Optimizing treatment for tuberculosis infection and multidrug-resistant tuberculosis

Nicholas Winters

Department of Epidemiology, Biostatistics, and Occupational Health

Faculty of Medicine

McGill University, Montreal

June 2023

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Doctor of Philosophy (PhD)

Table of Contents

Front Matter

Abstract

Tuberculosis is one of the oldest documented infectious diseases and to this day remains a substantial global health burden. An estimated one quarter of the global population is living with tuberculosis infection, of which roughly 5-10% will develop tuberculosis disease in their lifetime. Despite advances in tuberculosis diagnosis and treatment, there were 10.6 million cases of tuberculosis disease and nearly 1.6 million deaths reported by the World Health Organization (WHO) in 2021. Of the reported cases of tuberculosis disease, 450,000 were cases of multidrug-resistant tuberculosis (MDR-TB) for which current treatment entails a high patient burden due the long treatment duration and use of less effective drugs than those used for drug-susceptible tuberculosis. Optimal tuberculosis preventive therapy is crucial to meeting the goals of the WHO End-TB strategy of achieving an 80% reduction in tuberculosis incidence by 2030, while more effective and shorter treatment regimens are needed to combat the devastating impact of MDR-TB. The overall aim of this thesis is to improve treatment for tuberculosis infection and MDR-TB.

The first manuscript is network meta-analysis of individual patient data that was used to compare the completion, safety, and efficacy of two tuberculosis preventive treatments: 3 months of weekly rifapentine plus isoniazid (3HP) and 4 months of daily rifampicin (4R). We found that more participants completed a regimen of 3HP than 4R, and the adjusted risk ratios (aRR) and adjusted risk differences (aRD) along with their 95% confidence intervals (CI) for treatment completion were aRR 1.06 (95% CI: 1.02, 1.10) and aRD of 0.05 (95% CI: 0.02, 0.07). However, 3HP was also associated with increased risk of treatment-related grade 3 to 4 adverse events leading to treatment discontinuation than 4R (aRR 3.46

[95% CI: 2.09, 6.17]; aRD 0.02 [95% CI: 0.01, 0.03]). The increase in completion with 3HP over 4R must be weighed against the serious concerns for adverse events.

My second manuscript shifts focus to treatment of MDR-TB. In this study, I applied causal methods to compare the efficacy between two core drug regimens received in addition to other drugs being concurrently prescribed for MDR-TB treatment at the discretion of the provider: all three WHO group A drugs (bedaquiline, linezolid, and a fluoroquinolone) without clofazimine compared to the same three group A drugs plus clofazimine. This study was a target trial emulation using observational data where the core regimens were compared in an intention-to-treat analysis using baseline inverse probability of treatment weights and four per-protocol analyses that used time-varying inverse probability of censoring weights to estimate the average treatment effect (ATE). In this study, we found some evidence that adding clofazimine increases successful treatment outcomes when added to the group A drugs in the intention-to-treat analysis (baseline treatment weighted ATE: 0.07 [95% CI: -0.01, 0.15]) with results being similar in the second and fourth per-protocol analyses. However, the ATE was substantially attenuated in the first and third per-protocol analyses where patients were censored for deviating from their respective treatment strategies with additional censoring of patients in the control group who started clofazimine after baseline treatment assignment: in the first per-protocol analysis the censoring weighted ATE was -0.01 (95% CI: -0.10, 0.07), while in the third per-protocol analysis the censoring weighted ATE was 0.02 (95% CI: -0.05, 0.09).

In my third manuscript, continuing in MDR-TB treatment I used individual patient data from observational studies to identify patient characteristics and treatment factors that are associated with treatment duration. We used a novel outcome to address biases inherent in

studying duration of MDR-TB treatment: the individual's deviation in treatment duration from the mean duration of their treatment site, using only patients with successful outcomes. In this study we showed that bedaquiline use was associated with a 0.51 (95% CI: 0.15, 0.87) month decrease in duration of treatment, which was consistent across subgroups, while MDR-TB with fluoroquinolone resistance was associated with 0.78 (95% CI: 0.36, 1.21) months increase. Our results were consistent with the literature as we showed that many factors known to be associated with poor treatment outcomes were also associated with longer treatment durations. The results of this study may allow future research of shorter MDR-TB treatment to broaden inclusion criteria to patients who have been previously excluded from trials and guidelines.

Improving treatment for all forms of tuberculosis is imperative to meeting the WHO EndTB goals and improving patient outcomes. The work outlined in this thesis provides evidence that current practitioners can draw on for treatment of tuberculosis infection and that future researchers can use to improve treatment regimens for MDR-TB patients.

Résumé

La tuberculose est l'une des plus anciennes maladies infectieuses documentées et reste à ce jour un fardeau important pour la santé mondiale. On estime qu'un quart de la population mondiale vit avec une infection tuberculeuse et qu'environ 5 à 10 % d'entre eux développeront une tuberculose maladie au cours de leur vie. Malgré les progrès réalisés dans le diagnostic et le traitement de la tuberculose, l'Organisation mondiale de la santé (OMS) a recensé 10,6 millions de cas de tuberculose maladie et près de 1,6 million de décès en 2021. Parmi les cas de tuberculose déclarés, 450 000 étaient des cas de tuberculose multirésistante (TB-MR) pour lesquels le traitement actuel représente un lourd fardeau pour les patients en raison de la longue durée du traitement et de l'utilisation de médicaments moins efficaces que ceux utilisés pour la tuberculose sensible aux médicaments. Un traitement préventif optimal de la tuberculose est essentiel pour atteindre les objectifs de la stratégie de lutte contre la tuberculose de l'OMS, qui consiste à réduire de 80 % l'incidence de la tuberculose d'ici 2030, tandis que des régimes de traitement plus efficaces et plus courts sont nécessaires pour lutter contre l'impact dévastateur de la TB-MR. L'objectif global de cette thèse est d'améliorer le traitement de l'infection tuberculeuse et de la TB-MR.

Le premier manuscrit est une méta-analyse en réseau de données individuelles de patients qui a été utilisée pour comparer l'achèvement, la sécurité et l'efficacité de deux traitements préventifs de la tuberculose : 3 mois de rifapentine hebdomadaire plus isoniazide (3HP) et 4 mois de rifampicine quotidienne (4R). Nous avons constaté qu'un plus grand nombre de participants ont suivi un régime de 3HP que de 4R, et les rapports de risque ajustés (RRa) et les différences de risque ajustées (RDa) ainsi que leurs intervalles de confiance à 95 % (IC) pour l'achèvement du traitement étaient un RRa de 1,06 (IC 95 % : 1,02, 1,10) et un RDa de

0,05 (IC 95 % : 0,02, 0,07). Cependant, 3HP a également été associé à un risque plus élevé d'événements indésirables de grade 3 à 4 liés au traitement et conduisant à l'arrêt du traitement que 4R (RRa 3,46 [IC 95 % : 2,09, 6,17] ; RDa 0,02 [IC 95 % : 0,01, 0,03]). L'augmentation du taux d'achèvement avec 3HP par rapport à 4R doit être mise en balance avec les préoccupations sérieuses concernant les événements indésirables.

Mon deuxième manuscrit se concentre sur le traitement de la TB-MR. Dans cette étude, j'ai appliqué des méthodes causales pour comparer l'efficacité de deux traitements de base reçus en plus d'autres médicaments prescrits simultanément pour le traitement de la TB-MR, à la discrétion du prestataire : les trois médicaments du groupe A de l'OMS (bédaquiline, linézolide et une fluoroquinolone) sans clofazimine, comparés aux trois mêmes médicaments du groupe A plus clofazimine. Cette étude était une émulation d'essai ciblé utilisant des données d'observation où les régimes de base ont été comparés dans une analyse en intention de traiter utilisant des pondérations de probabilité inverse de traitement de base et quatre analyses per-protocole qui ont utilisé des pondérations de probabilité inverse de censure variant dans le temps pour estimer l'effet de traitement moyen (ETA). Dans cette étude, nous avons trouvé des preuves que l'ajout de la clofazimine augmente les résultats positifs du traitement lorsqu'elle est ajoutée aux médicaments du groupe A dans l'analyse en intention de traiter (l'ETA pondéré par le traitement de base : 0,07 [IC 95 % : -0,01, 0,15]), les résultats étant similaires dans les deuxième et quatrième analyses per-protocole. Cependant, l'ETA a été considérablement atténué dans les première et troisième analyses per-protocole où les patients ont été censurés pour avoir dévié de leurs stratégies de traitement respectives avec une censure supplémentaire des patients du groupe témoin qui ont commencé à prendre de la clofazimine après l'assignation du traitement de base : dans la

première analyse per-protocole, l'ETA pondéré par la censure était de -0,01 (IC à 95 % : -0,10, 0,07), tandis que dans la troisième analyse per-protocole, l'ETA pondéré par la censure était de 0,02 (IC à 95 % : -0,05, 0,09).

Dans mon troisième manuscrit, poursuite du traitement de la TB-MR, j'ai utilisé des données individuelles de patients provenant d'études d'observation pour identifier les caractéristiques des patients et les facteurs de traitement qui sont associés à la durée du traitement. Nous avons utilisé un nouveau résultat pour remédier aux biais inhérents à l'étude de la durée du traitement de la TB-MR: l'écart entre la durée du traitement individuel et la durée moyenne de leur site de traitement, en n'utilisant que les patients ayant obtenu de bons résultats. Dans cette étude, nous avons montré que l'utilisation de la bédaquiline était associée à une diminution de la durée du traitement de 0,51 (IC 95 % : 0,15, 0,87) mois, ce qui était cohérent entre les sous-groupes, tandis que la TB-MR avec résistance aux fluoroquinolones était associée à une augmentation de 0,78 (IC 95 % : 0,36, 1,21) mois. Nos résultats sont cohérents avec la littérature, car nous avons montré que de nombreux facteurs connus pour être associés à de mauvais résultats thérapeutiques étaient également associés à des durées de traitement plus longues. Les résultats de cette étude pourraient permettre aux futures recherches sur les traitements plus courts de la TB-MR d'élargir les critères d'inclusion aux patients qui ont été précédemment exclus des essais et des lignes directrices.

Il est impératif d'améliorer le traitement de toutes les formes de tuberculose pour atteindre les objectifs de l'OMS en matière de lutte contre la tuberculose et améliorer les résultats pour les patients. Les travaux décrits dans cette thèse fournissent des preuves sur lesquelles les praticiens actuels peuvent s'appuyer pour le traitement de l'infection tuberculeuse et que les

futurs chercheurs pourront utiliser pour améliorer les schémas thérapeutiques pour les patients atteints de TB-MR.

Acknowledgements

I need to thank Dr. Dick Menzies for his supervision during my PhD, completing this thesis would not be possible without his clinical insight and research experience. Dick provided many opportunities for me over the years I have worked with him and gave me a chance to show that anyone willing can complete a PhD despite small obstacles. Without Dr. Andrea Benedetti I would have struggled to conduct the analyses necessary for my first objective, I appreciate the clear and concise help you provided and for everything else you have done for me. I need to thank Dr. Mireille Schnitzer for her methodological input and Dr. Jon Campbell for answering hundreds of questions, many of them multiple times, and for his support throughout my PhD. Also thanks to Dr. Robert Platt and Dr. Kristian Filion for being on my thesis committee and helping to guide my research.

I must thank everyone in the Menzies TB group. Fede, Chantal, Olivia: thank you for not just helping me with all the admin tasks and questions, but always being the kindest people one can work with. To the trainees: Jon, Mayara, Mercedes, Tommy, Saeedeh, and Hannah – I was fortunate to have such cooperative colleagues that are also my good friends.

To the 2018 epidemiology cohort, you all made slogging through this PhD a little easier. The TH crew: Martha and Rip you made an experience that should be traumatic into something I will look back on fondly. Rip you helped me more than you'll ever know. To my MSc folk: Aemal, Ola, and Jana, thank you for the support both then and now.

Finally, I want to thank my family for putting up with me during this process and for everything they have done in my life that enabled me to get here. I am only able to write this because of their understanding and support.

Statement of financial support

My thesis work was supported by a Doctoral Fellowship from the Fonds de Recherche du Québec Santé from May 2019 to August 2023. I also received supplemental funding from my supervisor, Dr. Dick Menzies through his Canadian Institutes of Health Research Foundation Grant.

Contributions of authors

As first author on each manuscript, I developed the research questions and study design in collaboration with my supervisor Dr. Menzies and co-authors Dr. Jonathon Campbell, Dr. Andrea Benedetti, and Dr. Mireille Schnitzer.

Manuscript 1: *Completion, safety, and efficacy of tuberculosis preventive treatment regimens containing rifampicin or rifapentine: an individual patient data network meta-analysis.*

Dr. Menzies, Dr. Campbell, and I conceptualized the study. I coordinated the project and wrote the protocol, contacted authors, acquired data, and then cleaned, combined, and harmonized all data into the individual patient data set. With feedback from Dr. Benedetti, I designed the statistical analyses. I conducted the statistical analyses and drafted the manuscript, while all authors provided critical feedback.

Manuscript 2: *Efficacy of adding clofazimine to WHO group A drugs in treatment of rifampicin- and multidrug-resistant tuberculosis: an emulated target trial*.

Dr. Menzies and I conceptualized the study. Statistical analyses were conceived by Dr. Schnitzer and myself, while I conducted the statistical analysis with programming assistance from Dr. Liu. I drafted the manuscript, while all authors provided critical feedback on substantive content and interpretation.

Manuscript 3: *Identifying patients with multidrug-resistant tuberculosis who may benefit from shorter durations of treatment.*

Dr. Menzies and I conceptualized the study. I conceptualized the statistical analyses and received feedback from Dr. Schnitzer and Dr. Ripley. I performed all statistical analyses and

drafted the manuscript. All authors provided critical and substantive feedback on the manuscript.

Statement of originality

The manuscripts included in this thesis are all original scholarship that contribute to the scientific literature on treatment of tuberculosis infection and MDR-TB.

In my first manuscript I evaluated the completion, safety, and efficacy between 3HP and 4R. To my knowledge, this was the first study that was used to compare these two tuberculosis preventive treatments, and the first to use individual patient data and network meta-analysis methods for this comparison. The second manuscript involves the comparison of two regimens to treat MDR-TB that have yet to be compared in adequately powered randomized controlled trials. Further, I used a target trial framework and applied causal methods to account for time-varying confounding to assess the impact that treatment changes have over time, approaches that are rarely used in tuberculosis research. Finally, in the final manuscript I used a novel outcome for the treatment of MDR-TB to assess associations between treatment duration and both clinical and treatment characteristics. Although previous studies have tried to assess treatment duration, my study aimed to account for the complexities inherent in investigating duration of MDR-TB treatment by using the deviation from the mean treatment duration of an individual's treatment site as the outcome. We also described associations between treatment duration and site-level factors, which had not previously been described.

Although I have received guidance from my supervisor, thesis committee, co-authors, and colleagues on methodological and substantive aspects of my thesis, I declare that the research presented herein is of my own original work.

List of tables

Supplemental Table 4.S6. Adjusted risk ratios (aRR) and risk differences (aRD) with their 95% confidence intervals (CI) from direct and network meta-analysis models for the comparison of the incidence of adverse events (various definitions) between 3HP and 4R in the safety population.

Supplemental Table 4.S7. In the entire and per-protocol populations: the adjusted risk ratio (aRR) and risk difference (aRD) with their 95% confidence intervals (CI) from both direct and network meta-analysis estimates for the comparison of treatment-related grade 3 or 4 adverse events that led to permanent drug discontinuation between 3HP and 4R.

Supplemental Table 4.S8. Number and rate (%) of adverse events for all definitions, by HIV status in the safety population.

Table 5.1. Baseline characteristics of the overall study population by treatment group (intention-to-treat analysis).

Table 5.2. Drugs received in addition to drugs in the core regimens, by treatment group in the overall population.

Table 5.3. Average treatment effects and 95% confidence interval (CI) between the clofazimine group and the control group among each analysis population and by method of analysis.

Table 5.4. Drugs received that were permanently stopped due to an adverse event at any time during treatment.

Table 5.5. Reasons for earliest censoring event in all per-protocol analyses.

Supplemental Table 5.S1. Target trial specification and emulation using EndTB cohort.

Supplemental Table 5.S2A. Baseline characteristics of patient in the per-protocol analyses by censor status.

Supplemental Table 5.S2B. Baseline characteristics of patients in the per-protocol analyses by censor status.

Supplemental Table 5.S3. Stratified analyses for effect measure modification by drug susceptibility results for fluoroquinolones in the intention-to-treat analysis.

Supplemental Table 5.S4. Baseline characteristics of the patients in the control group who started clofazimine at any time during treatment and those who did not.

Supplemental Table 5.S5. Adverse events leading to permanent drug stoppage for patients in the control group who started clofazimine at any time during treatment and those who did not.

Table 6.1. Site-level characteristics and their univariable and multivariable associations with the site-specific mean treatment duration in patients with successful treatment outcomes. Estimates and 95% confidence interval (CI) from linear regression models.

Table 6.2A. Description of patient characteristics and their association (adjusted for age and sex) with deviation in treatment duration from site mean.

Table 6.2B. Description of drugs used in treatment and their association (adjusted for age and sex) with deviation in treatment duration from site mean.

Table 6.3. Associations of individual deviation in treatment duration from site mean with patient characteristics, resistance categories, and drugs used, within specified subgroups. Estimates and 95% confidence interval (CI) from multivariable linear mixed models including all variables shown (unless otherwise specified).

Supplemental Table 6.S1. Comparison of patients in the IPD who were included and excluded from this analysis.

Supplemental Table 6.S2. Mean total treatment duration and total number of patients at each site included in the study population.

Supplemental Table 6.S3. Associations between deviation in treatment duration from site mean and patient characteristics, resistance categories, and drugs used from the full multivariable model with inverse probability of selection weights and without (primary analysis), including all variables listed.

Supplemental Table 6.S4. E-values for selected characteristics with largest effect estimates for deviation in treatment duration.

Supplemental Table 6.S5. Associations between deviation in treatment duration from site mean and patient characteristics, resistance categories, and drugs used by different combinations of subgroups, and their regression estimates and 95% confidence intervals (CI) from a multivariable linear mixed model including all variables listed (unless specified otherwise).

