB Officer) akuthandizani kuloza komwe mumakhala mu Mzinda wa Blantyre, pogwilitsa ntchito mapu omwe atha kuwonetsedwa pa galasi la computer.

Kodi omwe tikuwapempha kuti atengepo mbali pa kafukufukuyu ndi ndani?

Odwala onse omwe ali ndi zaka khumi ndi zisanu ndi zitatu (18) kunka mtsogolo omwe akuyamba kulandila mankhwala a TB mu zipatala za ku Blantyre zomwezi kutengapo gawo pa kafukufukuyu.

Kusunga chinsisi

Nkhani zonse zomwe zitengedwe kudzela mu kafukufukuyu zikasungidwa mwa chitetezo pa mapepala komanso mu computer. Otsogolera kafukufukuyu, a Rebecca Harris, awonetsetsa kuti nkhani zonse zokhudza inu zisungidwe mwachinsisi. Nkhani zanu zomwe zisungidwe mu computer sizikhala ndi dzina lanu: tigwilitsa ntchito chizindikilo m'malo mwa dzina lanu kuti wina aliyense asathe kudziwa kuti nkhanizo ndi zokhudza inu.

Kodi kuwopsya kotenga mbali mu kafukufukuyu ndikuti? Nanga ubwino wotenga nawo mbali pa kafukufukuyu ndiwanji?

Simupeleka kanthu kalikonse kuti mutengenawo mbali pa kafukufukuyu. Kutenga nawo mbali pa kafukufukuyu sikungakuyikeni pa ziwopsyezo zina ziri zonse zodziwika. Palibe phindu la inu nokha lomwe mupeze potenga mbali pa kafukufukuyu komabe kutenga nawo mbali kwanu pa kafukufukuyu kuthandizira kupanga ndondomeko zabwino zolondolera komanso kuthandiza anthu omwe akudwala matenda a TB mu Blantyre.

Kutenga mbali ndi kusiya kutenga mbali mosakakamizidwa

Kutenga mbali kwanu mu kafukufukuyu ndikosakakamiza. Mutha kuleka kutenga nawo mbali pa kafukufukuyu pa nthawi iliyonse opanda kupeleka chifukwa cholekela. Iz isizingasokoneze thandizo la zaumoyo lomwe mumalandila nthawi zonse.

Kodi omwe akupanga kafukufukuyu akuchokela kuti?

Omwe akupanga kafukufukuyu amagwila ntchito ku Malawi-Liverpool-Wellcome Trust (MLW) Clinical Research Programme ndiponso ku Sukulu ya Ukhondo ndi Mankhwala ya ku London. Bungwe la MLW limapanga kafukufuku wamatenda, othandiza ku dziko la Malawi komanso mayiko ozungulira Malawi.

Magulu audindo omwe aloleza kuti kafukufukuyu achitike ndi awa:

Gulu lowona kufunikila kwa kafukufuku la ku Chipatala cha Mankhwalaku Blantyre m'dziko la Malawi komanso Sukulu ya Ukhondo ndi Mankhwala yaku London m'dziko la UK.

Mafunso

Ngati muli ndi mafunso okhudza kutenga mbali kwanu mu kafukufukuyu, chonde khalani omasuka kundifunsa. Ngati mutaganizila mafunso ena aliwonse titasiyana, mutha kulumikizana ndi otsogolera kafukufukuyu a Rebecca Harris kapena a Dr V. Mwapasa omwe ndi wa pampando wa gulu lowona kufunikila kwa kafukufuku la ku Chipatala cha Mankhwala ku Blantyre. Muthakugwilitsa ntchito ma nambala awa 01876444 ndi 01871911 poyimba lamya ndikulankhula ndi a Rebecca Harris (PI) kapena aku COMREC Secretariat.

Appendix C – Written consent form and information sheet for evaluation group

Patient Information Sheet

The TB-ePAL Study: Clinic-based electronic capture of TB case residence

Rebecca Harris

TBP02: Evaluation component of using ePAL to locate residence of patients initiating TB treatment

Why are we doing this study?
This is a research project that we hope will help us to develop a tool to better understand the spread of TB in Blantyre. The Government of Malawi records the address of all TB patients as part of providing care, but spoken addresses are not very accurate. We want to see if a new electronic map will be more accurate. In this study we want to compare the location identified through the map with the exact location of your home.