List of figures

List of abbreviations

AE	Adverse event
AFB	Acid-fast bacilli smear results
aIRD	Adjusted incidence rate difference
aIRR	Adjusted incidence rate ratio
aRD	Adjusted risk difference
aRR	Adjusted risk ratio
ART/ARV	Antiretroviral therapy
ATE	Average treatment effect
BMI	Body mass index
CDC	Centres for Disease Control
CI	Confidence interval
CTCAE	Common Terminology Criteria for Adverse Events
DOT	Directly observed therapy
DST	Drug susceptibility testing
ECOG	Eastern Cooperative Oncology Group functional status
HIV	Human immunodeficiency virus
IGRA	Interferon-gamma release assays
IPCW	Inverse probability of censor weight
IPD	Individual patient data
IPTW	Inverse probability of treatment weight
IQR	Interquartile range
IRD	Incidence rate difference
IRR	Incidence rate ratio
ITT	Intention-to-treat
MDR-TB	Multidrug-resistant tuberculosis
NMA	Network meta-analysis
PLHIV	People living with human immunodeficiency virus
PS	Propensity score
RCT	Randomized controlled trial
RD	Risk difference
RR	Risk ratio

RR-TB	Rifampin-resistant tuberculosis
SD	Standard deviation
TB	Tuberculosis
TBTC	Centres for Disease Control's Tuberculosis Trials Consortium
TPT	Tuberculosis preventive treatment
TST	Tuberculin skin test
WHO	World Health Organization

Drug abbreviations

3HP	3 months rifapentine plus isoniazid
4R	4 months of rifampicin
RIF	Rifampin/rifampicin
6H/9H	6 or 9 months of isoniazid
3HR	3 months of rifampicin plus isoniazid
Bdq	Bedaquiline
Lzd	Linezolid
Cfx	Ciprofloxacin
Gfx	Gatifloxacin
Lfx	Levofloxacin
Mfx	Moxifloxacin
Ofx	Ofloxacin
Cfz	Clofazimine
Cs	Cycloserine
Trd	Terizidone
E	Ethambutol
Am	Amikacin
Amx-Clv	Amoxicillin-clavulanate
Clr	Clarithromycin
Cm	Capreomycin
Dlm	Delamanid
Eto	Ethionamide
Imp	Imipenem
Km	Kanamycin

Mpm	Meropenem
PAS	ρ-aminosalicylic acid
Pto	Prothionamide
S	Streptomycin
Thz	Thioacetazone
Z	Pyrazinamide
BPaL	Bedaquiline, pretomanid, and linezolid
BPaLM	Bedaquiline, pretomanid, linezolid, and moxifloxacin
BPaLC	Bedaquiline, pretomanid, linezolid, and clofazimine
SLI	Second-line injectable
FQ	Fluoroquinolone

Chapter 1 – Introduction and thesis objectives

Section 1.1 – Introduction

Tuberculosis is a substantial global health burden, accounting for an estimated 10.6 million cases and 1.6 million deaths in 2021.[1] Tuberculosis exists on a continuum and is thought of as being in two main states: tuberculosis infection and active tuberculosis disease.

Tuberculosis infection is a dormant, non-contagious, and asymptomatic state of tuberculosis. The number of people with tuberculosis infection is estimated to be about one quarter of the global population.[2,3] About 5-10% of those infected will develop tuberculosis disease in their lifetime,[4] thus treating tuberculosis infection is an essential part of reducing the burden of tuberculosis disease.[5,6] Two shorter regimens, 3 months of rifapentine plus isoniazid (3HP) and 4 months of rifampicin (4R), have recently been recommended for tuberculosis preventive treatment by the World Health Organization (WHO). However, as 3HP and 4R have not been tested head-to-head in a clinical trial, questions remain regarding their relative rates of completion and safety.

Multidrug-resistant tuberculosis (MDR-TB) is defined as tuberculosis resistant to both the antibiotics isoniazid and rifampicin, and there were an estimated 450,000 cases in 2021.[1] Treatment outcomes for MDR-TB are much worse than in those with drug-susceptible tuberculosis.[1,7] Currently, the WHO recommends up to four drugs for 18-24 months to treat MDR-TB patients with extensive disease (defined by cavitation or bilateral disease on x-ray) and additional drug resistance. Recently, new drugs showing high efficacy have been added to the list of WHO group A, B and C drugs for MDR-TB treatment.[8,9] The WHO recommends that an effective regimen consists of all three group A drugs (bedaquiline,

linezolid, and a fluoroquinolone) and at least one group B drug (clofazimine and cycloserine/terizidone) so that treatment is initiated with at least four likely effective drugs.[9] However, uncertainty remains regarding the efficacy of adding group B drugs to the recommended regimen of all three group A drugs, and specifically the effect that adding clofazimine has on treatment outcomes.

For MDR-TB patients without past tuberculosis treatment or extensive disease, shorter 9-12 month treatment regimens have been recommended by the WHO.[9] However, the trials informing these decisions excluded patients with low body mass index, low HIV CD4 cell counts, serious comorbidities, and additional resistance.[10-13] Thus, questions remain whether patients with these factors could also benefit from shorter treatment regimens.

Section 1.2 – Thesis objectives

The overall objective of this thesis is to improve treatment for tuberculosis infection and MDR-TB. There are several available regimens for tuberculosis preventive treatment, but there has previously been no study comparing the two shorter regimens, 3HP and 4R. For MDR-TB, studying the efficacy of drug regimens and the duration of treatment is difficult due to complexities of the disease, multiple drugs used, and the unique resistance profile of each patient which require an individualized approach to treatment. For those reasons, there is a lack of evidence for efficacy of drug regimens and what factors are associated with shorter treatment duration. In this manuscript-based thesis I aim to address these research gaps.

Section 1.2.1 – Objective 1: Tuberculosis preventive treatment

To conduct a network meta-analysis using individual patient data to compare completion, safety, and efficacy between 3HP and 4R for the treatment of tuberculosis infection **(manuscript 1).**

Section 1.2.2 – Objective 2: Time-varying analysis of MDR-TB treatment

To determine the average treatment effect between two core MDR-TB treatment regimens of the group A drugs with and without the addition of clofazimine, using a target trial approach and to determine the impact of time-varying confounding with use of inverse probability of censor weights **(manuscript 2)**.

Section 1.2.3 – Objective 3: Factors associated with treatment duration of MDR-TB

To determine which clinical characteristics, drug susceptibility testing results, and drugs used in MDR-TB treatment are associated with shorter treatment duration in patients who had successful treatment outcomes using their individual deviation in treatment duration from the mean treatment duration of their treatment site to identify those who may benefit from shorter treatment **(manuscript 3)**.

Chapter 2 – Review of the literature

Section 2.1 Tuberculosis

Tuberculosis disease has been documented in humans for thousands of years[14] and persists today as a major global health burden. The cause of the disease was identified as the bacteria *Mycobacterium tuberculosis* by Robert Koch in 1882.[15] Tuberculosis is an airborne infectious disease that is spread person-to-person when tuberculosis-infected individuals expel bacteria into the air by coughing or speaking.[16,17] Risk of transmitting tuberculosis increases with severity of the index case (bacterial load, smear positivity, cavitary disease, or severity of cough) and the proximity and duration of an exposed person (i.e. close and casual contact).[17]

Despite advances in diagnosing and treating tuberculosis, it remains one of the most prevalent and lethal communicable diseases in the world, accounting for an estimated 10.6 million cases and 1.6 million deaths in 2020.[1] The incidence of tuberculosis varies widely across the globe (Figure 2.1), and in 2021 eight countries contributed to two thirds of the global total of estimated cases: India, China, Indonesia, the Philippines, Pakistan, Nigeria, Bangladesh, and South Africa.[1] Although many factors affect the geographical distribution of tuberculosis, one of the primary factors is poverty[18-20] with low- and middle-income countries having the highest rates of tuberculosis.[1] Although the overall tuberculosis incidence in Canada is low at about 5/100,000 population in 2020, the effect of income disparity is reflected in our disease distribution as well, which remains considerably high in

Inuit (70.3/100,000 population) and First Nations on-reserve populations (20/100,000 population).[21]

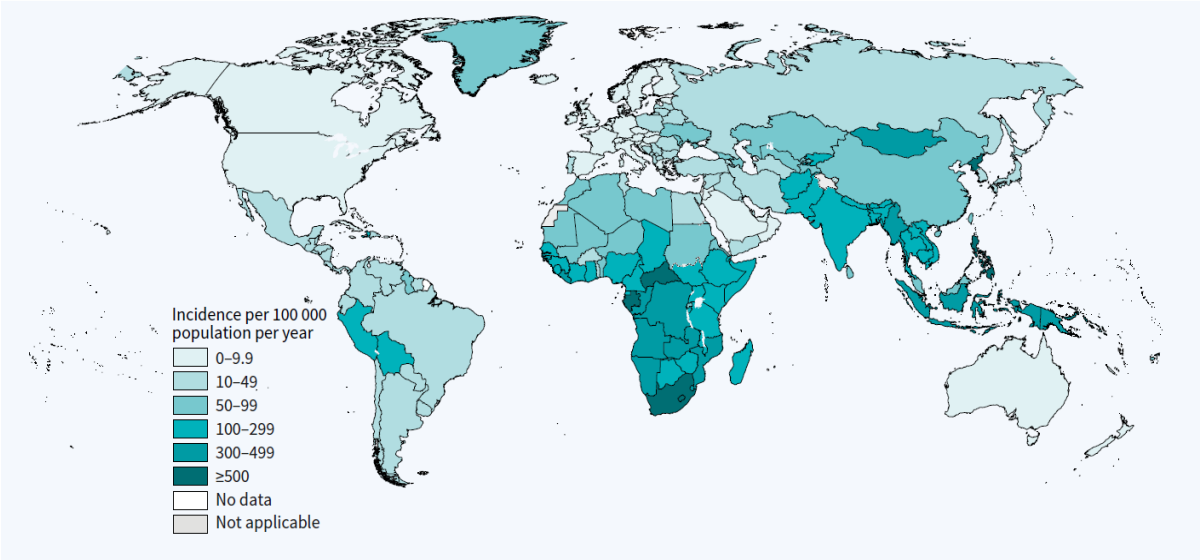


Figure 2.1. Estimated global incidence of tuberculosis per 100,000 people in 2021[1]

Testing to diagnose tuberculosis disease, including drug-resistant tuberculosis, is performed in patients exhibiting symptoms consistent with tuberculosis disease (fever, night sweats, weight loss, cough longer than three weeks, pain in chest, and/or coughing up blood or sputum[22]) or in those who are at high risk of tuberculosis disease (recent immigrants from areas with high endemicity of tuberculosis, HIV infected, immune compromised, or on dialysis).[23] Diagnosis is made using: 1) chest radiography; 2) microbiological culture and/or acid-fast bacilli smear; and/or 3) detection of *M. tuberculosis* using nucleic acid amplification tests such as Xpert MTB/RIF which uses the GeneXpert platform from Cepheid.[23]

Once infection and disease are confirmed, treatment should be initiated. For treatment of drug-susceptible tuberculosis disease, the WHO in their 2022 guideline recommend a 6 month regimen (2 months of isoniazid, rifampicin, pyrazinamide, and ethambutol followed

by 4 months of isoniazid and rifampicin) and conditionally recommend a 4 month regimen (2 months of isoniazid, rifapentine, moxifloxacin, and pyrazinamide followed by 2 months of isoniazid, rifapentine, and moxifloxacin).[24] In 2020, the WHO reported that approximately 86% of people treated for drug-susceptible tuberculosis had a successful treatment outcome.[1]

Although improved case identification and treatment initiation are important,[25,26] completing the prescribed doses in a regimen is essential to successful treatment outcomes and reducing the impact of tuberculosis disease. However, current treatments are burdensome on both patients and treatment programmes.[4,27,28] In situations already exacerbated by socioeconomic factors affecting tuberculosis care, the high number of pills in regimens[29,30] and adverse events[31-33] are considered by patients to be barriers for treatment completion and success. More effective, shorter regimens including drugs with better safety profiles are crucial to reducing the burden of tuberculosis treatment.

Tuberculosis exists on a continuum and is thought of as being in two main states: tuberculosis infection and active tuberculosis disease. Tuberculosis infection was previously referred to as latent tuberculosis infection, however the WHO now refer to this simply as tuberculosis infection.[34] The active state of tuberculosis is a symptomatic disease that is further categorized based on the drug resistance profiles of the patient. Tuberculosis that is resistant to both the antibiotics isoniazid and rifampicin[8] is referred to as multidrug resistant tuberculosis (MDR-TB).

Section 2.2 Tuberculosis infection

Tuberculosis infection is a dormant, non-contagious, and asymptomatic state of tuberculosis.[35,36] In those who have been exposed to tuberculosis and have tuberculosis infection, risk of progressing to tuberculosis disease is highest in the first two years, but only a few healthy people will develop tuberculosis disease[17] as host immune defenses usually contain the infection. Factors determining progression to tuberculosis disease include the time from infection and the state of an individual's immune system.[36,37] Risk of progression to tuberculosis disease is increased in young children and those with HIV infection, silicosis, and immunocompromising conditions[36] (Table 2.1).

Table 2.1. Risk of progression to tuberculosis disease among different populations testing positive for tuberculosis infection stratified by group (adapted from Campbell et al. Chapter 4: Diagnosis of tuberculosis infection[36,37])

Risk Factors	Annual risk of TB disease for the first 2-3 years after testing positive (%)
VERY HIGH RISK	
People living with HIV	1.7 to 2.7
Child or adolescent (<18y) tuberculosis contact	2.9 to 14.6
Adult (≥18y) tuberculosis contact	0.8 to 3.7
Silicosis	3.7
HIGH RISK	
Stage 4 or 5 chronic kidney disease with or without dialysis	0.3 to 1.2
Transplant recipients (solid organ or hematopoietic)	0.1 to 0.7
Fibronodular disease	0.2 to 0.6
Receiving immunosuppressing drugs (e.g., TNFα inhibitors or steroids)	0.5
Cancer (lung, sarcoma, leukemia, lymphoma or gastrointestinal)	0.1 to 0.4
MODERATE RISK	
Granuloma on chest x-ray	0.1
Diabetes	0.1 to 0.2
Heavy alcohol use (at least 3 drinks/day)	0.1 to 0.2
Heavy tobacco cigarette smoker (at least 1 pack/day)	0.1
LOW RISK	
General (adult) population with no known risk factor	0.03
Persons with a positive two-step TST booster and no known risk factor	0.02

Abbreviations: TB, tuberculosis; HIV, human immunodeficiency virus; TST , tuberculin skin test; TNF: tumour necrosis factor.

The geographical distribution of tuberculosis infection is likely similar to that of tuberculosis disease[38] and is correlated with the same risk factors. The number of people with tuberculosis infection is estimated to be roughly one quarter of the global population.[2,3] Given this large reservoir and estimates that 5-10% of those infected will develop tuberculosis disease in their lifetime,[4] treating tuberculosis infection is an essential part of meeting the goals of the WHO's End-TB strategy: achieving an 80% reduction in the 2015 incidence rate of tuberculosis by 2030.[5,6]

Identifying patients and initiating treatment are important steps in the cascade of care for tuberculosis prevention.[25,26] Testing for tuberculosis infection is recommended for individuals at high-risk of progressing to tuberculosis disease,[35,36] and the test results will help identify infection in those who may benefit from treatment.

Although no gold standard exists, tuberculosis infection is in part diagnosed using two tests that are recommended by the WHO.[4] The first is the tuberculin skin test which is an intradermal injection of a purified protein derivative of non-specific *M. tuberculosis* antigens.[36,39] A previous infection with tuberculosis causes cell-mediated immunity and re-exposure to tuberculin antigens through the injection results in a hypersensitivity reaction that manifests as an induration on the skin. The second test is the interferon-gamma release assay,[40,41] which is another test of cell-mediated immune response but uses blood samples to measure interferon-gamma released from T-cells following stimulation with *M. tuberculosis* specific antigens. Additionally, a negative chest x-ray or if a chest x-ray is abnormal a negative mycobacterial sputum culture may be used to rule out tuberculosis disease.[42] Once

tuberculosis disease is ruled out, tuberculosis preventive treatment should be initiated in those indicated for screening and testing positive for tuberculosis infection.

Section 2.2.1 Current treatments for tuberculosis infection

Historically, for tuberculosis preventive treatment the WHO[35] recommended isoniazid for 6 months (6H) and the Canadian Thoracic Society recommended 9 months of isoniazid (9H).[42] Although effective at preventing tuberculosis disease, these regimens are associated with poor completion rates[43,44] and significant liver toxicity.[45]

In 2020, the WHO recommended 3 months of rifapentine plus isoniazid (3HP) and a conditional recommendation for 4 months of rifampicin (4R)[35] for tuberculosis preventive treatment. The US Centres for Disease Control and Prevention in 2020[46] and the Canadian Thoracic Society in 2022[42] also recommended 3HP and 4R for tuberculosis preventive treatment. Additionally, 3 months of rifampicin plus isoniazid (3HR) was also recommended by the WHO.[35] These recommendations were based on the results of several randomized controlled trials conducted over the past 20 years.[47-50] Overall, these trials indicated similar efficacy for prevention of tuberculosis disease between the longer regimens and the newer shorter regimens.

Treatment completion in shorter tuberculosis preventive treatment regimens

The PREVENT TB study was a randomized controlled trial conducted by the Centres for Disease Control's Tuberculosis Trials Consortium (TBTC), in which 3HP was compared to 9H.[47] The results of this trial indicated that in 7,731 patients older than 12 years of age, 3HP had higher treatment completion rates than 9H.[47] The TBTC's iAdhere trial found that self-administered 3HP had a lower completion rate (74%) than directly observed treatment with

3HP (87.2% completion rate).[51] Similar results for completion of 3HP compared to 9H were found in 905 patients 2-17 years of age[52] who were enrolled in the PREVENT TB study that included children. In another subset of the PREVENT TB study, 3HP had higher completion than 9H in people living with HIV.[53] Furthermore, in a randomized controlled trial of 1,148 HIV-infected patients, both 3HP and 3HR had higher completion than 6H.[49]

Other shorter regimens using rifampicin have also been investigated. Three randomized controlled trials were used to compare 4R against 9H for treatment of tuberculosis infection in 829 children[50] and 6,859 adults (847 from a phase 2 trial[45] and 6012 from a phase 3 trial[48]). In these trials, investigators showed that 4R had substantially better completion than 9H.

Adverse events in shorter tuberculosis preventive treatment regimens

One of the primary considerations when choosing tuberculosis preventive treatment is safety,[31,54-56] and questions remain regarding which treatment has fewer adverse events. Although 4R is well tolerated with lower rates of grade 3 to 5 adverse events than 9H,[48,50] 3HP has been shown to be associated with higher rates of discontinuation than 9H due to adverse events[47] and higher rates of flu-like syndrome than 9H as reported in the PREVENT TB study.[57] In the iAdhere trial,[51] rates of adverse events in the 3HP treatment groups were similar to the PREVENT TB study.[47] Conversely, in HIV patients, 3HP was shown to be as tolerable as 9H.[49,53] Results of a meta-analysis assessing 3HP indicated similar rates of adverse events as with use of the conventional tuberculosis preventive treatments (6H, 9H, continuous isoniazid, 4 months of rifampicin plus isoniazid, and 2-3 months of rifampicin plus pyrazinamide), however the majority of included studies were observational cohort designs which may have greater risk of bias than randomized trials.[44] In a network meta-

analysis of 61 randomized controlled trials of tuberculosis preventive treatment, few studies reported data for hepatoxicity between regimens which limited inference (while no other adverse events were assessed), and the reported comparisons only show indirect evidence.[58] Such relatively high rates of adverse events for 3HP are concerning, considering the significantly shorter treatment duration and use as prophylaxis, and evidence comparing 3HP to 4R is needed.

Research gaps for tuberculosis preventive treatment

There is a demand from clinicians and patients for shorter and more tolerable regimens for treatment of tuberculosis infection.[27,28,59] As 4R and 3HP have not been compared directly in a randomized trial the question regarding which of these shorter regimens has better completion and safety, remains. My first objective aims to address this gap in the literature.

Effective and tolerable tuberculosis preventive treatments are crucial to reducing the burden of tuberculosis disease. However, long treatment durations or drugs with high risk of adverse events can prevent patients from completing (and accepting) the required dosage, which may result in tuberculosis disease.[31,60] A consequence beyond developing tuberculosis disease, is that improper dose and/or non-adherence to intolerable treatments may be associated with the acquisition of drug resistance,[61,62] which results in tuberculosis that is substantially more difficult to treat.

Section 2.3 Multidrug-resistant tuberculosis

As defined previously, MDR-TB is tuberculosis that is resistant to both rifampicin and isoniazid, the two most effective first-line tuberculosis antibiotics.[9] Globally, MDR-TB is

devastating with nearly 41% of untreated cases dying.[1] Although treatment success has increased over time,[1,7] the number of estimated MDR-TB cases has increased from previous years to 450,000 in 2021.[1] For those who started treatment for rifampicin resistant or MDR-TB in 2018, approximately 10% failed treatment, 16% were lost to follow-up, and 14% died.[63] Treatment outcomes are worse for individuals with HIV infection,[64] underweight (body mass index <18.5 kg/m^2),[65] and malnutrition[62] when compared to MDR-TB patients without these respective comorbidities.

Countries with an MDR-TB incidence greater than 10 notified cases per 100,000 people in 2021 were Kazakhstan, Kyrgyzstan, Turkmenistan, Russia, and South Africa while other countries including India, Peru, Namibia, the Philippines, and many Eastern European countries had an incidence rate between 4-10 notified cases per 100,000.[66] Canada has a substantially lower occurrence of MDR-TB with only a total of 22 notified cases in 2021, compared to the higher burden countries like India which had nearly 60,000 and China and Russia which had nearly 20,000 total cases each.[1,66]

Drug resistance in tuberculosis occurs primarily due to genetic mutations in the genes carrying drug targets or enzymes.[62,67,68] The two main ways people contract drug-resistance is either by primary infection with a drug-resistant strain or acquired resistance which develops during treatment of a drug-susceptible strain.[8,69,70] Factors that may increase the risk of acquiring drug resistance include poor treatment adherence, inadequate dosing, intermittent ingestion of drugs, and/or previous exposure to tuberculosis drugs.[62,69,71] Evidence is uncertain regarding characteristics that may increase the risk of primary infection with MDR-TB, however those with MDR-TB tend to be younger, live in locations with high endemicity, and have had exposure to people with MDR-TB.[62] Additionally, use

of intravenous drugs, incarceration, and homelessness are associated with an increased risk of MDR-TB. Infection with HIV is also associated with increased risk of MDR-TB, however this may be due to common risk factors listed previously or nosocomial transmission.[69]

Diagnosis of MDR-TB follows a similar algorithm as described previously for drug-susceptible tuberculosis, where additional results of phenotypic, culture-based drug susceptibility testing for resistance and/or molecular methods to predict possible resistance (such as Xpert MTB/RIF) are used to define drug resistance patterns.[23,68] Phenotypic, culture-based methods involve growing bacterial cultures in media that contain specific antibiotic drugs and observing if growth is inhibited.[72] Phenotypic, culture-based testing is the gold standard, but can take nearly two months to obtain results.[23] Molecular testing is an alternative to culture-based testing that provides results much faster and is performed either with targeted assays or whole-genome sequencing.[23,73] For targeted assays, such as Gene Xpert MTB/RIF and line probe assays, amplification of genetic sequences are performed to find a single mutation of a specific gene sequence known to be associated with drug resistance (for example the genes katG and rpoB for isoniazid and rifampicin resistance respectively).[62,67,68,74] However, the sensitivity is not perfect, as these assays target specific sections of a gene and not others where mutations could also occur.[23] Further, targeted assays cannot confirm whether a mutation confers drug resistance. Alternatively, whole genome sequencing uses the entire DNA of *M. tuberculosis* and can determine if mutations confer drug resistance as well as identify multiple mutations at once.[23,62,68] The choice of these diagnostics will depend on resources available and local burden of disease.[75] Once

drugs have been identified to which a strain of tuberculosis is resistant, treatment with effective drugs can tailored to fit the patient.

Section 2.3.1 Treatment for multidrug-resistant tuberculosis

Those treated for MDR-TB have much worse outcomes than those treated for drug-susceptible tuberculosis, with the global rate for treatment success being only 60-70%[1,7] compared to 85%[1] in drug-susceptible tuberculosis. Such poor rates of success are in part due to the arduous treatment length, lack of effective drugs, and drug-related severe adverse events.[30,32,76] Until 2020, the WHO had recommended up to five drugs for 18-24 months to treat MDR-TB,[77] while drug-susceptible tuberculosis treatment is only 6 months.[24]

Treatment for MDR-TB with advanced disease and extensive resistance

Treatment of MDR-TB has historically been highly individualized to account for the unique drug resistance profile of each patient. Several programmatic aspects determine which treatment a patient will receive including drug availability and resources for drug susceptibility testing.[75] Additionally, the majority of MDR-TB patients are in resource limited settings, making it difficult to conduct the very long and expensive trials required to assess treatment efficacy. Consequently, there are few randomized controlled trials assessing MDR-TB drugs and most evidence comes from observational studies in programmatic settings. Furthermore, the treatment regimens recommended by the WHO vary by extent of disease (defined by presence of cavitation and/or bilateral disease on chest radiography and/or acid-fast bacilli smear results[9]) and drug resistance patterns, which is why MDR-TB has conventionally been treated with individualized regimens.[77]

The current drugs recommended for treatment of MDR-TB are presented in Table 2.2, along with their WHO grouping. Recently, the WHO has added the use of new drugs, delamanid and bedaquiline, to the list of WHO group A, B, and C drugs used for treatment of MDR-TB.[8] Bedaquiline is the first new antibiotic for tuberculosis to be approved by the US Food and Drug Administration (FDA) since the 1970s, while delamanid is only approved by the European Medicines Agency and not the FDA.[78]

Table 2.2. WHO recommended MDR-TB drugs, groupings, and class.[8,24]

	Drug names	Class	Abbreviation
Group A	Levofloxacin OR moxifloxacin	Fluoroquinolones (FQ)	Lfx OR Mfx
	Bedaquiline	Diarylquinolines	Bdq
	Linezolid	Oxazolidinone antibiotics	Lzd
Group B	Clofazimine	Leprostatics	Cfz
	Cycloserine OR terizidone	Streptomyces derivatives	Cs OR Trd
Group C	Ethambutol	Antitubercular agent	E
	Delamanid	Nitroimidazole	Dlm
	Pyrazinamide	Antitubercular agent	Z
	Imipenem OR meropenem	Carbapenems	Imp OR Mpm
	Amikacin OR streptomycin	Second-line injectables (SLI), Aminoglycosides	Am OR S
	Ethionamide OR prothionamide	Pyridines and derivatives	Eto OR Pto
	ρ-aminosalicylic acid	Aminosalicylates	PAS
Drugs no longer recommended as first-line treatment	Amoxicillin-clavulanate	Beta-lactamase inhibitors	Amx-Clv
	Thioacetazone	Benzene and substituted derivatives	Thz
	Kanamycin, Capreomycin	Second-line injectables (SLI), Aminoglycosides	Km, Cm
	Clarithromycin	Macrolide	Clr
	Ofloxacin, Ciprofloxacin, Gatifloxacin	Fluoroquinolones (FQ)	Ofx, Cfx, Gfx

In those with more extensive disease and resistance in addition to MDR-TB (i.e. with resistance to a fluoroquinolone and/or second-line injectable), recommended treatment can be as long as 18-20 months. The WHO recommends that an effective regimen consists of all three group A drugs and at least one group B drug so that treatment is initiated with at least

four likely effective drugs.[8,24] If only one or two group A agents can be used, then both group B agents are to be included as part of the regimen. If an effective regimen cannot be made from groups A and B alone, group C agents are added until five drugs are used.

In an individual patient data meta-analysis,[79] the group A drugs (bedaquiline, linezolid, and fluoroquinolones) as well as carbapenems and clofazimine have shown to be individually associated with treatment success (defined according to WHO 2013[80] or Laserson et al.[81] criteria as completion or cure) compared to death, failure, relapse, and loss to follow-up[79]. In the same individual patient data meta-analysis amikacin, cycloserine/terizidone, pyrazinamide, and streptomycin were associated with modest benefit but only in those without resistance. However, the use of kanamycin and capreomycin were associated with negative outcomes, while ethionamide or prothionamide, capreomycin, kanamycin, ρ-aminosalicylic acid, macrolides, and amoxicillin-clavulanate were associated with no benefit or poor outcomes. With such few effective drugs available, it is already difficult to start patients on regimens that are likely to result in successful treatment outcomes.

Treatment is further complicated when considering the safety of available drugs. The group A drugs bedaquiline and fluoroquinolones have relatively better safety profiles with few patients discontinuing treatment due to adverse events, but other drugs such as linezolid, second-line injectables, and ρ-aminosalicylic acid are associated with higher rates of drug discontinuation.[76] Although associated with treatment success, the group A drug linezolid has a poor safety profile when used in extended durations and is associated with increased risk of adverse events (about 14% of MDR-TB patients discontinued the drug[76]), which include myelosuppression and neurotoxicity.[76,82] Use of ρ-aminosalicylic acid is associated with an increased risk of adverse events, and the second-line injectables are associated with

potentially permanent hearing loss and nephrotoxicity,[76,82] and are no longer recommended as first-line treatment of MDR-TB by the WHO.[8,24] The group B drug clofazimine is associated with a relatively lower rate of adverse events but can cause skin discolouration,[76,82,83] the stigmatization of which can lead some patients to discontinue the drug.[84] However, there is a lack of randomized controlled trial data and limited evidence accessing the efficacy and safety of clofazimine.[83,85] Only a few small randomized controlled trials which included fewer than 450 patients across all trials have been used to assess clofazimine in MDR-TB patients, but the regimens used did not contain bedaquiline.[85] Given that clofazimine has been recommended by the WHO to be added to the three group A drugs, more evidence regarding the efficacy and safety of this drug is greatly needed.

Research gaps for treatment for MDR-TB with advanced disease and extensive resistance

The highly individualized treatments make it difficult to assess drug regimens against a standard comparator in a randomized trial. The long duration, lack of effective drugs, and risk of adverse events have a substantial impact on completion and success in the treatment of MDR-TB.[30,33] Patients with MDR-TB can be receiving many different drugs and applying methods for causal inference evokes unique challenges that require additional consideration to account for the multiple concurrent drugs in a regimen as well as drugs added later in treatment.[86] Furthermore, time-varying factors throughout treatment (such as adverse events, changes in health status of the patient, bacterial culture results, and acquired resistance) can cause changes in treatments being used and are also predictive of outcomes.[87,88] As the majority of studies on treatment outcomes in MDR-TB patients involve assessing correlates of efficacy for individual drugs given at treatment initiation only,[64,79,89] the impact that time-varying confounding may have on estimates of effect has

rarely been assessed, limiting causal interpretations. Thus, there are still questions regarding which drug combinations are optimal for treatment of MDR-TB patients. To account for this potential bias and help identify causal effects without randomization, methods such as inverse probability of treatment and censoring weights can be used in combination with both the counterfactual framework and emulation of target trials[90-93] (the detail of which are outlined in methods Section 3.3).

Shorter treatment for MDR-TB

In the last few years there have been some promising changes in MDR-TB treatment not just with the development of newer drugs but also investigations into shorter regimens. In the past 10 years several studies[10,12,13,94,95] have been used to investigate shorter regimens for treatment of MDR-TB and are outlined in Table 2.3. The "Bangladesh regimen", a short 9 to 12-month standardized course of seven drugs showed promising results in an observational study.[95] This led to initiation of the STREAM trials, which were used to assess the efficacy of this shorter treatment under randomization. The STREAM stage 1 trial investigators found that the 9 to 12-month Bangladesh regimen was non-inferior to the longer 18 to 20-month WHO regimen at achieving a favourable outcome.[13] In the subsequent STREAM stage 2 trial, investigators reported that 9- and 6-month all-oral bedaquiline containing regimens were superior to the "Bangladesh regimen" in achieving a favourable outcome.[12] The results from the STREAM stage 1 trial[13] trial indicated similar safety between the shorter regimens and standard WHO regimens, while in the STREAM stage 2 trial[12] results indicated similar safety but with fewer cases of hearing loss in the short, all-oral regimen compared to the second-line injectable containing "Bangladesh regimen."

Table 2.3 Description of studies and trials investigating shorter treatment regimens for multidrug-resistant tuberculosis

Author (year)	Excluded patients and population notes	Primary outcome	Size (n)	Regimens	Months of treatment	Successful primary outcome % (95%CI)
Studies of MDR-TB						
Van Deun (2010)[95] "Bangladesh regimen"	**Excluded**: Previous second-line TB treatment **Notes**: Low HIV prevalence	Successful outcome defined as: Completion (or cure) compared to death, failure, LTFU, default, relapse (WHO 2013)	427	1) Gfx-based, Pto+H (intensive phase), Cfz throughout 2) Ofx-Based, Pto+H throughout 3) Ofx-based, Pto throughout, no H 4) Ofx-based, Pto (intensive phase), H throughout 5) Ofx-based, Pto (intensive phase), H+Cfz throughout	9 12 12 12 12	88 (83, 92) 69 (78, 58) 57 (41, 72) 66 (52, 75) 84 (68, 92)
Nunn (2019)[13] STREAM Trial 1	**Excluded**: MDR-TB with resistance to FQ or SLI, critically ill (unlikely to survive 4 months of treatment)	Favourable status defined as: negative cultures for *M. tuberculosis* 132 weeks post-randomization and a prior point during trial period, with no positive cultures in between	424	1) Mfx/Cfz/Eth/Z + [Km/H/Pto (3 months)] 2) Standard WHO regimens	9 20	79.8 78.8
Goodall (2022)[12] STREAM Trial 2	**Excluded**: MDR-TB with resistance to FQ or SLI, critically ill (unlikely to survive 4 months of treatment), previous treatment past 12 weeks, CD4 <50cells/mm3	Favourable status defined as: negative cultures for *M. tuberculosis* 76 weeks post-randomization and on proceeding visit with no positive cultures in between	588	1) Standard WHO regimens 2) Mfx/Cfz/Eth/Z + [Km/H/Pto (3 months)] 3) Lfx/Cfz/Eth/Z/Bdq + [H/Pto (3 months)] 4) Lfx/Cfz/Z/Bdq + [H/Km (2 months)]	20 9 9 6	Not reported 71 & 69* 83 91
Nyang'wa (2022)[96] TB-PRACTECAL	**No exclusions made** for HIV, CD4 count, or BMI, or past treatment. **Notes:** RR-TB patients	Favourable status defined as no unfavorable event (death, treatment discontinuation, failure, LTFU, or relapse) at 72 weeks post- randomization.	552	1) Bdq/Pretomanid/Lzd/Mfx (BPaLM) 2) Bdq/Pretomanid/Lzd/Cfz (BPaLC) 3) Bdq/Pretomanid/Lzd (BPaL) 4) Standard WHO regimens	6 6 6 9 to 20	76 / 96† 72 / 90 66 / 88 47 / 88
Studies of MDR-TB with additional resistance to FQ and SLI						
Conradie (2020)[11] NixTB Trial	**Excluded**: any comorbidity likely to compromise protocol assessments, Karnofsky score <50, low BMI, CD4 <=50 cells/uL	favourable outcome defined as: clinical TB disease had resolved, negative culture status at 6 months post-treatment end, and no unfavourable outcome (failure or relapse)	109	Bdq/Pretomanid/Lzd (BPaL)	6	90 (83, 95)
Conradie (2022)[10] ZeNix Trial	**Excluded**: any comorbidity likely to compromise protocol assessments, Karnofsky score <60, low BMI, CD4 <=100 cells/uL	Favourable outcome defined as: continued negative culture status to the end of follow up and no previous unfavourable outcome (failure or relapse)	181	1) Bdq/ Pretomanid /Lzd (1200mg for 26 weeks) 2) Bdq/ Pretomanid /Lzd (1200mg for 9 weeks) 3) Bdq/ Pretomanid /Lzd (600 mg for 26 weeks) 4) Bdq/ Pretomanid /Lzd (600 mg for 9 weeks)	6 6 6 6	93 (81, 99) 89 (76, 96) 91 (79, 98) 84 (70, 93)

* Two different randomizations of the same arm. † Percentages are for intention-to-treat / per-protocol arms TB: tuberculosis; RR-TB: rifampicin-resistant TB; LTFU: lost to follow-up; FQ: fluoroquinolone; SLI: second-line injectable; BMI: body mass index; Gfx: gatifloxacin; Ofx: ofloxacin; Mfx: moxifloxacin; Cfz: clofazimine; Eth: ethambutol; H: isoniazid; Km: kanamycin; Pto: prothionamide; Z: pyrazinamide; Bdq: bedaquiline; Lzd: linezolid