Why are we asking you to take part in this study?
Your participation will help us to see if the new map tool is accurate enough for use in Blantyre.

What will happen if I decide to take part in the study?
First the TB officer will ask some questions to record information about you on a small computer, including all of the information recorded routinely by the TB services, and some additional information for the study about the your household size and income. The TB officer will then help you find where you live on a map of Blantyre that can be shown on the screen of the computer. We will then arrange a home visit where we will record the exact location. If you agree, we can also provide TB screening for your household members as recommended by the Government of Malawi.

Who are we asking to participate?
All patients aged 18 years and above initiating TB treatment in the participating health clinics in Blantyre.

Confidentiality
All information obtained during the study will be held securely and stored on paper and computer files. The PI, Rebecca Harris, will take responsibility for keeping your information confidential. The computer files will use a code instead of your name to ensure that no one can identify you.

What are the risks and benefits of the study?
Taking part in the study will not cost you anything and we will provide 1000 MK to compensate you for your time. There are no direct risks from taking part in this study, however there is a risk that if people find out about your TB that you could be treated badly because of this, or they may make assumptions about your HIV status because TB is sometimes linked to HIV. There are no direct individual benefits to taking part in this study; however, you will be contributing to the development of better tools for follow up and support of people with TB in Blantyre.

Voluntary participation and withdrawal
Your participation is voluntary. You may withdraw from the study at any time without giving a reason and without affecting your usual health care.

Where do we come from?
We work at the Malawi-Liverpool-Wellcome Trust (MLW) Clinical Research programme and the London School of Hygiene and Tropical Medicine. The MLW programme conducts research on diseases of local importance to Malawi and the region.

The Ethics Committees that have approved the study are:
The College of Medicine Research Ethics Committee, Blantyre, Malawi, and the London School of Hygiene and Tropical Medicine, London, UK.

Questions
If you have any questions concerning participation in this study, please feel free to ask me. If you think of any questions after you have left, you can contact: Rebecca Harris (PI) or the COMREC secretariat, telephone: 01876444 or 01871911.

Malawi-Liverpool-Wellcome Trust
Clinical Research Programme
Queen Elizabeth Central Hospital, Blantyre

Version 1.0; 24 Feb 2014

Tel: 01876444
Fax: 01875774
www.mlw.medcol.mw

Chikalata cha zomwe odwala ayenela kudziwa

Kafukufukuwa TB-ePAL: Katengedwe ka makono ka kokhala kwa odwala TB,kochitikira kuchipatala.

Rebecca Harris

Kawuniwuni wa kugwilitsa ntchito ePAL pofuna kudziwa komwe odwala omwe akuyamba kulandila mankhwala a TB amakhala.

Chifukwa chiyani tikuchita kafukufukuyu?

Uyu ndi kafukufuku amene tikukhulupilira kuti atithandiza kupanga ndondomeko yothandizira kumvetsetsa kufala kwa matenda a TB mu Blantyre. Boma la Malawi limalemba malo okhala anthu onse omwe akudwala matenda a TB, ngati gawo la ndondomeko yosamalira odwala, koma kupeleka malo okhala pogwilitsa ntchito kulankhula sikukhala kolondola kwenikweni.Tikufuna tiwone ngati mapu atsopano, ogwilitsa ntchito njira za makono, ati akhale olondola kwambiri. Mukafukufukuyu, tikufuna tifanizire malo omwe tidziwapeza pogwilitsa ntchito mapundi malo enieni omwe mumakhala.

Chifukwa chiyani tikukupemphani kuti mutengeko mbali pa kafukufukuyu?

Kutenga mbali kwanu kutithandiza kuti tidziwe ngati mapu atsopanowa ndiwolondola bwino lomwe kut itiwangwilitse ntchito mu Blantyre.