However, the success of these regimens may be predicated on susceptibility to drugs in the standardized regimen and no additional resistance to second-line injectables or fluoroquinolones, as patients with such resistance were excluded from these trials.[12,13] This is important, as evidence from an individual patient data meta-analysis indicated an increased risk of treatment failure in patients who are receiving drugs to which they have resistance on drug susceptibility testing.[79,89] Further, the WHO does not recommend the 9-month all-oral regimen to patients who have had previous treatment or extensive disease (as defined by presence of cavitation or bilateral disease on chest radiography).[77] In addition to the randomized trials describe above, evidence from an observational study assessing a 9 to 12-month injectable containing regimen consisting of a fluoroquinolone, clofazimine, ethambutol, pyrazinamide, ethionamide or prothionamide, high dose isoniazid, and an injectable compared to a 9 to 12-month bedaquiline containing regimen consisting of the same drugs but 6 months of bedaquiline replacing the injectable, indicated that the bedaquiline containing regimen had 14% (95% CI: 8, 20) higher treatment success compared to all negative outcomes, in a programmatic setting.[97]

These studies in shorter regimens led the WHO to recommended the use of an all-oral bedaquiline containing regimen lasting 9-12 months in programmatic settings for patients without exposure to second-line injectables, without resistance to fluoroquinolone and/or second-line injectable, and with non-extensive tuberculosis disease.[98]

Recently, the TB-PRACTECAL,[96] Nix-TB,[11] and ZeNix[10] trials have investigated the use of shorter treatment regimens as well as a novel therapeutic pretomanid (a nitroimidazooxazine developed by the TB Alliance and approved by the US FDA in 2019 under the Limited Population Pathway for Antibacterial and Antifungal Drugs[99]), for

treatment of rifampicin-resistant tuberculosis (RR-TB), MDR-TB, and MDR-TB plus additional resistance to a fluoroquinolone and/or second-line injectables. In the TB-PRACTECAL trial conducted by Médecins Sans Frontières,[96] the efficacy and safety for treatment of RR-TB was evaluated for three different all-oral regimens: 6 months of bedaquiline, pretomanid, and linezolid (BPaL), BPaL plus moxifloxacin (BPaLM), and BPaL plus clofazimine (BPaLC), compared to standard WHO regimens. Compared to standard WHO regimens, all three shorter all-oral 6-month regimens showed superior efficacy for successful treatment outcomes in the intention-to-treat analyses and similar efficacy in the per-protocol analyses, with substantially fewer grade 3 or higher adverse events and serious adverse events.[96] In 2022, the WHO recommended BPaLM/BPaL regimens for treatment of MDR/RR-TB.[9]

The Nix-TB[11] and ZeNix[10] trials were used to assess the BPaL regimens for treatment of MDR-TB and MDR-TB plus additional resistance to a fluoroquinolone and/or a second-line injectable. These trials indicated that 90-93% of enrolled subjects treated with BPaL regimens had successful treatment outcomes (see Table 2.3). However, these trials had either no comparator arm or no comparator arm with standard care or conventional treatment. There were also a significant number grade 3 to 5 adverse events in both trials, with a total of six deaths due to adverse events in the NixTB trial.[11] Further, these trials had strict inclusion criteria and excluded patients with low body mass index, low HIV CD4 cell counts, and those with any comorbidity likely to compromise protocol assessments.[10,11] Additionally, of those in the ZeNix trial who had poor treatment outcomes nearly all had cavitation on chest radiography. Therefore, the populations used in these trials do not

necessarily reflect those treated in programmatic settings and inference on relative efficacy and safety are not possible due to lack of comparator arms.

Nonetheless, in 2022 the WHO recommended the use of BPaL regimens for treatment of some forms of MDR-TB plus additional resistance to fluoroquinolones and/or second-line injectables despite the strict patient populations, adverse events, poor outcomes in those with extensive disease, and no comparator arms used in the trials informing their decision.[24]

Shorter regimens are promising for patients and treatment programmes as they will help reduce the burden associated with longer MDR-TB treatment.

Research gaps for shorter treatments for MDR-TB

There have been advances in reducing duration of treatment for some patients with MDR-TB. Questions remain regarding whether the shorter regimens can be used effectively in the patients with past first-line or second-line tuberculosis treatment, more extensive tuberculosis disease, additional drug resistance, and presence of comorbidities who have previously been excluded from randomized trials informing treatment guidelines.[10-13] What is also not known is which patient characteristics and drugs used in treatment are associated with the success observed in MDR-TB patients treated for shorter durations. Addressing these questions may help broaden eligibility for trials of shorter regimens and improve treatment experience and outcomes for more MDR-TB patients.

Section 2.4 Rationale for thesis

Evidence is needed to address the gaps outlined in my literature review, but new trials and observational studies for tuberculosis are lengthy and expensive. However, there are existing data in the form of published randomized controlled trials, observational cohorts, and individual patient datasets that can be used to answer these questions. In this thesis I use existing data and apply novel methods to generate evidence that can be used to address these gaps in the literature for treatment of tuberculosis infection and MDR-TB.

Chapter 3 – Methodology used in thesis objectives.

In this chapter I will discuss some methods that were not described in detail in the manuscripts. All other methods are covered within the presented manuscripts.

Section 3.1 Network meta-analysis with individual patient data

In traditional meta-analysis, aggregate data from studies of the same treatment comparison (A vs. B) can be combined, which generate a more precise effect estimate than a single study. However, for studies of different treatments (A vs. B and C vs. B) traditional meta-analysis for comparisons of treatments that were not tested head-to-head is not appropriate (i.e. A vs. C).[100] A network meta-analysis approach can generate indirect, study-level effect estimates for scenarios like A vs. C above, by comparing effects between their common comparator (B)[101]. Although newer network meta-analysis approaches can make some adjustment for covariates,[102] they are limited by the same heterogeneity of the published outcomes and lack of ability to adjust for individual level characteristics of patients that exist with meta-analysis of aggregate data.

Use of individual patient data is the best way to overcome these limitations of heterogeneity as it allows standardization of inclusion criteria, exposure, confounders, and outcomes.[103,104] An individual patient-data meta-analysis starts by conducting a literature review to identify studies eligible for inclusion. However, unlike traditional meta-analyses, the aggregate data from identified studies are not analyzed. Rather, study authors are contacted to request their original patient data which are then combined into a single dataset.[105] Although individual

patient data meta-analyses have mainly been limited to combined studies of the same treatments, incorporating network meta-analysis methods with the use of individual patient data will provide the ability to indirectly compare regimens that have not been compared directly head-to-head[103] while allowing for the standardization of definitions across studies.[106] This should result in estimates that are more robust to study- and patient- level differences than those from traditional aggregate data meta-analyses and network meta-analysis.[107,108]

The network meta-analysis of individual patient data for my first objective was conducted in two stages. In the first stage, one-step individual patient data meta-analyses were conducted separately for studies of 3HP compared to 6H or 9H and for studies of 4R compared to 6H or 9H to obtain their direct effect estimates. The estimates for direct risk ratios (RR) were calculated using Poisson regression in generalized linear mixed models (GLMM) with a random intercept for study and a log link, while risk differences (RD) were calculated with a Gaussian distribution and identity link. To account for differences between study populations, estimates were adjusted for covariates considered a priori to be important predictors of the outcomes used.

In the second stage, the network meta-analysis of the indirect effects between 3HP and 4R were estimated as outlined by Morton et al.[24] and Veroniki et al.[25] using the estimates calculated from the direct, one-step individual patient data meta-analysis models outlined above, as follows:

$$\text{RR: } \log[RR_{3HPvs4R}] = \log[\text{direct } RR_{3HPvs6\text{-}9H}] - \log[\text{direct } RR_{4Rvs6\text{-}9H}], \text{ and}$$

$$\text{RD: } [RD_{3HPvs4R}] = [\text{direct } RD_{3HPvs6\text{-}9H}] - [\text{direct } RD_{4Rvs6\text{-}9H}]$$

Assuming no covariance, the variance is calculated as:

$$\text{Var}(\log[RR_{3HPvs4R}]) = \text{Var}(\log[\text{direct } RR_{3HPvs6\text{-}9H}]) + \text{Var}(\log[\text{direct } RR_{4Rvs6\text{-}9H}])$$

and the corresponding 95% CI would be calculated as:

$$RR_{3HPvs4R} +/- Z_{0.95} * \text{Sqrt}(\text{Var}(\log[RR_{3HPvs4R}])$$

It is important to note that variance is additive, however in my first objective confidence intervals were calculated using bootstrap methods.

Section 3.2 Data collection for individual patient data

The most challenging aspect of undertaking an individual patient data analysis of any kind is the collection and harmonization of data from included studies. This requires a lot of patience and time, and in the case of the individual patient dataset used in my first objective it took from Sept 2018 to June 2022 to have all data received, queries answered, and outcomes/variables harmonized into the final dataset used in analyses. Authors of identified studies were contacted in 2018 and invited to share their original trial data. The time from contact to a signed data sharing agreement varied from a few days to several months. Once data sharing agreements were finalized and signed, I began receiving the datasets that were currently available. Some data were transferred to public repositories and required third parties to address queries. In certain instances it would take months of back and forth between investigators only to be informed that the data cannot be locate by anyone involved in the projects or that those who conducted the study had retired. For the data that was received, many months (and even years) of emails were required to answer questions about the data, and some queries could not be resolved as the study analysts with the required

information were no longer working with the investigator. For reference, once data was finalized it took from June 2022 to September 2022 to have the manuscript written and submitted to the CDC for clearance (as CDC data was used) and then accepted for publication in February 2023. All this to say, the most challenging part of an individual patient data analysis is the construction of the data itself, and I only included six trials.

Section 3.3 Causal inference, inverse probability weighting, time-varying confounding, and target trials

A causal effect can be described in simple terms as the difference in an outcome because of a preceding event compared to what would have happened had that event not occurred, given all other factors are held constant.[109] In epidemiology an event is an exposure or treatment. In the potential outcomes framework proposed by Rubin[110,111], at the individual level each person has two potential outcomes, their outcome had they received treatment $Y_i(1)$ and their outcomes had they been untreated $Y_i(0)$. In an individual, the causal effect of treatment is $Y_i(1) - Y_i(0)$. In reality we can only observe what happened under the treatment they actually received, and we cannot observe what happened had they received a different treatment, i.e. their *counterfactual outcome*. To determine the effect of a treatment in the real world, populations need to be compared in which one group receives treatment and the other does not. This allows identification of the average treatment effect (ATE) which is the difference in the expectation of outcomes between groups: $E[Y_i(1) - Y_i(0)]$. As the ATE is unbiased only when treatment (A) assignment is independent of outcome (i.e. $Y \coprod A$), randomized trials provide the gold standard for identifying a causal effect.[112,113]

In observational studies treatment is not randomly assigned, therefore other methods must be used to create independence between treatment assignment and outcome.[113] As

randomization aims to ensure that the probability of treatment between two arms is 0.5, if the probability of receiving treatment between groups in an observational study can be balanced so there are no differences in probability of treatment between groups, it can be said that patients are effectively randomized to treatment based on measured confounders (note that most methods besides randomization or use of instrumental variables cannot account for unmeasured confounding[114]). One method of obtaining this balance is through use of propensity scores and inverse probability of treatment weighting (IPTW).[115,116] A propensity score is the probability of an individual in a study population receiving their assigned treatment conditional on a set of covariates (W), which is P(A=1 | W).[112] The IPTW of an individual is the inverse of their propensity score and is defined as 1/P(A | W) for those with A=1 and 1/[1-P(A | W)] for those with A=0. To account for extreme weights, the IPTW can be stabilized by including the marginal probability of treatment in the numerator and/or truncated at specified percentiles.[112] However, truncation of weights may introduce bias and methods such as truncating based on sample size[117] may help reduce the potential bias that percentile truncation my create. By weighting subjects a "pseudo-population" is created where an individual subject represents "copies" of themselves based on their weight (for instance, if their weight = 3 they represent 3 subjects). If the propensity score models are properly specified, the IPTW creates a population where $Y \coprod A \mid W$, i.e. the outcome is independent of treatment conditional on W, and the IPTW-ATE can be estimated as E[Y*i*(1)/P(A=1 | W) – Y*i*(0)/(1 – P(A=1 | W))].[112]

For causal inference to be made, the following assumptions must be met: i) consistency of outcome: the counterfactual outcome of the individual is the same as what was observed given the observed treatment (and treatment is well defined); ii) positivity of treatment: all

individuals in the subpopulation have a positive probability of having each treatment, given that they are available (i.e. the drug was available at the centre where the patient was treated) and conditional on patient-level confounders; iii) exchangeability of treatment: the counterfactual outcomes are independent of treatment and treatment availability conditional on a set of confounders.[118]

In longitudinal observational studies, including studies for treatment of MDR-TB, time-varying confounding is a concern. Figure 3.1 shows a hypothetical directed acyclic graph depicting such time-varying confounding. At time 0 a treatment regimen is assigned ($A_{t=0}$) which can be influenced by baseline confounders ($W_{t=0}$). Over the treatment course, there are changes in some confounders (L_t) that are both caused by treatment (for example adverse events) and can then affect the treatment being received at time 1 ($A_{t=1}$) as well as affect the outcome Y.

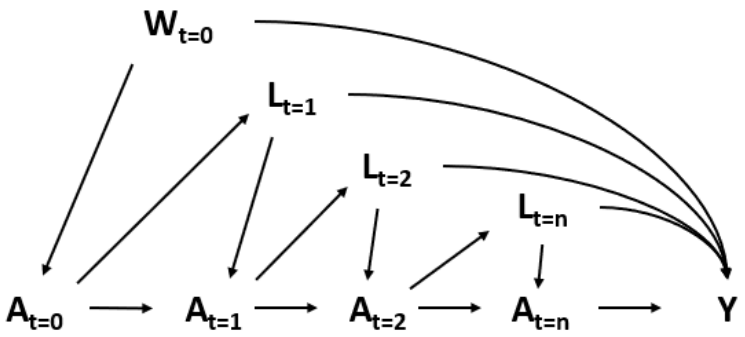


Figure 3.1 Directed acyclic graph depicting hypothetical time-varying confounding. A: treatment; W: baseline confounders; L: time-varying confounders; Y: outcome; t: time point during follow up where 0 is baseline up to last observed time point n.

As adherence to treatment is not randomized, adjustments should be made for prognostic factors and time-varying confounders that predict whether someone remains on assigned

treatment to de-confound the relationship between treatment and outcome.[119] There are several methods that can be applied to handle time-varying confounding, including g-computation,[90,92] longitudinal targeted maximum likelihood estimation,[120,121] and in the case of my second objective IPTW and inverse probability of censoring weights (IPCW).[90-92,122] In my second objective I censor subjects once they stop taking any drug in their assigned treatment intervention (but not for stopping any other concurrent MDR-TB drugs they received in addition to their invention regimen). The censor probabilities are calculated at specified time-points in the data and defined as the probability of being uncensored given baseline, time-varying confounders (L_t) and being previously uncensored (C_{t-1}) which is expressed as $P(C_t \mid W_0, L_{t-1}, C_{t-1}=0)$. Similar to IPTW, those who are uncensored receive a weight of $1/1 - P(C_t \mid W_0, L_{t-1}, C_{t-1}=0)$, while those who are censored receive a weight of 1 from the point they are censored (stop taking their regimen) and onward throughout their remaining treatment (censored subjects are excluded from the statistical analysis). The final weight represents the cumulative product of all IPCW and the baseline IPTW, defined as:

$$\prod_{i=n} 1/P(A_0 \mid W_0, L_{t=0}) * 1/P(C_t \mid W_0, L_t, C_{t-1}=0).$$

The previously outlined methods help identify the causal effect using statistical approaches. The target trial[123,124] is framework that enables a researcher to improve an observational analysis by emulating the desirable features of randomized trial.[125] A target trial is emulated in two steps: the first involves explicitly stating the protocol criteria of a hypothetical randomized trial. This includes protocol elements: eligibility criteria, treatment strategies, assignment procedures, follow-up period, outcomes, and the causal contrast of interest. These elements are then emulated in the observational data as presented in Table 3.1, as they apply to my target trial emulation in objective 2.

Table 3.1. Target trial specification and emulation for comparing MDR-TB treatments using the EndTB cohort

Protocol component	Target trial	Emulation in EndTB data
Eligibility	Patients of any age initiating MDR-TB treatment who are eligible at baseline to receive all drugs in each treatment regimen, disregarding drug susceptibility testing results.	Same
Treatment strategies	Initiation at baseline, in addition to other drugs being concurrently received for MDR-TB treatment at the discretion of the provider, of the core regimens of all three WHO group A drugs (bedaquiline, linezolid, and a FQ) without clofazimine (control group) or the same three WHO group A drugs (bedaquiline, linezolid, and a FQ) plus the addition of clofazimine (clofazimine group). 1st per-protocol analysis: the treatment strategy involved remaining on the assigned treatment but allowing for permanent drug stoppages for these reasons: adverse events, planned treatment changes, reintroduction or replacement of stopped drug, resistance to drug, drug supply or drug administration issues, or pregnancy but censoring for any other reason a patient stopped a drug as well as censoring patients in the control group who start clofazimine. 2nd Per-protocol analysis: the treatment strategy was the same as in the 1st per-protocol analysis but allowed patients in the control to start clofazimine without being censored. 3rd Per-protocol analysis: the treatment strategy involved remaining on the assigned treatment but censoring for any of the above-mentioned reasons for drug stoppages, as well as censoring patients in the control group who start clofazimine. 4th Per-protocol analysis, the treatment strategy was the same as in the 3rd per-protocol analysis but allowed patients in the control to start clofazimine without being censored.	Same
Assignment procedures	MDR-TB patients would be randomized to the control group or the clofazimine group	Individuals are assigned to each treatment group at baseline
Follow-up period	Patients are followed from treatment initiation until their end of treatment outcome.	Patients are followed until their end of treatment outcome or censor
Outcome	Treatment success was defined as cure or completion of treatment, compared to all other negative outcomes (death, treatment failure [defined as a change in any two drugs received including the supplemental, concurrent MDR-TB drugs in addition to core regimens due to an adverse event], or lost to follow up) as defined in WHO 2013.[80]	Same
Causal contrast	We will estimate the intention-to-treat (ITT) effect and the per-protocol effect	Observational analogue of the ITT and per-protocol effects.
Analysis plan	The ITT analysis involves direct comparison of the proportion of patients with treatment success among those assigned to each treatment. The per-protocol analyses will censor any patient deviating from the respective treatment strategies described above. The per-protocol analyses will adjust for both baseline and post-baseline confounders of treatment censoring.	Same

MDR-TB: multidrug-resistant tuberculosis; ITT: intention-to-treat; FQ: fluoroquinolone; IPTW: inverse probability of treatment weight; IPCW: inverse probability of censor weight.

Then an analysis plan is presented outlining the statistical methods being used to attempt to estimate the ATE in each analysis population in the emulated trial, which includes the

model specifications for IPTW, IPCW, time points used, and model specification for estimation of the final weighted estimates.

Although not all bias (especially unmeasured confounding) can be eliminated when using non-randomized data, a the target trial framework aims to eliminate biases that are common in observational research by specifically outlining a hypothetical target trial, so that focus can be on controlling for measured confoudning.[123,124] The target trial framework can help explicitly identify and account for biases common in observational studies, for instance outlining eligibility criteria that includes timing of treatment initiation can help prevent immortal time bias[126] and the specification of treatment strategies can prevent ill-defined interventions.[127] Well defined treatment strategies can then help identify selection biases due to non-adherence which can be accounted for using time-varying IPCW,[90-92,119] which is particularly important when studying the sustained treatment strategies[128,129] inherent in MDR-TB treatment.[88] Additionally, when trials are unable to be conducted due to cost, time, ethics, or feasibility an emulation of a target trial is a useful alternative to obtain a causal effect estimate for a research question using observational data.[124]

Combining all these methods is just an attempt to reduce potential bias as much as possible in order to identify a more valid causal effect of treatment using observational data.

Chapter 4 – Completion, safety, and efficacy of tuberculosis preventive treatment regimens containing rifampicin or rifapentine: an individual patient data network meta-analysis.

Section 4. 1 Preface

As described in Chapter 2, there is no randomized controlled trial comparing the two primary shorter tuberculosis preventive treatments, 3HP and 4R, that are recommended by the World Health Organization, Canadian Thoracic Society, and the US Centres for Disease Control and Prevention. Such a trial would be expensive and time consuming. However, evidence is still needed for clinicians to decide which treatment is optimal for people with tuberculosis infection.

In this manuscript I address this gap by using existing data from completed trials of 3HP compared to 6H or 9H and trials of 4R compared to 6H or 9H, to construct an individual patient data set. With this individual patient data from randomized trials I used network meta-analysis methods to generate indirect comparisons for treatment completion, safety, and efficacy between 3HP and 4R.

This work was published in *Lancet Respiratory Medicine* 2023; **11**(9): 782-90.[130]

Section 4.2 Manuscript 1

Completion, safety, and efficacy of tuberculosis preventive treatment regimens containing rifampicin or rifapentine: an individual patient data network meta-analysis.

Nicholas Winters MSc[1], Robert Belknap MD[2], Andrea Benedetti PHD[1], Andrey Borisov MD[3], Jonathon R. Campbell PHD[4,5,6], Richard E. Chaisson MD[7], Pei-Chun Chan PHD[8,9], Neil Martinson MBBCh[7], Payam Nahid MD[10], Nigel A. Scott MSc[3], Erin Sizemore MPH[3], Timothy R. Sterling MD[11], M. Elsa Villarino MD[12], Jann-Yuan Wang PHD[13], and Dick Menzies MD[1]

Affiliations:

1. Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montreal, Québec, Canada
2. Denver Health and Hospital Authority and Division of Infectious Diseases, Department of Medicine, University of Colorado, Denver, Colorado, USA
3. US Centers for Disease Control and Prevention, Atlanta, Georgia
4. Department of Medicine & Department of Global and Public Health, Faculty of Medicine and Health Sciences, McGill University, Montreal, Canada
5. McGill International TB Centre, Montreal, Canada
6. Respiratory Epidemiology and Clinical Research Unit, Centre for Outcomes Research & Evaluation, Research Institute of the McGill University Health Centre, Montreal, Canada
7. Johns Hopkins University School of Medicine, Center for Tuberculosis Research, Baltimore, Maryland, USA
8. Institute of Epidemiology and Preventive Medicine, College of Public Health, National Taiwan University, Taipei, Taiwan
9. Division of Chronic Infectious Disease, Taiwan Centers for Disease Control, Taipei City, Taiwan
10. UCSF Center for Tuberculosis, University of California, San Francisco, CA, USA.
11. Department of Medicine, Vanderbilt University Medical Center, Nashville, Tennessee, USA.
12. California Department of Public Health, Sacramento, California, USA
13. Department of Internal Medicine, National Taiwan University Hospital, Taipei, Taiwan

ABSTRACT

Background Three months of rifapentine plus isoniazid (3HP) and four months of rifampicin (4R) are recommended for tuberculosis preventive treatment (TPT). As these regimens have not been compared directly, we used individual patient data (IPD) and network meta-analysis (NMA) methods to compare completion, safety, and efficacy between 3HP and 4R.

Methods We conducted an IPD NMA by searching PubMed for randomized controlled trials (RCTs) published between Jan 1, 2000, and Mar 1, 2019. Eligible studies compared 3HP or 4R to isoniazid (H) for six or nine months (6-9H) and reported treatment completion, adverse events (AE), and/or incidence of tuberculosis. Deidentified IPD from eligible studies were provided by study investigators and outcomes were harmonized. Methods for NMA were used to generate indirect adjusted risk ratios (aRR) and risk differences (aRD) with their 95% confidence intervals (CI).

Findings We included 17572 participants from 14 countries in six trials. In the NMA, treatment completion was higher for those on 3HP than 4R: aRR 1.06 (95% CI: 1.02, 1.10); aRD 0.05 (95% CI: 0.02, 0.07). For treatment-related AEs leading to drug discontinuation, risks were higher for 3HP than 4R for AEs of any severity (aRR 2.86 [95% CI: 2.12, 4.21]; aRD of 0.03 [95% CI: 0.02, 0.05]) and grade 3-4 AEs (aRR 3.46 [95% CI: 2.09, 6.17]; aRD 0.02 [95% CI: 0.01, 0.03]). Similar increased risks with 3HP were observed with other definitions of AEs and were consistent across age groups. No difference in incidence of tuberculosis disease between 3HP and 4R was found.

Interpretation In absence of RCTs, our IPD-NMA indicated 3HP provided an increase in treatment completion over 4R but was associated with higher risks of AEs. Although findings should be confirmed, the trade-off between completion and safety must be considered when deciding TPT.

Funding None.

RESEARCH IN CONTEXT

Evidence before: Historically, the World Health Organization (WHO) has recommended six months of isoniazid (6H) for tuberculosis preventive treatment (TPT). In 2020, the WHO added a recommendation for three months of rifapentine plus isoniazid (3HP) and a conditional recommendation for four months of rifampicin (4R) for TPT. In randomized controlled trials (RCTs), these shorter regimens showed non-inferior efficacy for prevention of tuberculosis disease and better completion when compared to longer isoniazid regimens. Compared to nine months of isoniazid (9H) 4R was well tolerated, but trials of 3HP compared to 9H indicated an increased risk of grade 3-4 adverse events in those receiving 3HP. A meta-analysis indicated that 3HP had similar rates of adverse events as 6-9H, but mostly included observational studies, and in a network meta-analysis there was not enough data on hepatotoxicity to allow comparisons between regimens, and no comparisons between 3HP and 4R were reported. As 3HP and 4R have not been directly compared in an RCT, questions remain regarding optimal regimen selection. We conducted this individual patient data network meta-analysis to compare treatment outcomes between 3HP and 4R.

Added value of this study: In absence of direct head-to-head trials, we were able to generate evidence comparing 3HP to 4R. We showed that 3HP administered under directly observed therapy (DOT) had a higher proportion of participants completing this therapy than participants receiving 4R. Importantly, we found that 3HP was associated with higher proportion of participants with adverse events than 4R, a finding that was consistent in age subgroups and using different definitions of adverse events. We found no difference in efficacy for prevention of tuberculosis disease between 3HP and 4R.

Implications of all the available evidence: Our findings provide evidence for clinicians to draw on when deciding which shorter regimen to prescribe for TPT. The proportion completing therapy was higher with 3HP, which was administered under DOT in the clinical trials included in our analysis. The higher risk of serious adverse events associated with 3HP is of importance as safety is paramount for preventive treatments and must be considered in deciding between regimens. The trade-off between treatment completion and risk of adverse events needs to be considered when choosing TPT. Although ideally these findings would be confirmed in randomized trials directly comparing 3HP to 4R, such trials would be expensive and time consuming. Evidence from this study may assist clinicians on deciding optimal treatment, which will help improve efforts to reduce the global burden of tuberculosis disease.

INTRODUCTION

Tuberculosis is a global health burden with nearly 10.6 million reported cases and 1.6 million deaths estimated in 2021.[1] An estimated one quarter of the global population is living with tuberculosis infection (TBI),[2] of whom 5-10% will develop tuberculosis disease in their lifetime. Hence, treating TBI is essential to meet the goals of the World Health Organization's (WHO) End-TB strategy.[3,4]

Historically, the WHO has recommended daily isoniazid for six (6H) or nine months (9H)[5] for tuberculosis preventive treatment (TPT). Although these regimens have shown good efficacy, they are associated with poor completion rates[6,7] and significant hepatotoxicity.[8] There is a demand from clinicians and patients for shorter and more tolerable TPT regimens.[5,9,10]

In 2020, the WHO recommended a three-month regimen of rifapentine plus isoniazid (3HP) and 4 months of rifampicin (4R),[2] based on the results of several randomized controlled trials (RCTs), over the past 20 years.[11-15] In the trials supporting these recommendations, compared to mono-isoniazid regimens, 4R had significantly fewer grade 3-5 adverse events (AE), including hepatotoxicity,[11,13] while 3HP had lower hepatotoxicity, but higher overall rates of grade 3-4 AEs and AE-related drug discontinuation.[14,16] A meta-analysis concluded that proportions of AEs with 3HP and 6-9H were similar,[6] while a network meta-analysis (NMA) of 61 studies found no direct comparisons of 3HP and 4R and little evidence of difference in hepatoxicity between these two regimens.[17]

As 3HP and 4R have not yet been compared directly in a trial, uncertainty about optimal regimen selection remains. Therefore, we used existing data from completed RCTs to

perform an NMA of individual patient data (IPD), to generate indirect estimates of relative treatment completion, safety, and efficacy between 3HP and 4R.

METHODS

The protocol for this IPD-NMA was registered on PROSPERO (CRD42019124635).

Search strategy and selection criteria

We conducted a structured review of the literature to identify RCTs comparing treatment for TBI, published since 2000 (as we considered that individual-level patient data published earlier would be difficult to locate and obtain). A list of keywords and medical subject headings terms relating to TBI, drug regimens, and treatment outcomes were used to search PubMed for RCTs published between Jan 1, 2000, and Mar 1, 2019 (see Supplement 4.1 for detailed search). In addition, we identified relevant articles from the references in retrieved studies and from previously published reviews.

Eligible RCTs compared either 3HP or 4R to 6H or 9H, were published in peer-reviewed journals, and reported at least one of the following outcomes: treatment completion, treatment-related AEs, or incidence of tuberculosis disease. We also searched for studies of three months of rifampicin plus isoniazid (3HR) but were unable to acquire data for analyses. We included RCTs with participants of all ages who had a documented positive tuberculin skin test (TST) or interferon-gamma release assays (IGRA), or other conditions associated with increased risk of tuberculosis disease. We excluded observational studies, grey literature or unpublished data, and populations where participants were exposed to persons with isoniazid and/or rifampicin resistant tuberculosis strains.

Authors of all eligible studies were contacted and invited to contribute their deidentified individual-level patient data. Information requested included: i) baseline characteristics, ii) risk factors and indication for treatment, iii) treatment regimens, and iv) treatment outcomes: treatment completion, AEs, and incidence of tuberculosis disease (for full detail see Supplement 4.2).

To assess comparability of outcomes between studies, we also asked for study protocols and standard operating procedures to determine diagnostic methods and outcome assessment. For treatment completion, we requested participant pill counts and treatment durations. Information was abstracted for AEs to determine definitions (grading system, investigator defined, etc.), attribution to drug, and whether assessed by blinded, independent committee. For incidence of tuberculosis disease, methods for diagnosis were abstracted including laboratory tests and whether case records were adjudicated by an independent committee. Demographic characteristics and treatment outcomes were harmonized across studies, the accuracy of these procedures was validated by comparing to results reported in the original publications. All AEs were harmonized according to the grading criteria in the Common Terminology Criteria for Adverse Events (CTCAE) v5.0.[18] The criteria used to describe grading categories of AEs in other systems were matched to that of the CTCAE, and if discrepant were reassigned a grade to conform to what they would be classified as in the CTCAE (see Supplement 4.3).

Risk of bias was assessed using the Cochrane Risk of Bias 2 (RoB2) Tool[19] for randomized trials.

Our outcome of treatment completion was defined as taking more than 80% of the prescribed doses (using pill counts) in 120% of allowed time, dichotomized into completers or non-completers, and analyzed as risks among the entire population of all subjects randomized and included in the data sent for our IPD (see Supplement 4.4 for detailed descriptions of each regimen). We analyzed AE outcomes in the safety population (participants who took ≥1 dose of study drug) which were defined as (i) any treatment-related AE (i.e., adjudicated to be definitely, possibly, or probably related to study drug) that led to permanent drug discontinuation; and (ii) any treatment-related grade 3-4 AEs that led to permanent drug discontinuation. In secondary analysis of AEs, we compared the risk of (i) any AEs that led to permanent drug discontinuation (regardless of relationship to treatment); (ii) any grade 3-4 AE (regardless of relationship to treatment or drug discontinuation); and (iii) treatment-related grade 3-4 AEs (regardless of impact on drug discontinuation). To assess our outcome of efficacy for prevention of tuberculosis disease we estimated the relative incidence rate of all forms of tuberculosis disease, per 1000 person-years of follow-up, by pooling suspected, microbiologically confirmed, or clinically diagnosed tuberculosis disease in the entire population.

Data analysis

The power for each outcome in our NMA was determined using both traditional and indirect methods as described by Thorlund et al[20] (detailed in Supplement 4.5). We determined we had an adequate number of participants to detect an indirect difference between 3HP and 4R of 13-15% in the proportion of treatment completion with 80% power (0.05 alpha), using an intraclass correlation coefficient (ICC) of 0.004 calculated from data in included RCTs. For AEs, based on study ICCs of 0.0005, and the number of participants

in the datasets we obtained, we had 80% power (0.05 alpha) to detect indirect differences in proportion of AEs of 1.8 to 3.1% between 3HP and 4R. Due to inadequate power, all analyses of efficacy for prevention of tuberculosis disease were exploratory.

We conducted a NMA using IPD to estimate indirect treatment effects between 3HP and 4R using the estimates generated from direct analyses with their common comparator of 6-9H. As included studies were few, we pooled those receiving 6H or 9H and assumed clinical equivalence in the absence of trials directly comparing these two durations. The IPD-NMA was done in two stages. In the first stage, one-step IPD meta-analyses were conducted separately for studies of 3HP compared to 6-9H and for studies of 4R compared to 6-9H to obtain their direct effect estimates. The estimates for direct risk ratios (RR) were calculated using Poisson regression in generalized linear mixed models (GLMM) with a random intercept for study and a log link, while risk differences (RD) were calculated with a Gaussian distribution and identity link. The estimates for direct incidence rate ratios (IRR) were calculated using Poisson regression in a GLMM with a random intercept for study, person-time for follow-up incorporated as an offset, and a log link (incidence rate differences (IRD) were estimated as outlined in Bagos et al.[21]). To account for differences between study populations, estimates were adjusted for covariates considered a priori to be important predictors; each outcome was adjusted for different covariates and missing data for categorical variables were included as a 'not available' (NA) category (see Supplement 4.6 for detail). In our model building diagnostics, we assessed the impact that different specifications of random intercepts and the use of propensity scores (for confounders of AEs and tuberculosis incidence, as few events limited adjustment sets) had on both model fit and

variance. When no substantial differences were observed the model with simplest interpretation was chosen.

In the second stage, the NMA of the indirect RRs and RDs between 3HP and 4R were calculated from the estimates of the direct models as log[direct $RR_{3HPvs.6\text{-}9H}$] – log[direct $RR_{4Rvs.6\text{-}9H}$] for RRs and [direct $RD_{3HPvs.6\text{-}9H}$] – [direct $RD_{4Rvs.6\text{-}9H}$] for RDs (IRRs and IRDs were calculated similarly).[22,23] The 95% confidence intervals (CIs) were estimated with bootstrap resampling methods on 1000 replications and calculated using the 2.5th and 97.5th percentiles of the sampling distribution. As methods for assessing heterogeneity of adjusted IPD-NMA are not available and the number of included studies are few, this was not assessed statistically. All analyses were conducted using R, version 4.1.2.[24]

We conducted post-hoc sensitivity analyses of treatment completion in the NMA between 3HP and 4R, stratified by age (<18 compared to ≥18 years of age), age among adults only (<35, 35 to 65, >65 years of age), and HIV status. In addition, we investigated treatment completion only in studies using 9H as the comparator regimen. We analyzed differences in treatment-related AEs that led to permanent drug discontinuation and treatment-related grade 3-4 AEs that led to permanent drug discontinuation in the NMA between 3HP and 4R stratified by age (<50 and ≥50 years of age). We also analyzed treatment-related grade 3-4 AEs that led to permanent drug discontinuation in the entire population and per-protocol population (defined as all participants with an AE but excluding participants without AEs who did not complete >80% of prescribed doses).