Chichitike ndi chaani ngati nditasankha kuti nditengepo mbali pa kafukufukuyu?
Poyamba, ogwila ntchito za TB, akufunsani mafunso kuti alembe zokhudzana ndiinu pa computer yaying'ono. Zomwe zitalembedwezi ziphatikizira kuzomwe iwo amalemba nthawi zonse zokhudza TB komanso zina zowonjezela zokhudza kukula kwa banja lanu ndi chuma chanu zomwe zikufunikira pa kafukufukuyu. Atatha izi, ogwila ntchito za TB, akuthandizani kuti muloze komwe mumakhala pogwilitsa ntchito mapu omwe awonetsedwe pa computer. Kenako, tipanga ndondomeko yoti tipite kunyumba kwanu kuti tikathe kulemba malo enieni omwe mumakhala. Ngati mungavomele, titha kukayeza anthu a kumyumba kwanu kuti tikawone ngati nawonso akudwala matenda a TB: tikachita izi motsatira ndondomeko ya Boma la Malawi pa nkhani ya matenda a TB.

Kodi omwe tikuwapempha kuti atengepo mbali pa kafukufukuyu ndi ndani?

Odwala onse omwe ali ndi zaka khumi ndi zisanu ndi zitatu (18) kunka mtsogolo omwe akuyamba kulandila mankhwala a TB mu zipatala za mu Blantyre zomwe zikutengapo gawo pa kafukufukuyu.

Kusunga chinsisi

Nkhani zonse zomwe zitengedwe kudzela mu kafukufukuyu zikasungidwa mwa chitetezo pa ma pepala komanso mu computer. Otsogolera kafukufukuyu, a Rebecca Harris, awonetsetsa kuti nkhani zonse zokhudza inu zisungidwe mwachinsisi. Nkhani zanu zomwe zisungidwe mu computer sizikhala ndi dzina lanu: tigwilitsa chizindikilo m'malo mwa dzina lanu kuti wina aliyense asathe kudziwa kuti nkhanizo ndi zokhudzainu.

Kodi kuwopsya kotenga mbali mu kafukufukuyu ndikuti , nanga ubwino wotenga nawo mbali pa kafukufukuyu ndi wanji?

Simupeleka kanthu kalikonse kuti mutenge nawo mbali pa kafukufukuyu ndipo ife tikupatsani ndalama zokwana MK1000 pokuthokozani kamba kopatula nthawi yanuo kuti mutengepo mbali pa kafukufukuyu. Kutenga nawo mbali pa kafukufukuyu sikungakuyikeni pa ziwopsyezo zina zirizonse zodziwikiratu , Komabe, pali chiwopsyezo choti ngati anthu atadziwa kuti mukudwala matenda a TB, atha osakusamalan ibwino komanso anthuwa atha kukuganizila kuti muli ndikachilombo ka HIV poti nthawi zina pamakhala ubale pakati pa matenda a TB ndi kachilombo ka HIV. Palibe phindu la inu nokha lomwe mupeze potenga mbali pa kafukufukuyu komabe kutenga nawo mbali kwanu pa kafukufukuyu kuthandizira kupanga ndondomeko zabwino zolondolera komanso kuthandiza anthu omwe akudwala matenda a TB mu Blantyre.

Kutenga mbali ndikusiya kutenga mbali mosakakamizidwa

Kutengambali kwanu mu kafukufukuyu ndikosakakamiza. Mutha kuleka kutenga nawo mbali pa kafukufukuyu pa nthawi iliyonse opanda kupeleka chifukwa cholekela. Izi sizingasokoneze thandizo la zaumoyo lomwe mumalandila nthawi zonse.

Kodi omwe akupanga kafukufukuyu akuchokela kuti?

Omwe akupanga kafukufukuyu amagwila ntchito ku Malawi-Liverpool-Wellcome Trust (MLW) Clinical Research Programme ndiponso ku Sukulu ya Ukhondo ndi Mankhwala yaku London. Bungwe la MLW limapanga kafukufuku wa matenda, othandiza kudziko la Malawi komanso mayiko ozungulira Malawi.

Magulu audindo omwe aloleza kuti kafukufukuyu achitike ndi awa:

Gulu lowona kufunikila kwa kafukufuku la ku Chipatala cha Mankhwala ku Blantyre m'dziko la Malawi komanso Sukulu ya Ukhondo ndi Mankhwala yaku London m'dziko la UK.