RESULTS

The literature search identified data sets from 12 trials described in 17 publications,[11-16,25-35] of which six trials described in ten publications[11-16,26,27,30,32] were included: three trials that compared 3HP to 6-9H[12,14-16,26,27,30] and three trials that compared 4R to 6-9H[11,13,32] (Figure 4.1). Note, participants of one trial were reported in several publications[14-16,27,30]; we refer to this trial as CDC Study 26 (2011). Six trials described in seven publications[25,28-31,33-35] were not included: of these, data from four trials described in five publications could not be located or were no longer available,[25,28,31,34,35] one trial had no comparator arm of 6-9H,[33] and data from one trial was not included because corresponding authors did not respond.[29]

In total, we included 17,572 participants: 4,897 received 3HP; 4,055 received 4R; and 8,620 received 6-9H. Participants in the included data sets were enrolled in 14 countries in six WHO regions (Supplemental Table 4.S1). In one trial DOT was used for both 3HP and 9H.[26] In a second trial DOT was used for both 4R and 6H,[32] but was excluded from completion analyses due to insufficient data obtained to estimate completion. In two trials DOT was used for 3HP but the comparator arms of 6-9H were self-administered,[12,14] and in the final two trials, 4R and the 9H comparator arms were self-administered.[11,13] The average age of participants was similar across trials with exception of one trial in a pediatric population,[11] but the proportion with HIV ranged from 0[11,26,32] to 100%.[12] All three outcomes were available in five trials, while only AEs were available in one.[32] Supplemental Table 4.S2 presents the study level descriptions and outcomes of trials excluded from our IPD.[25,28,29,31,33-35] Study characteristics and outcomes were similar between included and excluded studies; the majority of excluded studies compared 3HR to 6-9H.

Overall RoB2 assessment indicated some concerns in four data sets, due to lack of blinded outcome assessments in these four trials[12,14,26,32]; and two open label trials had blinded and independent adjudication of the AEs[11,13] (Figure 4.S1).

For the overall population included in the NMA of 3HP and 4R (Table 4.1), the mean age was 34.9 years, 50.8% were female, and the mean body mass index was 25.4. Age, sex, body mass index, and recreational drug use were similar across treatment groups. The majority of participants were contacts (82.9%), while 68.6% of participants were close contacts (≥4 hours per week of contact with confirmed active TB case). The prevalence of people living with HIV (PLHIV) was 7.2% and was higher in studies of 3HP, while antiretroviral therapy (ART) use was higher in 4R (49%) than 3HP (1%).

Treatment completion

In the studies of 3HP compared to 6-9H, the number of participants completing treatment was 3963/4897 (80.9%) for those receiving 3HP and 2856/4614 (61.9%) for those receiving 6-9H (Table 4.2), resulting in an adjusted RR (aRR) of 1.30 (95% CI: 1.24, 1.37) and an adjusted RD (aRD) of 0.19 (95% CI: 0.17, 0.21). In the studies reporting completion of 4R compared to 9H, the number completing treatment was 2828/3865 (73.2%) for those receiving 4R and 2270/3823 (59.4%) for those receiving 9H resulting in an aRR of 1.23 (95% CI: 1.17, 1.30) and an aRD of 0.14 (95% CI: 0.12, 0.16).

In the NMA of the indirect effect between 3HP and 4R, treatment completion was more likely with 3HP with an aRR of 1.06 (95% CI: 1.02, 1.10) and aRD of 0.05 (95% CI: 0.02, 0.07). When only including studies of 9H as the comparator, the indirect aRR was 1.02 (95% CI: 0.98, 1.07) and the aRD was 0.03 (95% CI: 0.00, 0.06).

In sensitivity analyses, those under 18 years of age had a higher completion of 3HP than in the entire study population, with indirect aRR and aRD between 3HP and 4R of 1.12 (95% CI: 1.01, 1.23) and 0.07 (95% CI: 0.00, 0.15) respectively (Supplemental Table 4.S3). In those 18 years and older, the indirect aRR and aRD from the NMA between 3HP and 4R were similar to that of the overall study population. Completion between 3HP and 4R in those under 35 years of age was similar to those under 18 (aRR: 1.09 [95% CI: 1.02, 1.15]; aRD: 0.07 [95% CI: 0.03, 0.11]), but in those 35 to 65 and above 65 years of age there were no significant differences.

For the 1,271 PLHIV (Supplemental Table 4.S4), treatment completion was substantially higher for those receiving 3HP compared to 4R in indirect NMA. In those without HIV (n=11,817) differences in treatment completion between 3HP and 4R were similar to that of the overall study population.

Separate specifications of models with random intercept for country with missing category, random slope for treatment effects, or PS for adjustment had negligible impact on variance and model fit.

Adverse events

As presented in Table 4.3, the number of participants who experienced any treatment-related AE that led to permanent drug discontinuation were slightly higher with 3HP than 6-9H, and lower with 4R than 6-9H in direct comparisons. As a result, in the NMA 3HP had higher risk than 4R, with an aRR of 2.86 (95% CI: 2.12, 4.21) and the aRD of 0.03 (95% CI: 0.02, 0.05). Results were similar for treatment-related grade 3-4 AEs that led to

permanent drug discontinuation, and in the NMA 3HP had higher risk than 4R, with an aRR of 3.46 (95% CI: 2.09, 6.17) and aRD of 0.02 (95% CI: 0.01, 0.03).

For the indirect NMA stratified by age (Supplemental Table 4.S5), 3HP had greater risk than 4R for both treatment-related AEs of any grade and grade 3-4 events that led to permanent drug discontinuation, regardless of age category.

Using other definitions of AEs, differences were similar between 3HP and 4R (Supplemental Table 4.S6). Findings were similar in analyses using the entire population and per-protocol populations (Supplemental Table 4.S7). For rates of AEs by HIV status see Supplemental Table 4.S8.

Tuberculosis disease

In direct comparisons, the rate of tuberculosis disease was similar between 3HP and 6-9H, as well as between 4R and 9H (Table 4.4). In the NMA of the indirect effect, the rate of tuberculosis disease with 3HP was similar to that with 4R, with an aIRR of 1.16 (95% CI: 0.40 3.58) and an aIRD of 0.8/1000 person-years of follow-up (95% CI: -2.3, 7.0).

DISCUSSION

Our NMA comparing treatment outcomes between 3HP and 4R using IPD from six trials with 17,572 participants indicated that people treated with 3HP have about 5% higher treatment completion than those receiving 4R. However, compared to 4R, those treated with 3HP had 3% higher treatment-related AE that led to permanent drug discontinuation and 2% higher treatment-related grade 3-4 AEs that led to permanent drug discontinuation.

Our results suggest no difference in efficacy for prevention of tuberculosis disease between regimens, although this analysis was limited by the low number of disease occurrences.

Interpreting treatment completion between 3HP and 4R requires certain considerations. Differences in the regimens compared, including treatment scheduling (3HP taken once weekly and 4R taken daily) and site-level clinical practices, will affect completion. Importantly, for analysis of completion, treatment was self-administered in both arms of the included studies of 4R[11,13] whereas in studies of 3HP, all 3HP arms were under DOT but the comparator (6-9H) could be either self-administered or DOT.[26] Since DOT may increase treatment completion[33] the structure of the included trials may have differentially affected our analysis of completion favouring 3HP over 4R. Additionally, when excluding the single study using 6H[12], the difference in completion between 3HP and 4R was no longer significant; we could not distinguish whether this was due to the comparator arm regimen or because all those receiving 6H were PLHIV.

As only 5-10% of persons with TBI will progress to tuberculosis disease, and TBI is an asymptomatic condition, treatment safety is paramount. In this analysis the risk of AEs was higher among those who received 3HP compared to 4R, using different definitions of AEs. We were unable to compare risks of AEs in pediatric populations as events were too few, but the low numbers of AEs and high completion in those under 18 years of age may indicate better tolerability of all regimens in this age group.

Our study has some limitations to consider. The number of persons with HIV or other comorbidities (including diabetes or other immunosuppressive conditions), were too few to adequately analyze safety (including drug-drug interactions with ART) in these important subgroups. Imbalance in ART availability for PLHIV between 4R and 3HP adds complexity

in comparing outcomes in PLHIV in our IPD, but AE rates were actually lower among PLHIV, while the rate of TB disease was very low in all groups. Overall, we do not think the low numbers of PLHIV resulted in biased estimates, but certainly less precision. Hence, further research is needed to assess relative safety of 3HP and 4R in persons with HIV, or other comorbidities. Although our analysis suggests no major difference in efficacy between the two regimens, this important analysis was limited due to few persons who developed TB disease. Propensity scores (PS) may be inappropriate for prediction of randomly assigned treatment, thus adjusting for between study differences using variables with substantial missing data (such as renal failure, use of biologics, and immune-suppression other than HIV) was limited. However, in our model selection we assessed a PS that predicted the probability of a participant being in their given study and no substantial differences were observed between a model with this PS and our fully adjusted model. We could not include treatment site or country as random intercepts in our models (although we could include study) as this data was missing for a large portion of the population, leading to an under-estimation of variance. However, we assessed both fit (using AIC and BIC) and changes in variance between a model fit using a country variable with a missing category specified as a random intercept and our model fit with just a random intercept for study and observed no substantial differences. All treatment arms were unblinded in the included trials, and the consequent bias must be considered when interpreting results, notably the ascertainment of AEs with novel treatments such as 3HP. Additionally, calendar dates were unavailable as substantial data received were de-identified, precluding assessment of temporal trends within trials.

Despite these limitations, our study has several strengths. This is the first study to combine IPD and NMA approaches to provide adjusted indirect estimates of the relative completion, safety, and efficacy between 3HP and 4R, two treatment regimens that have not been directly compared in a randomized trial. The availability of individual level patient data from randomized trials enabled adjustment for study level differences and harmonization of outcomes across studies, resulting in estimates that are more robust to study/patient level differences than those from a traditional aggregate data NMA. In a previous NMA using aggregate data,[17] authors were unable to analyze AEs other than hepatotoxicity. Having access to the individual patients' data allowed us to harmonize all AE outcomes and assess differences between the two regimens using different definitions of AEs . Further, our sample size and number of events provided adequate power to make precise estimates for comparisons of treatment completion and AEs for age stratified analyses, although not for TB prevention, as noted above.

In the absence of trials directly comparing 3HP and 4R, this IPD-NMA from randomized trials of TPT provides evidence that 3HP under DOT had significantly higher treatment completion but also significantly higher risks of treatment-related AEs compared to 4R. This trade-off between completion and risk of AEs must be considered when deciding the optimal treatment for TBI.

Acknowledgments: The authors thank Dr. Andrew Vernon for his assistance with the CDC Study 26 data and research insight. The authors also thank the Critical Path Institute (C-PATH) members Mr. Dan Hartley, Pavan Kumar Sudhakar, and Ahmad Faizan for their assistance with CDC Study 26 data.

Data Sharing Agreement: As this is an analysis of individual patient data shared by contributing authors, they retain the right to review requests for sharing of their data. The data from Prevent TB is publicly accessible, for access contact C-PATH institute. For all other data, please contact the corresponding author of this paper (Dr. Menzies) with all data sharing inquiries.

Author Contributions: Mr. Winters and Dr. Menzies had full access to all the data and take responsibility for the integrity of the data and accuracy of data analysis.
Concept and Design: Menzies, Winters.
Acquisition, Analysis, or Interpretation of Data: All authors.
Drafting the Manuscript: Winters, Menzies.
Critical Revisions: All authors.

Conflict of Interest Disclosures: The authors have no relevant conflicts of interest to declare.

Funding/Support: There was no funding source for this study. Mr. Winters is supported by Fonds de recherche du Québec (FRQS) Doctoral Award and the Canadian Institutes of Health Research (CIHR) Foundation Grant. Dr. Menzies is supported by CIHR Tier 1 Canada Research Chair in Tuberculosis.

Ethics Committee Approval: This study was approved by an ethics committee of the Research Institute of the McGill University Health Center.

Figure 4.1 Study flow chart for identified studies and their inclusion status.

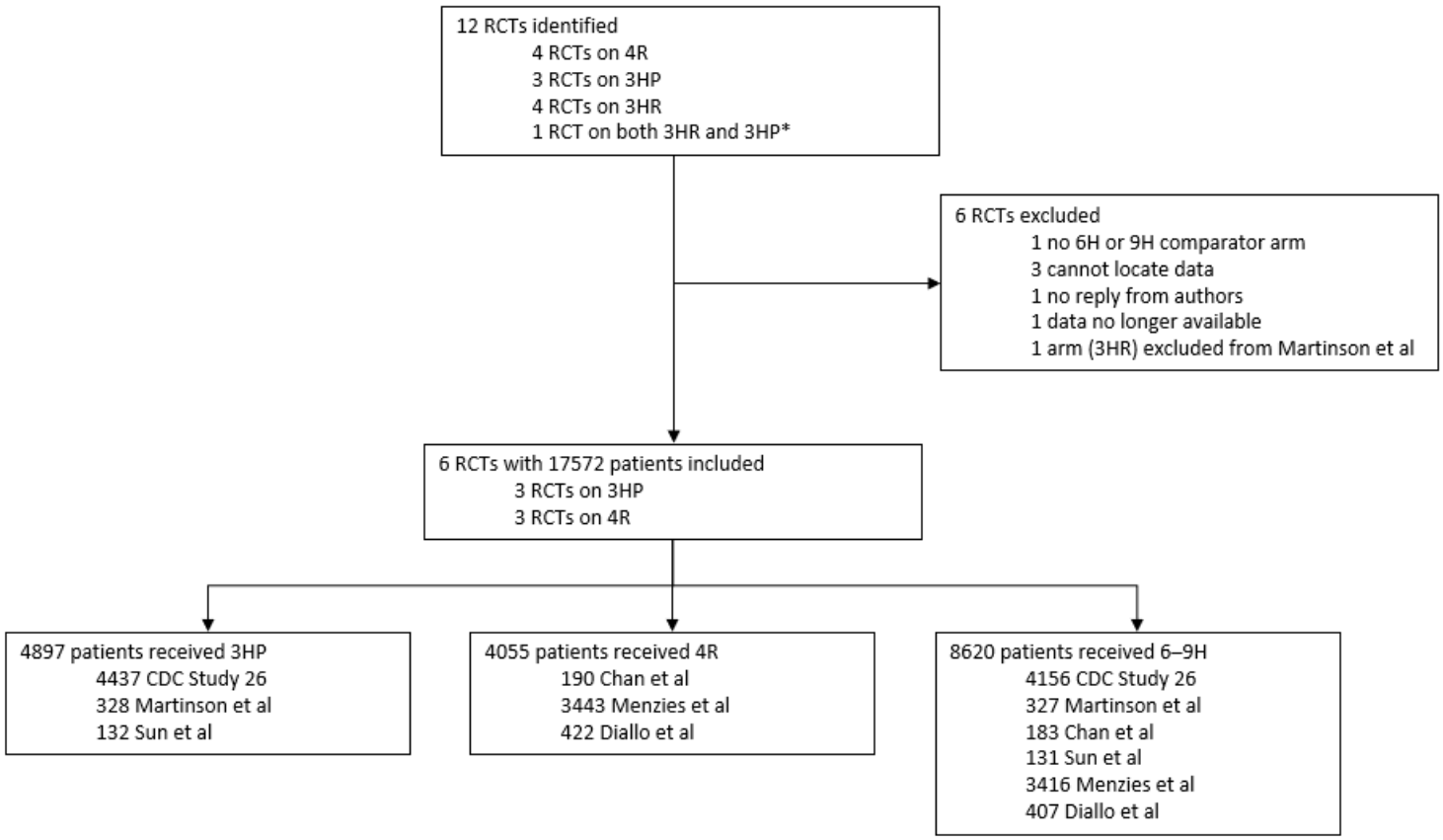


Table 4.1. Baseline characteristics of the participants included in the IPD by treatment received.

	3HP n = 4897	4R n = 4055	6-9H n = 8620	Overall n = 17572
Sex (%)				
Female	2298 (46.9)	2224 (54.8)	4413 (51.2)	8935 (50.8)
Male	2598 (53.1)	1831 (45.2)	4207 (48.8)	8636 (49.1)
NA	1 (0)	0 (0)	0 (0)	1 (0)
Age (mean (SD))	34.8 (15.4)	35.2 (15.7)	34.9 (15.3)	34.9 (15.4)
Body mass index (mean (SD))	26.7 (6.5)	23.8 (5.5)	25.4 (6.5)	25.4 (6.4)
Diabetes (%)				
No	459 (9)	3937 (97.1)	4349 (50.5)	8745 (49.8)
Yes	1 (0)	118 (3)	115 (1)	234 (1)
NA	4437 (90.6)	0 (0)	4156 (48.2)	8593 (48.9)
Renal failure (%)				
No	459 (9)	3592 (88.6)	4010 (46.5)	8061 (45.9)
Yes	1 (0)	43 (1)	27 (0)	71 (0)
NA	4437 (90.6)	420 (10)	4583 (53.2)	9440 (53.7)
Contact of active tuberculosis case (%)				
No	1033 (21.1)	190 (5)	1242 (14.4)	2465 (14.0)
Yes	3595 (73.4)	3865 (95.3)	7111 (82.5)	14571 (82.9)
NA	269 (6)	0 (0)	267 (3)	536 (3)
Type of contact of an active tuberculosis case (%)				
Not a contact	1033 (21.1)	190 (5)	1242 (14.4)	2465 (14.0)
Casual	0 (0)	402 (10)	358 (4)	760 (4)
Close	3518 (71.8)	2649 (65.3)	5891 (68.3)	12058 (68.6)
NA	346 (7)	814 (20)	1129 (13.1)	2289 (13.0)
Recent converter (%)				
No	2907 (59.4)	3300 (81.4)	5936 (68.9)	12143 (69.1)
Yes	1266 (25.9)	145 (4)	1362 (15.8)	2773 (15.8)
NA	724 (15)	610 (15)	1322 (15.3)	2656 (15.1)
Biologic use (%)				
No	132 (3)	3129 (77.2)	3212 (37.3)	6473 (36.8)
Yes	0 (0)	34 (1)	36 (0)	70 (0)
NA	4765 (97.3)	892 (22)	5372 (62.3)	11029 (62.8)
Immune suppression (%)				
No	459 (9)	3809 (93.9)	4232 (49.1)	8500 (48.4)
Yes	1 (0)	246 (6)	232 (3)	479 (3)
NA	4437 (90.6)	0 (0)	4156 (48.2)	8593 (48.9)
HIV infection (%)				
Negative	2204 (45.0)	3923 (96.7)	6063 (70.3)	12190 (69.4)
Positive	510 (10)	132 (3)	629 (7)	1271 (7.2)
HIV status unknown	2183 (44.6)	0 (0)	1928 (22.4)	4111 (23.4)
If HIV positive, on ART = Yes (%)	7 (1)	65 (49)	76 (12)	148 (12)

Table 4.1. Continued.