Mafunso

Ngati muli ndi mafunso okhudza kutenga mbali kwanu mu kafukufukuyu, chonde khalani omasuka kundifunsa. Ngati mutaganizila mafunso ena aliwonse titasiyana, muthakulumikizana ndi otsogolera kafukufukuyu a Rebecca Harris kapena a Dr V. Mwapasa omwe ndi wapampando wagulu lowona kufunikila kwa kafukufuku la ku Chipatala cha Mankhwala ku Blantyre. Muthakugwilitsa ntchito ma nambala awa 01876444 ndi 01871911 poyimba lamya ndikulankhula ndi a Rebecca Harris (PI) kapena aku COMREC Secretariat.

Informed Consent Form

The TB-ePAL Study: Clinic-based electronic capture of TB case residence

Rebecca Harris

TBP02: Evaluation component of using ePAL to locate residence of patients initiating TB treatment - Study participant home visit consent form

Please sign below if you agree with the following statements:

1. I have read or had read the information sheet about the study, which explains what you are trying to find out and why you would like to talk to me.
2. I have discussed and understood the purpose of the research, and of the electronic tool and home visit to record my home address.
3. I have asked all the questions that I have about the purpose of the research, the electronic tool and home visit and feel happy that I have enough information about it.
4. I understand that, by completing this consent sheet, I am agreeing to allow a member of the research team to contact me by telephone to arrange a visit to my home.
5. I understand that you will keep the information I give you confidential and not use names linked to the information.
6. I know that my TB diagnosis and HIV status will not be shared with anyone during the home visit unless I agree to this.
7. If I do not agree to take part in this study, I understand that it will not affect my right to TB care at the health centre now or in the future.
8. I understand that I can choose not to continue participation in the study at any time without having to give a reason for this and without affecting my usual health care.

You will be given a copy of the information sheet and a signed consent form to keep.

I voluntarily agree to take part in this study.

Participant _____ TB Officer _____
(Name in BLOCK CAPITALS) (Name in BLOCK CAPITALS)

_____ _____
Signature or thumbprint Signature

_____ _____
Date Date

If the participant gave verbal consent, please enter the name of person who witnessed the consent here, and their signature:

Witness _____ _____
(Name in BLOCK CAPITALS) Date

_____
Signature

Chikalatachovomelezakuchitanawokafukufuku.
Kafukufukuwa TB-ePAL: Katengedwekamakonokakokhalakwaodwala TB, kochitikirakuchipatala.

Rebecca Harris

KawuniwuniwakugwilitsantchitoePALpofunakudziwakomweodwalaomweakuyambakulandilamankhwala a TB amakhala. **- Chikalata cha wotengambali pa kafukufuku, chovomelezakuyendeledwakumyumbakwawo.**

Chondelembanisiginetchayanupamusipangatimukugwilizanandizomwezalembedwapansipa.

1. Ndawelengakapenandinawelengeledwa mbiriyakafukufukuyu, imeneikulongosolazimenemukufunakupezakomansochifukwachaani mukufunakulankhulanane.
2. Takambitsanandipondamvetsetsazolingazakafukufukuyu, njirayamakonoyakalondolondowamapukomansokupitakunyumbakwangakutimukalembezothandizilakudziwandikufikakwathuko.
3. Ndafunsamafunsoonseomwendinalinawookhudzacholinga cha kafukufukuyu, njirayamakonoyakalondolondowamapukomansondondomekoyotimubwelekunyumbakwangandipondiliokondwakutindadziwaizi.
4. Kutikulemba pa pepala lino kukutanthauzakutindavomelakutim'modziwaanthu a kafukufukuathakundiyimbilalamyakutitipangendondomekoyotiadzabwelekumyumbakwanga .
5. Ndamvetsetsakutimusungazimenendikuwuzenimwachinsinsindiposimugwilitsanthitomayinaaanthumokhudzanandizomwemulembe.
6. Ndikudziwa kuti zotsatila za kuyezedwa kwanga kwa TB komanso HIV sizidzawululidwa kwa wina aliyense panthawi imene ndayendeledwa kunyumba, pokhapokha nditavomeleza kuti ziwululidwe.
7. Ndikumvetsetsakutingatindikanakutenganawombali pa kafukufukuuyu, sizikhudzaufuluwangawopezachithandizo cha matenda a TB kuchipatala,tsopanokapenam'tsogolo.
8. Ndikumvetsakutindithakusankhakulekakutenganawombali pa kafukufukuyu pa nthawiiliyonseopandakupelekachifukwacholekelandipoizisizingasokonezethandizo la zaumoyolomwendimalandilanthawizonse.

Mupatsibwa chikalatachomwe chili ndimbiriyakafukufukuyikomansomupatsidwachikalatachosayinidwachosonyezachilolezokutimwatenga nawombali pa kafukufukuyu.Ndikuvomeleza, mosakakamizidwa, kutinditengenawombali pa kafukufukuyu.

Wotengagawo pa kafukufuku_____
(Dzinalolembedwa mu malembaakuluakulu)

_____Tsiku_____
Sayinikapenachidindo cha chala

_____

Ogwilantchitoya TB_____ _____
(Dzinalolembedwa mu malembaakuluakulu)

Sayini_____ Tsiku_____

Ngati otenga mbali pa kafukufuku walola kutenga nawo mbali mu kafukufuku pongolankhula chabe, chonde lembani dzina la munthu amene anayikila izi umboni. Dzina la mbonindiponsosayiniyawozilembedwem'musimu:

Mboni _____ _____Tsiku _____

(Dzina lolembedwa mu malemba akulu akulu)

_____

Sayini

Appendix D: Permissions for reproduction of figures

From: **DYE, Christopher M.** dye
Subject: RE: Permission for use of figure
Date: 24 July 2017 20:29
To: Rebecca Harris Rebecca.Harris@lshtm.ac.uk

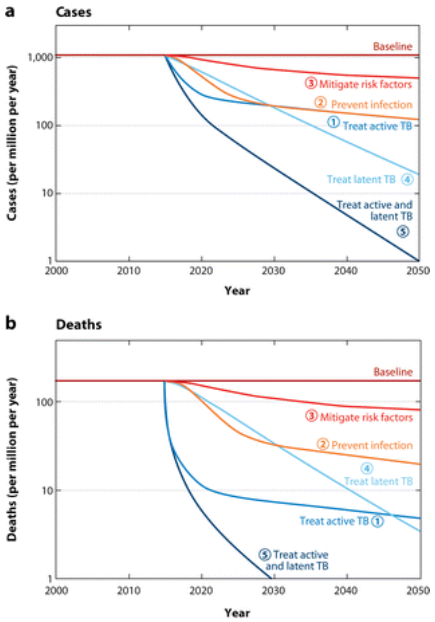
Rebecca, you are welcome to reproduce the Figures in your thesis.
Best regards… Chris

From: Rebecca Harris [mailto:Rebecca.Harris@lshtm.ac.uk]
Sent: 24 July 2017 18:40
To: DYE, Christopher M.
Subject: Re: Permission for use of figure

Dear Chris,

Apologies to email against soon, but was wondering if I could also ask for permission to use the below image from your 2013 paper for the same purpose?

Many thanks,
Rebecca



Dye C, et al. 2013.
Annu. Rev. Public Health. 34:271–86

Dear Chris,

I hope this email finds you well.

I was wondering whether I could seek permission for including the following image from your 2013 prospects for TB elimination paper within the printed and electronic versions of my thesis (for submission to the London School of Hygiene and Tropical Medicine). The thesis will be made available on the the LSHTM Research Online institutional repository, which is non-commercial and available to all: http://researchonlint.lshtm.ac.uk/

If you are not the rights holder for this material, I would be grateful if you could advise me who to contact.

Many thanks in advance,

Rebecca

From: permissions@who.int
Subject: ID: 233421 Permission authorization for WHO copyrighted material
Date: 24 July 2017 13:20
To: rebecca.harris@lshtm.ac.uk
Cc: permissions@who.int

Dear Ms Harris

Thank you for your request for permission to reproduce, reprint or translate certain WHO copyrighted material.

On behalf of the World Health Organization, we are pleased to authorize your request to reproduce the WHO materials as detailed in the form below, subject to the terms and conditions of the non-exclusive licence below.

If you have questions regarding this authorization, please contact permissions@who.int.

We thank you for your interest in WHO published materials.