	3HP n = 4897	4R n = 4055	6-9H n = 8620	Overall n = 17572
Smoking status (%)				
Never	3365 (68.7)	2895 (71.4)	6030 (70.0)	12290 (69.9)
Current	1184 (24.2)	509 (12.6)	1601 (18.6)	3294 (18.7)
Ever	345 (7)	284 (7)	624 (7)	1253 (7.1)
NA	3 (0)	367 (9)	365 (4)	735 (4)
Alcohol use (%)				
Never	2381 (48.6)	128 (3.2)	2278 (26.4)	4787 (27.2)
Current	81 (2)	128 (3)	230 (3)	439 (3)
Ever	2050 (41.9)	8 (0)	2014 (23.4)	4072 (23.2)
NA	385 (8)	3791 (93.5)	4098 (47.5)	8274 (47.1)
Recreational drug use (%)				
No	4731 (96.6)	3387 (83.5)	7809 (90.6)	15927 (90.6)
Yes	157 (3)	57 (1)	198 (2)	412 (2)
NA	9 (0)	611 (15)	613 (7)	1233 (7.0)
TST performed = Yes (%)	4632 (94.5)	4010 (98.9)	8316 (96.4)	16958 (96.5)
IGRA performed = Yes (%)	132 (3)	481 (12)	594 (7)	1207 (6.9)
Chest x-ray result at baseline (%)				
Normal	4310 (88.0)	3195 (78.8)	7230 (83.9)	14735 (83.9)
Abnormal	253 (5)	395 (10)	620 (7)	1268 (7)
Abnormal Not TB	0 (0)	306 (8)	300 (4)	606 (3)
NA	334 (7)	159 (4)	470 (6)	963 (6)

NA: data not available. IPD: individual patient data. Close contact defined as those spending ≥4 hours per week of contact with confirmed active tuberculosis case.

Table 4.2. Adjusted risk ratio (aRR) and risk difference (aRD) with their 95% confidence intervals (CI) from both direct and network meta-analysis models for the comparison of treatment completion between 3HP and 4R.

	Intervention		Comparator		Direct IPD-MA†	IPD-NMA†	IPD-NMA†
	Completing n / N	%	Completing n / N	%	aRR & aRD (95% CI)	3HP vs 4R aRR & aRD (95% CI)*	Only studies of 9H§ aRR & aRD (95% CI)*
3HP vs 6-9H	**3HP**		**6-9H**		**3HP vs 6-9H**		
CDC Study 26	3545 / 4437	79.9	2609 / 4156	62.8	**aRR**		
Martinson	300 / 328	92	143 / 327	44	1.30 (1.24, 1.37)		
Sun	118 / 132	89	104 / 131	79	**aRD**	**aRR**	**aRR**
Total	**3963 / 4897**	**80.9**	**2856 / 4614**	**61.9**	0.19 (0.17, 0.21)	1.06 (1.02, 1.10)	1.02 (0.98, 1.07)
4R vs 9H	**4R**		**9H**		**4R vs 6-9H**	**aRD**	**aRD**
Menzies	2476 / 3443	71.9	1965 / 3416	57.5	**aRR**	0.05 (0.02, 0.07)	0.03 (0.00, 0.06)
Diallo	352 / 422	83	305 / 407	75	1.23 (1.17, 1.30)		
Total	**2828 / 3865**	**73.2**	**2270 / 3823**	**59.4**	**aRD** 0.14 (0.12, 0.16)		

† Risk ratios and risk differences adjusted for age, sex, body mass index category, diabetes, smoking, HIV infection, and alcohol use. Note: cannot adjust for contact/close contact, recreational drug use, or use of ART. § Martinson et al. removed (only study with 6H arm); no study with 6H arm included for 4R comparison of treatment completion. *Confidence intervals (CIs) were estimated with bootstrap resampling methods on 1000 replications and calculated using the 2.5th and 97.5th percentiles of the sampling distribution. Treatment completion defined as taking >80% of prescribed doses in 120% of allowed time.

Table 4.3. Adjusted risk ratios (aRR) and risk differences (aRD) with their 95% confidence intervals (CI) from direct and network meta-analysis models for the comparison of the incidence of treatment-related adverse events that led to permanent drug discontinuation between 3HP and 4R in the safety population.

	Intervention		Comparator		Direct IPD-MA†	IPD-NMA†
	Events n / N	%	Events n / N	%	aRR &a RD (95% CI)	3HP vs 4R aRR & aRD (95% CI)**
Any treatment-related adverse event that led to permanent drug discontinuation*						
3HP vs 6-9H	**3HP**		**6-9H**			
CDC Study 26	247/4343	5.7	170/4066	4.2	**aRR**: 1.37	
Martinson	0/328	0	2/326	1	(1.13 , 1.66)	
Sun	12/132	9	7/131	5	**aRD**: 0.01	**aRR**: 2.86 (2.12, 4.21)
Total	259/4803	5.4	179/4523	4.0	(0.01 , 0.02)	**aRD**: 0.03 (0.02, 0.05)
4R vs 6-9H	**4R**		**6-9H**			
Chan	2/190	1	13/183	7	**aRR**: 0.48	
Menzies	68/3281	2.1	131/3231	4.1	(0.36 , 0.63)	
Diallo	0/420	0	0/397	0	**aRD**: -0.02	
Total	70/3891	1.8	144/3811	3.8	(-0.03 , -0.01)	
Treatment-related Grade 3 or 4 adverse events that led to permanent drug discontinuation*						
3HP vs 6-9H	**3HP**		**6-9H**			
CDC Study 26	104/4343	2.4	75/4066	1.8	**aRR**: 1.24	
Martinson	0/328	0	2/326	1	(0.93 , 1.66)	
Sun	2/132	2	4/131	3	**aRD**: 0.00	**aRR**: 3.46 (2.09, 6.17)
Total	106/4803	2.2	81/4523	1.8	(0.00 , 0.01)	**aRD**: 0.02 (0.01, 0.03)
4R vs 6-9H	**4R**		**6-9H**			
Chan	2/190	1	13/183	7	**aRR**: 0.36	
Menzies	29/3281	0.9	72/3231	2.2	(0.24 , 0.54)	
Diallo	0/420	0	0/397	0	**aRD**: -0.01	
Total	31/3891	0.8	85/3811	2.2	(-0.02, -0.01)	

† Risk ratios and risk differences adjusted for age, sex, body mass index category, and HIV infection.
*Judged to be possibly, probably, or definitely related to study drug in primary studies with harmonization conducted for meta-analysis. **Confidence intervals were estimated with bootstrap resampling methods on 1000 replications and calculated using the 2.5th and 97.5th percentiles of the sampling distribution.

Table 4.4. Adjusted incidence rate ratios (aIRR) and incidence rate differences (aIRD) with their 95% confidence intervals (CI) from both direct and network meta-analysis models for the comparison of incidence rate of active tuberculosis between 3HP and 4R.

	Intervention			Comparator			Direct IPD-MA†	IPD-NMA†
	n	TB events per person years	Rate per 1000 Person years	n	TB events per person years	Rate per 1000 Person years	aIRR & aIRD (95% CI)	3HP vs 4R aIRR & aIRD (95% CI)*
3HP vs 6 - 9H		**3HP**			**6 - 9H**			
CDC Study 26	4437	14 / 11326	1	4156	22 / 10511	2	**aIRR** 0.84 (0.54, 1.25)	**aIRR** 1.16 (0.40, 3.58)
Martinson	328	28 /1167	24	327	24 / 1129	21	**aIRD per 1000** -0.1 (-0.4, 0.2)	**aIRD per 1000** 0.8 (-2.3, 7.0)
Sun§	132	0 / 289	0	131	0 / 344	0		
Total	**4897**	**42 / 12782**	**3**	**4614**	**46 / 11984**	**4**		
4R vs 9H		**4R**			**9H**			
Menzies	3443	8 / 7986	1	3416	9 / 7908	1	**aIRR** 0.72 (0.29, 1.79)	
Diallo	422	0 / 546	0	407	2 / 523	4	**aIRD per 1000** -0.9 (-3.9, 2.0)	
Total	**3865**	**8 / 8532**	**1**	**3823**	**11 / 8431**	**1**		

† Incidence rate ratios and differences adjusted for age, sex, body mass index category, HIV infection, and TST size category. *Confidence intervals were estimated with bootstrap resampling methods on 1000 replications and calculated using the 2.5th and 97.5th percentiles of the sampling distribution. § Incidence of tuberculosis not reported in publication, but additional data provided by study authors.

Section 4.3 Supplemental material

Completion, safety, and efficacy of tuberculosis preventive treatment regimens containing rifampicin or rifapentine: an individual patient data network meta-analysis.

Supplement 4.1. PubMed search terms

To identify data sets from randomized controlled trials (RCTs) that we could potentially include in our individual patient data set, we searched PubMed between Jan 1, 2000 (as individual-level patient data published earlier would be difficult to locate and obtain) and Mar 1, 2019, to identify RCTs published in the peer-reviewed literature using the following keywords and medical subject headings (we did not specify any search limits besides date range above):

(latent tb[title/abstract] OR latent tuberculos*[title/abstract] OR LTBI[title/abstract] OR LTB*[title/abstract] or latent mycobacterium[title/abstract])

AND (treatment[title/abstract] OR safety[title/abstract] OR completion[title/abstract] OR adherence[title/abstract] OR activation[title/abstract] OR adverse event*[title/abstract])

AND (randomized controlled trial[title/abstract] OR RCT[title/abstract] OR controlled trial[title/abstract] OR trial[title/abstract] OR control trial[title/abstract] OR randomized[title/abstract])

AND (rif*[title/abstract] OR rifampin[title/abstract] OR rifamycin[title/abstract] or Rifampin[title/abstract] OR rifapentine[title/abstract] OR 3HP[title/abstract] OR INH[title/abstract] or isoniazid[title/abstract] OR 3HR[title/abstract] OR (rifampin plus isoniazid[title/abstract]) OR (rifampin plus[title/abstract]) OR (rifamycin plus[title/abstract]))

Supplement 4.2. Detail on data requested from collaborators for inclusion in the IPD.

i) baseline characteristics: age, sex, height (m), weight (kg) and/or BMI, race/ethnicity, country, TST reaction size or IGRA positivity, chest x-ray results (to exclude active TB), and comorbidity (diabetes, immune-suppressive conditions, liver disease, etc.);

ii) risk factors and indication for treatment: HIV status (including ARV use), close/casual contact of confirmed active TB, recent conversion to positive TST, smoking status, alcohol use, history of incarceration, injection drug use, and biological treatment for immune or chronic inflammatory disorders;

iii) treatment: drug(s) used, duration (months/days, start and stop dates), dosage, duration of follow-up, and deviations from protocol (per-protocol, intention to treat (ITT), modified-ITT); and

iv) outcomes: completion of treatment (pill counts), drug related AE (any, grade 3-4 and drug discontinuations that were restarted or permanently stopped), incidence of active TB and person years of follow-up.

Supplement 4.3. Harmonization of adverse events grading systems for adverse events observed across included data.

Note: in CDC Study 26: 23 grade 3 hepatotoxicity AEs were reassigned as grade 4; in Sun set al.: one grade 2 rash reassigned as grade 3, one grade 3 fever reassigned as grade 2

	Grade 1	Grade 2	Grade 3	Grade 4
GI intolerance **Perm Stop GII gastro** **STOMACH PAINS** **Gastrointestinal Intolerance**				
CTCAE - nausea (v5/v4)	Loss of appetite without alteration in eating habits	Oral intake decreased without significant weight loss, dehydration or malnutrition	Inadequate oral caloric or fluid intake; tube feeding, TPN, or hospitalization indicated	-
CTC - nausea	able to eat	oral intake significantly decreased	no significant intake, requiring IV fluids	-
DAIDS - nausea	Transient (< 24 hours) or intermittent AND No or minimal interference with oral intake	Persistent nausea resulting in decreased oral intake for 24 to 48 hours	Persistent nausea resulting in minimal oral intake for > 48 hours OR Rehydration indicated (e.g., IV fluids)	Life-threatening consequences (e.g., hypotensive shock)
Overall **Harmonized to CTCAE**	**Same**	**Similar**	**Similar**	**G4 in DAIDS = G3 in CTCAE** **if G4 in DAIDS then G3 in IPD**

CTCAE - vomiting v4	1 - 2 episodes (separated by 5 minutes) in 24 hrs	3 - 5 episodes (separated by 5 minutes) in 24 hrs	>=6 episodes (separated by 5 minutes) in 24 hrs; tube feeding, TPN or hospitalization indicated	Life-threatening consequences
CTCAE - vomiting v5	Intervention not indicated	Outpatient IV hydration; medical intervention indicated	Tube feeding, TPN, or hospitalization indicated	Life-threatening consequences requiring parenteral nutrition; or physiologic consequences requiring intensive care; hemodynamic collapse
CTC - vomiting	1 episode in 24 hours over pretreatment	2-5 episodes in 24 hours over pretreatment	>=6 episodes in 24 hours over pretreatment; or need for IV fluids	
DAIDS - vomiting	Transient or intermittent AND No or minimal interference with oral intake	Frequent episodes with no or mild dehydration	Persistent vomiting resulting in orthostatic hypotension OR Aggressive rehydration indicated (e.g., IV fluids)	Life-threatening consequences (e.g., hypotensive shock
Overall Harmonized to CTCAE	**Same**	**Same**	**Same**	**Similar**
Hematologic				
CTCAE - thromboembolic event v5	Medical intervention not indicated (e.g., superficial thrombosis)	Medical intervention indicated	Urgent medical intervention indicated (e.g., pulmonary	Life-threatening consequences with hemodynamic or neurologic instability

			embolism or intracardiac thrombus)	
CTCAE - thromboembolic event V4	Venous thrombosis (e.g., superficial thrombosis)	Venous thrombosis (e.g., uncomplicated deep vein thrombosis), medical intervention indicated	Thrombosis (e.g., uncomplicated pulmonary embolism [venous], non-embolic cardiac mural [arterial] thrombus), medical intervention indicated	Life-threatening (e.g., pulmonary embolism, cerebrovascular event, arterial insufficiency); hemodynamic or neurologic instability; urgent intervention indicated
CTC - Thrombosis/embolism	-	deep vein thrombosis, not requiring anticoagulant	deep vein thrombosis, requiring anticoagulant therapy	embolic event including pulmonary embolism
DAIDS - Thrombosis or embolism	NA	Symptoms AND No intervention indicated	Symptoms AND Intervention indicated	Life-threatening embolic event (e.g., pulmonary embolism, thrombus)
Overall		**G2 CTCAE = G3 CTC/DAIDS**	**G3 CTCAE = G4 CTC/DAIDS**	**Similar**
Harmonized to CTCAE	**G1**	**if G3 CTC/DAIDS then G2 in IPD**	**if G4 in CTC/DAIDS then G3 in IPD**	**G4**

CTCAE - anemia (v5/v4)	Hemoglobin (Hgb) <LLN - 10.0 g/dL; <LLN - 6.2 mmol/L; <LLN - 100 g/L	Hgb <10.0 - 8.0 g/dL; <6.2 - 4.9 mmol/L; <100 - 80g/L	Hgb <8.0 g/dL; <4.9 mmol/L; <80 g/L; transfusion indicated	Life-threatening consequences; urgent intervention indicated
CTC - Hemoglobin	<LLN - 10.0 g/dL <LLN - 100 g/L <LLN - 6.2 mmol/L	8.0 - <10.0 g/dL 80 - <100 g/L 4.9 - <6.2 mmol/L	6.5 - <8.0 g/dL 65 - <80 g/L 4.0 - <4.9 mmol/L	<6.5 g/dL <65 g/L <4.0 mmol/L
DAIDS - Hemoglobin (g/dL; mmol/L)	10.0 to 10.9 6.19 to 6.76	9.0 to < 10.0 5.57 to < 6.19	7.0 to < 9.0 4.34 to < 5.57	< 7.0 < 4.34
Overall Harmonized to CTCAE	**Same G1**	**G2 DAIDS = G1 CTC/CTCAE if G2 DAIDS then G1 in IPD**	**G3 DAIDS = G2 CTC/CTCAE if G3 in DAIDS then G2 in IPD**	**G4 DAIDS = G3 CTC/CTCAE if G4 DAIDS then G3 in IPD**
CTCAE - Neutrophil count decreased (v4/5)	<LLN - 1500/mm3; <LLN - 1.5 x 10e9 /L	<1500 - 1000/mm3; <1.5 - 1.0 x 10e9 /L	<1000 - 500/mm3; <1.0 - 0.5 x 10e9 /L	<500/mm3; <0.5 x 10e9 /L
CTC - Neutrophils/granulocytes	>=1.5 - <2.0 x 109 /L >=1500 - <2000/mm3	>=1.0 - <1.5 x 109 /L >=1000 - <1500/mm3	>=0.5 - <1.0 x 109 /L >=500 - <1000/mm3	<0.5 x 109 /L <500/mm3
DAIDS - Absolute Neutrophil Count (ANC), Low (cells/mm3; cells/L)	800 to 1,000 0.800 x 10^9 to 1.000 x 10^9	600 to 799 0.600 x 10^9 to 0.799 x 10^9	400 to 599 0.400 x 10^9 to 0.599 x 10^9	< 400 < 0.400 x 10^9
Overall Harmonized to CTCAE	**G1 in DAIDS = G3 in CTCAE/CTC**	**G2 in DAIDS = G3 in CTCAE/CTC**	**G3 in DAIDS = G3 in CTCAE/CTC**	**Same**
		ALL ARE IN MENZIES SO GRADE IS CTCAE		
death Death DEATH				
CTCAE				Grade 5
CTC				Grade 5
DAIDS				Grade 5
Overall				**Grade 5**

Harmonized to CTCAE				Grade 5
Dizziness				
CTCAE v4/v5	Mild unsteadiness or sensation of movement	Moderate unsteadiness or sensation of movement; limiting instrumental ADL	Severe unsteadiness or sensation of movement; limiting self care ADL	-
CTC	not interfering with function	interfering with activities of daily living	interfering with activities of daily living	Bedridden or disabling
DAIDS	No Entry	No Entry	No Entry	No Entry
Overall	**same**	**different**	**same**	**different**
Harmonized to CTCAE		**ALL ARE IN MENZIES SO GRADE IS CTCAE**		
Drug induced pancreatitis				
CTCAE (v5/v4 [pancreatitis])	-	Enzyme elevation; radiologic findings only	Severe pain; vomiting; medical intervention indicated (e.g., analgesia, nutritional support)	Life-threatening consequences; urgent intervention indicated
CTC	-	-	abdominal pain with pancreatic enzyme elevation	complicated by shock (acute circulatory failure)
DAIDS	NA	Symptoms with hospitalization not indicated	Symptoms with hospitalization indicated	Life-threatening consequences (e.g., circulatory failure, hemorrhage, sepsis)
Overall	**same**	**CTCAE Grade 2=CTC Grade 3**	**CTCAE Grade 2=CTC Grade 3**	**same**