Kind regards,
WHO Permissions team

WORLD HEALTH ORGANIZATION (WHO)

Non-exclusive licence to use selected WHO published materials

You submitted a request, through WHO's online platform, for permission to reprint and reproduce certain WHO copyrighted material (the "Licensed Materials"). This is a legal agreement (the "Agreement") between you and WHO, granting you a licence to use the Licensed Materials subject to the terms and conditions herein.

Read this Agreement in its entirety before using the Licensed Materials.

By using the Licensed Materials, you enter into, and agree to be bound by, this Agreement.

This licence is granted only for original materials belonging to WHO. If any part of the WHO published materials you wish to reproduce are credited by WHO to a source other than WHO, those materials are not covered by this Agreement and are not part of the Licensed Materials. You are responsible for determining if this is the case, and if so, you are responsible for obtaining any necessary permission from the source of those third-party materials prior to their use.

If you enter into this Agreement on behalf of an organization, by using the Licensed Materials you confirm (represent and warrant) that you are authorized by your organization to enter into this Agreement on the organization's behalf. In such a case, the terms "you" and "your" in this Agreement refer to, and this Agreement applies to, the organization.

WHO grants this licence to you based on the representations and warranties you made in the licence request you submitted through WHO's online platform. If any of those representations and/or warranties are or become false or inaccurate, this licence agreement shall automatically terminate with immediate effect, without prejudice to any other remedies which WHO may have.

If you have questions regarding this Agreement, please contact permissions@who.int.

1. Licence. Subject to the terms and Conditions of this Agreement, WHO grants to you a worldwide, royalty free, non-transferable, non-sublicensable, non-exclusive licence to use, reproduce, publish, and display the Licensed Materials in the manner and using the media indicated in the Permissions Request Form you submitted to WHO (the "Licensed Use"). This licence is limited to the current edition of your publication. Future editions or a different use of the Licensed Materials will require additional permission from WHO. If your request includes translation into different languages, then non-exclusive permission is hereby granted to translate the Licensed Materials into the languages indicated.

2. Retained Rights. Copyright in the Licensed Materials remains vested in WHO, and WHO retains all rights not specifically granted under this Agreement.

3. Mandatory Acknowledgement. In every instance of the Licensed Use, you must make suitable acknowledgement of WHO, either as a footnote or in a reference list at the end of your publication, as follows:

"Reprinted from Publication title, Vol /edition number, Author(s), Title of article / title of chapter, Pages No., Copyright (Year)."

In addition, If the Licensed Materials originate from the WHO web site, you must also include the URL reference and the date accessed.

Translations of the Licensed Materials should be attributed as follows:

"Translated with permission of the publisher from Publication title, Vol /edition number, Author(s), Title of article / title of chapter, Pages No., Year."

4. Altering or Modifying the Licensed Materials. As part of the Licensed Use, you may minimally alter or adapt figures and tables in the Licensed Materials to match the style of your publication. Any other alteration or modification of the Licensed Materials (including abbreviations, additions, or deletions) may be made only with the prior written authorization of WHO.

5. Appropriate and Prohibited Uses. You must use the Licensed Materials in a factual and appropriate context. You may not use the Licensed Materials in association with any product marketing, promotional, or commercial activities, including, without limitation, in advertisements, product brochures, company-sponsored web sites, annual reports, or other non-educational publications or distributions.

6. No WHO endorsement. You shall not state or imply that WHO endorses or is affiliated with your publication or the Licensed Use, or that WHO endorses any entity, organization, company, or product.

7. No use of the WHO logo. In no case shall you use the WHO name or emblem, or any abbreviation thereof. Notwithstanding the foregoing, if the WHO name and/or emblem appear as an integral part of the Licensed Materials (e.g. on a map) you may use the name and/or emblem in your use of the License Materials, provided the name and/or logo is not used separately from the Licensed Materials.

8. No Warranties by WHO. All reasonable precautions have been taken by WHO to verify the information contained in the Licensed Materials. However, WHO provides the Licensed Materials to you without warranty of any kind, either expressed or implied, and you are entirely responsible for your use of the Licensed Materials. In no event shall WHO be liable for damages arising from your use of the Licensed Materials.

9. Your Indemnification of WHO. You agree to indemnify WHO for, and hold WHO harmless against, any claim for damages, losses, and/or any costs, including attorneys' fees, arising in any manner whatsoever from your use of the Licensed Materials or for your breach of any of the terms of this Agreement.

10. Termination. The licence and the rights granted under this Agreement shall terminate automatically upon any breach by you of the terms of this Agreement. Further, WHO may terminate this licence at any time with immediate effect for any reason by written notice to you.

11. Entire Agreement, Amendment. This Agreement is the entire agreement between you and WHO with respect to its subject matter. WHO is not bound by any additional terms that may appear in any communication from you. This Agreement may only be amended by mutual written agreement of you and WHO.

12. Headings. Paragraph headings in this Agreement are for reference only.

13. Dispute resolution. Any dispute relating to the interpretation or application of this Agreement shall, unless amicably settled, be subject to conciliation. In the event of failure of the latter, the dispute shall be settled by arbitration. The arbitration shall be conducted in accordance with the modalities to be agreed upon by the parties or, in the absence of agreement, with the rules of arbitration of the International Chamber of Commerce. The parties shall accept the arbitral award as final.

14. Privileges and immunities. Nothing in or relating to this Agreement shall be deemed a waiver of any of the privileges and immunities enjoyed by WHO under national or international law and/or as submitting WHO to any national court jurisdiction.

***

DataCol Web: Form for requesting permission to reproduce, reprint or translate WHO copyrighted material
==========
ID: 233421

Section: Contact details
---
* Title
* Ms
---
* First name
* Rebecca
---
* Family name
* Harris
---
* Organization/affiliation
* LSHTM
---
* Web site address
*
---
* Type of organization
* University/Academic
---

* If other, please specify
*
---
* If STM signatory, please select
---
* Position
* PhD student
---
* Telephone
* +447944301012
---
* Address
* LSHTM, Keppel Street, London, WC1E 7HT
---
* Country
* United Kingdom of Great Britain and Northern Ireland
---
* Email
* rebecca.harris@lshtm.ac.uk

Section: Information about WHO material to be reproduced
---
* Full title of WHO material from which the reproduction is to be made
* TB strategy targets gif from the End TB strategy website
---
* Website URL where WHO material is published
* http://www.who.int/entity/tb/strategy/tb-strategy-targets.gif
---
* ISBN / WHO Reference Number
*
---
* Please select the item(s) to be reproduced
* Figure/table
---
* Type of reuse
* Dissertation or thesis
---
* No of item(s) to be reproduced
* 5 items or less
---
* For each item, please provide a reference and page number. If entire document, please state "Entire document".
* published on end TB strategy website: http://www.who.int/tb/strategy/end-tb/en/

Section: Information about your publication
---
* Please provide the title of your publication that the above materials are to be published in
* Informing development strategies for new TB vaccines: mathematical modelling and novel epidemiological tools (PhD thesis)
---
* Publishing format
* Print, HTML
---
* Will you be translating?
* No
---
* If yes, please indicate languages
*
---
* If web please provide URL / If other, please specify
* Will be posted on the LSHTM thesis repository
---
* Number of copies (if applicable)
*
---
* Target audience and planned distribution
* PhD examiners and scientific community
---
* Planned publication/distribution date
* End 2017
---
* If your publication or the material is to be sold, indicate the planned selling price or subscription fee
* n/a
---
* Is your publication sponsored or funded by an organisation other than your own?
* No
---
* If yes, please provide additional information
*

---
* Will there be any advertising associated with your publication?
* No
---
* If yes, please provide additional information
*
---
* Subject(s) of interest that most correspond to your request
* Tuberculosis (TB)
---
* Additional information about your request
*
---
* Approval
* Auto permission
---
* Latest approval modification
---
* WHO Department
* ACP, ACT
---
* Correct WHO URL
* http://www.who.int/entity/tb/strategy/tb-strategy-targets.gif

Section: Terms and conditions
---
* By submitting this request you confirm that you will abide by the terms and conditions if WHO grants you permission.
* I have read and agree with the terms and conditions

Click the following link to access a format view of this record:
http://apps.who.int/datacol/survey.asp?survey_id=258&respondent_id=233421

This email was automatically sent to you by the WHO Intranet Data Collector.
The DataCol can send emails to accounts specified by the Form focalpoint.