Harmonized to CTCAE	ALL ARE IN MENZIES SO GRADE IS CTCAE			
FATIGUE AND MUSCLE PAIN **SYNDROME MYALGIA, CHEST PAIN, VOMITING TIRED AND VERY WEAK FOR 4 DAYS**				
CTCAE - v4/v5	Grade 1 Fatigue relieved by rest	Grade 2 Fatigue not relieved by rest; limiting instrumental ADL	Grade 3 Fatigue not relieved by rest, limiting self care ADL	-
CTC	Grade 1 increased fatigue over baseline, but not altering normal activities	Grade 2 moderate (e.g., decrease in performance status by 1 ECOG level or 20% Karnofsky or Lansky) or causing difficulty performing some activities	Grade 3 severe (e.g., decrease in performance status by >= 2 ECOG levels or 40% Karnofsky or Lansky) or loss of ability to perform some activities	Grade 4 bedridden or disabling
DAIDS	Grade 1 Symptoms causing no or minimal interference with usual social & functional activities	Grade 2 Symptoms causing greater than minimal interference with usual social & functional activities	Grade 3 Symptoms causing inability to perform usual social & functional activities	Grade 4 Incapacitating symptoms of fatigue or malaise causing inability to perform basic self-care functions
Overall Harmonized to CTCAE	**Same**	**Same**	**Same**	**Same**
FLUSHING, RED EYES, HEADACHE				

RED EYES, PALPITATIONS,DIARRHEA,VOMITING,BURNING FEET, HEADACHE, LOW BLOOD PRESS					
CTCAE v5		Asymptomatic; clinical or diagnostic observations only	Moderate symptoms; limiting instrumental ADL	Symptomatic, associated with hypotension and/or tachycardia; limiting self care ADL	-
CTCAE v4		Asymptomatic; clinical or diagnostic observations only; intervention not indicated	Moderate symptoms; medical intervention indicated; limiting instrumental ADL	Symptomatic, associated with hypotension and/or tachycardia; limiting self care ADL	
CTC - Flushing		Flushing present	-	-	-
DAIDS - Flushing		NA	NA	NA	NA
DAIDS - Headache		Symptoms causing no or minimal interference with usual social & functional activities	Symptoms causing greater than minimal interference with usual social & functional activities	Symptoms causing inability to perform usual social & functional activities	Symptoms causing inability to perform basic self-care functions OR Hospitalization indicated OR Headache with significant impairment of alertness or other neurologic function
	Overall Harmonized to CTCAE	**Similar**	**Similar**	**Similar**	**Similar**

Hepatotoxicity **HEPATOTOXICITY** **HEPATOTOXICITY; EVENT CAUSING PERMANENT DISCONTINUATION OF STUDY DRUGS** **HEPATOTOXICITY; GRADE 3 OR 4 TOXICITY DURING OR WITHIN 60 DAYS OF STUDY THERAPY** **Perm Stop GII hepatotox** **Perm Stop GIII or GIV hepatotox** **RISE IN AST** **INCREASED LFTS ON 8/3 AST 75 ALT 172 RUQ DISCOMFORT** **OTHER MEDICAL CONDITION; GRADE 3 OR 4 TOXICITY DURING OR WITHIN 60 DAYS OF STUDY THERAPY** **JAUNDICE, MALAISE**				
CTCAE (definition used in protocol)	ALT/AST 1 to ≤ 3 times upper limit of normal (ULN) plus symptoms as above OR ALT/AST 1 to ≤ 5 times ULN and no symptoms		ALT/ AST 3 to ≤ 10 times ULN plus symptoms as above OR ALT/AST 5 to ≤ 10 ULN and no symptoms.	ALT or AST > 10 times ULN.
CTC - SGOT(AST and ALT)	>ULN - 2.5 x ULN	>2.5 - 5.0 x ULN	>5.0 - 20.0 x ULN	>20.0 x ULN
DAIDS - AST and ALT (same)	1.25 to < 2.5 x ULN	2.5 to < 5.0 x ULN	5.0 to < 10.0 x ULN	≥ 10.0 x ULN
Overall **Harmonized to CTCAE**	**same**	**same**	**CTC Grade3 = CTCAE Grade4** **if CTC and G3 then G4 in IPD**	**CTC Grade3 = CTCAE Grade4**

CTCAE v4/v5 - bilirubin	**>ULN - 1.5 x ULN if baseline was normal; > 1.0 - 1.5 x baseline if baseline was abnormal**	>1.5 - 3.0 x ULN if baseline was normal; >1.5 - 3.0 x baseline if baseline was abnormal	>3.0 - 10.0 x ULN if baseline was normal; >3.0 - 10.0 x baseline if baseline was abnormal	>10.0 x ULN if baseline was normal; >10.0 x baseline if baseline was abnormal
CTC - bilirubin	>ULN - 1.5 x ULN	>1.5 - 3.0 x ULN	>3.0 - 10.0 x ULN	>10.0 x ULN
DAIDS - Total Bilirubin	1.1 to < 1.6 x ULN	1.6 to < 2.6 x ULN	2.6 to < 5.0 x ULN	≥ 5.0 x ULN
Overall	**same**	**same**	**different if G3/G4 in DAIDS then G3 in IPD**	**different**
Harmonized to CTCAE	**same**	**Same**	**Same**	**Same**
MILD ABDOMINAL PAIN RIGORS, ACHING ALL OVER, STOMACH PAIN, VOMITING, BP LOWER THAT USUAL				
CTCAE V4/V5	Mild pain	Moderate pain; limiting instrumental ADL	Severe pain; limiting self care ADL	-
CTC	mild pain not interfering with function	moderate pain: pain or analgesics interfering with function, but not interfering with activities of daily living	severe pain: pain or analgesics severely interfering with activities of daily living	disabling
DAIDS * NO abdominal specific, this is systemic pain	Pain causing no or minimal interference with usual social & functional activities	Pain causing greater than minimal interference with usual social & functional activities	Pain causing inability to perform usual social & functional activities	Disabling pain causing inability to perform basic self-care functions OR Hospitalization indicated

	Overall Harmonized to CTCAE	**same**	**similar**	**same**	**same**
Perm Stop GII rash **Perm Stop GIII or GIV rash** **Rash** **RASH** **rash acneiform**					
CTCAE * rash acneiform V5		Papules and/or pustules covering <10% BSA, which may or may not be associated with symptoms of pruritus or tenderness	Papules and/or pustules covering 10 - 30% BSA, which may or may not be associated with symptoms of pruritus or tenderness; associated with psychosocial impact; limiting instrumental ADL; papules and/or pustules covering > 30% BSA with or without mild symptoms	Papules and/or pustules covering >30% BSA with moderate or severe symptoms; limiting self-care ADL; associated with local superinfection with oral antibiotics indicated	Life-threatening consequences; papules and/or pustules covering any % BSA, which may or may not be associated with symptoms of pruritus or tenderness and are associated with extensive superinfection with IV antibiotics indicated
CTCAE * rash acneiform V4		Papules and/or pustules covering <10% BSA, which may or may not be associated with symptoms of	Papules and/or pustules covering 10-30% BSA, which may or may not be associated with symptoms of	Papules and/or pustules covering >30% BSA, which may or may not be associated with symptoms of pruritus or	Papules and/or pustules covering any % BSA, which may or may not be associated with symptoms of

	pruritus or tenderness	pruritus or tenderness; associated with psychosocial impact; limiting instrumental ADL	tenderness; limiting self care ADL; associated with local superinfection with oral antibiotics indicated	pruritus or tenderness and are associated with extensive superinfection with IV antibiotics indicated; life threatening consequences
CTC *Rash/desquamation	macular or papular eruption or erythema without associated symptoms	macular or papular eruption or erythema with pruritus or other associated symptoms covering <50% of body surface or localized desquamation or other lesions covering <50% of body surface area	symptomatic generalized erythroderma or macular, papular orvesicular eruption or desquamation covering ³50% of body surface area	generalized exfoliative dermatitis or ulcerative dermatitis
DAIDS *rash general	Localized rash	Diffuse rash OR Target lesions	Diffuse rash AND Vesicles or limited number of bullae or superficial ulcerations of mucous membrane limited to one site	Extensive or generalized bullous lesions OR Ulceration of mucous membrane involving two or more distinct mucosal sites OR Stevens-Johnson

					syndrome OR Toxic epidermal necrolysis
	Overall **Harmonized to CTCAE**	**Same**	**DAIDS different if G2 in DAIDS then G3 in IPD**	**DAIDS different**	**DAIDS different**
PRURITIC RASH **PRURITIS, NECK PAIN**					
CTCAE V4/V5		Mild or localized; topical intervention indicated	Widespread and intermittent; skin changes from scratching (e.g., edema, papulation, excoriations, lichenification, oozing/crusts); oral intervention indicated; limiting instrumental ADL	Widespread and constant; limiting self care ADL or sleep; systemic corticosteroid or immunosuppressive therapy indicated	-
CTC		mild or localized, relieved spontaneously or by local measures	intense or widespread, relieved spontaneously or by systemic measures	intense or widespread and poorly controlled despite treatment	-

DAIDS	Itching causing no or minimal interference with usual social & functional activities	Itching causing greater than minimal interference with usual social & functional activities	Itching causing inability to perform usual social & functional activities	-
Overall Harmonized to CTCAE	**Same**	**Same**	**Same**	**Same**
SUICIDAL IDEATION, VOMITING, NEUROPATHY				
CTCAE V4/V5	Increased thoughts of death but no wish to kill oneself	Suicidal ideation with no specific plan or intent	Specific plan to commit suicide without serious intent to die which may not require hospitalization	Specific plan to commit suicide with serious intent to die which requires hospitalization
CTC * mood alteration-anxiety/depression	mild mood alteration not interfering with function	moderate mood alteration interfering with function, but not interfering with activities of daily living	severe mood alteration interfering with activities of daily living	suicidal ideation or danger to self
DAIDS	Preoccupied with thoughts of death AND No wish to kill oneself	Preoccupied with thoughts of death AND Wish to kill oneself with no specific plan or intent	Thoughts of killing oneself with partial or complete plans but no attempt to do so OR Hospitalization indicated	Suicide attempted
Overall Harmonized to CTCAE	**same**	**similar**	**similar**	**similar**

Fever				
CTCAE V4/V5	38.0 - 39.0 degrees C (100.4 - 102.2 degrees F)	>39.0 - 40.0 degrees C (102.3 - 104.0 degrees F)	>40.0 degrees C (>104.0 degrees F) for <=24 hrs	>40.0 degrees C (>104.0 degrees F) for >24 hrs
CTC	38.0 - 39.0°C (100.4 - 102.2°F)	39.1 - 40.0°C (102.3 - 104.0°F)	>40.0°C (>104.0°F) for <24hrs	>40.0°C (>104.0°F) for >24hrs
DAIDS	38.0 to < 38.6°C or 100.4 to < 101.5°F	≥ 38.6 to < 39.3°C or ≥ 101.5 to < 102.7°F	≥ 39.3 to < 40.0°C or ≥ 102.7 to < 104.0°F	≥ 40.0°C or ≥ 104.0°F
Overall Harmonized to CTCAE	**same**	**similar if G3 in DAIDS then G2 in IPD**	**G4 DAIDS = G3 CTCAE if G4 in DAIDS then G3 in IPD**	**similar**

CTCAE: Common Terminology Criteria for Adverse Events Versions 4 or 5; CTC: Common toxicity criteria version 2.0; DAIDS: the Division of AIDS Table for Grading the Severity of Adult and Pediatric Adverse Events, Corrected Version 2.1

Supplement 4.4. Drug dosage and duration for each regimen analyzed.

Study	Regimen	Doses	Dosage maximum (by weight)	Treatment days	80% of pills	120% of expected time (days)
CDC Study 26[14-16,28]	3HP	12	900mg rifapentine (incremental adjustment for those ≤50kg) + 900mg isoniazid (15 to 25 mg/kg) once weekly	90	10	108
	9H	270	300mg (5 to 15 mg/kg) daily	270	216	324
Martinson et al.[12]	3HP	12	900mg rifapentine + 900mg isoniazid once weekly	90	10	108
	6H	180	300mg daily	180	144	216
Sun et al.[27]	3HP	12	900mg rifapentine (750 mg/week for 32.1 to 50 kg; 600mg/week for 25.1 to 32 kg; 450 mg/week for 14.1to 25.0 kg) + 900mg isoniazid (15 mg/kg) once weekly	90	10	108
	9H	270	300mg (5 mg/kg).daily	270	216	324
Menzies et al.[13]	4R	120	600mg (10 mg/kg) daily	120	96	144
	9H	270	300mg (5 mg/kg) daily	270	216	324
Diallo et al.[11]	4R	120	600mg (15 to 20 mg/kg) daily	120	96	144
	9H	270	300mg (10 to 15 mg/kg) daily	270	216	324
Chan et al.[33]	4R	120	600mg (10 mg/kg) daily	120	96	144
	6H	180	300mg (5 mg/kg) daily	180	144	216

Note Chan et al. not included in analyses of treatment completion as individual data was not available.

Supplement 4.5. Detailed power calculations

For treatment completion:

To our knowledge there is no method for calculating power for an IPD network meta-analysis. There are power calculations for direct comparisons and for indirect comparisons, but not a combination of both.

Hence, to assess power we used two approaches. *1)* we assessed sample size as if we were conducting a direct head-to-head comparison using the completion from the Menzies et al. 4R trial and from the Martinson et al. 3HP trial (and hypothetical scenarios) that account for clusters; and *2)* we assessed the power we would have to detect these rates in an indirect comparison using a network meta-analysis (NMA), by calculating the effective sample size (if the effective sample size is $\geq$ sample size for 80% power in a direct comparison, then we can claim to have the same power for an indirect comparison). Martinson et al. was chosen as they reported the highest rate of treatment completion in 3HP of the identified studies, while Menzies had the lowest rate of completion of 4R. Using the extreme values of completion would provide the most conservative estimates of power.

1) Direct power calculation (80% power, alpha 0.05, 10 clusters) :

Actual and hypothetical scenarios for the sample sizes required to detect differences in treatment completion with a power 80% (alpha 0.05, 10 clusters):

Study	INH completion	4R/3HP completion	difference	ICC	sample size
4R Menzies	0.63*	0.78*	0.15	0.004**	**2693**
3HP Martinson	0.83*	0.95*	0.13	0.004**	**1858**
Hypothetical	0.63	0.78	0.15	0.001	954
Hypothetical	0.63	0.78	0.15	0.0005	618
Hypothetical	0.83	0.95	0.13	0.001	597
Hypothetical	0.83	0.95	0.13	0.0005	388
Hypothetical	0.95	0.85	0.10	0.004	2880
Hypothetical	0.95	0.90	0.05	0.004	9016
Hypothetical	0.95	0.90	0.05	0.001	2902
Hypothetical	0.94	0.95	0.01	0.001	54785
Hypothetical	0.94	0.95	0.01	0.0005	35547

* Actual completion from these two studies
** Actual ICC calculated from Menzies et al. 2018

The above sample sizes are required for sufficient power to detect a difference of 13% to 15% in completion between arms.

2) Indirect power and effective sample size:

To calculate the "effective sample size" for an NMA for indirect effects, we use the following formula:

$$[N_{3HPv9INH}*(1-ICC)*N_{4Rv9INH}*(1-ICC)] / [N_{3HPv9INH*}(1-ICC) +N_{4Rv9INH}*(1-ICC)]$$

Here we can basically ignore the ICC part, since we expect to have a very small ICC of 0.001

We received a total of **17,572 participants**. For treatment completion, 17,119 were assessed.

The effective sample size we would have for treatment completion is:

[(9511)*(7688)) / ((9511)+(7688)] = **4251**

Our effective sample size for an NMA is larger than the required sample size for some direct comparisons, and so we should have more than enough power even for an indirect comparison in an NMA to detect a difference of 13-15% in treatment completion. However, if we expect the ICC to be low (very similar across sites) we may have power for differences of 5%, but likely not lower.

For adverse events:

Using the same 2 approaches above we have the following sample sizes for direct comparisons:

Actual and hypothetical scenarios for the sample sizes required to detect differences in adverse events with a power 80% (alpha 0.05, 10 clusters):

Study	INH proportion	4R/3HP proportion	difference	ICC	sample size
4R Menzies	0.03	0.012	0.018	0.0005	**4328**
3HP Martinson	0.043	0.012	0.031	0.0005	**1887**
4R Menzies	0.03	0.012	0.018	**0.004**	**20,725**
3HP Martinson	0.043	0.012	0.031	**0.004**	**81,505**
Hypothetical	0.04	0.01	0.03	0.0005	1893
Hypothetical	0.04	0.01	0.03	0.001	2835
Hypothetical	0.04	0.01	0.03	0.004	8807
Hypothetical	0.02	0.01	0.01	0.0005	10,095
Hypothetical	0.02	0.01	0.01	0.001	15,558
Hypothetical	0.02	0.01	0.01	0.004	48,339
Hypothetical	0.045	0.01	0.035	0.0005	1477
Hypothetical	0.045	0.01	0.035	0.001	2277
Hypothetical	0.045	0.01	0.035	0.004	7075

*Martinson et al. are proportions for Grade 4 AE

If we can assume an ICC of 0.0005, we will have sufficient power for indirect comparison in an NMA (effective sample size for AEs is 4363), but realistically the ICC will be around 0.001-0.004. And so, with our effective sample size we will not be powered for indirect comparisons when ICC is 0.004.

For active TB:

We will not be powered to detect a difference in rate of active TB, as these rates are very small in each group. Thus, all analyses of active TB will be considered exploratory only. However, the main questions involve safety and completion, and we should be able to conduct analyses for these outcomes.

Supplement 4.6. Detail on model fit and covariate adjustments.

For continuous variables, we assessed their functional forms by fitting natural cubic splines and assessed model fit using Akaike information criterion, Bayesian information criterion, and likelihood ratio tests. All continuous variables showing no deviation from linearity were included as linear functions. We had different adjustment sets for each outcome:

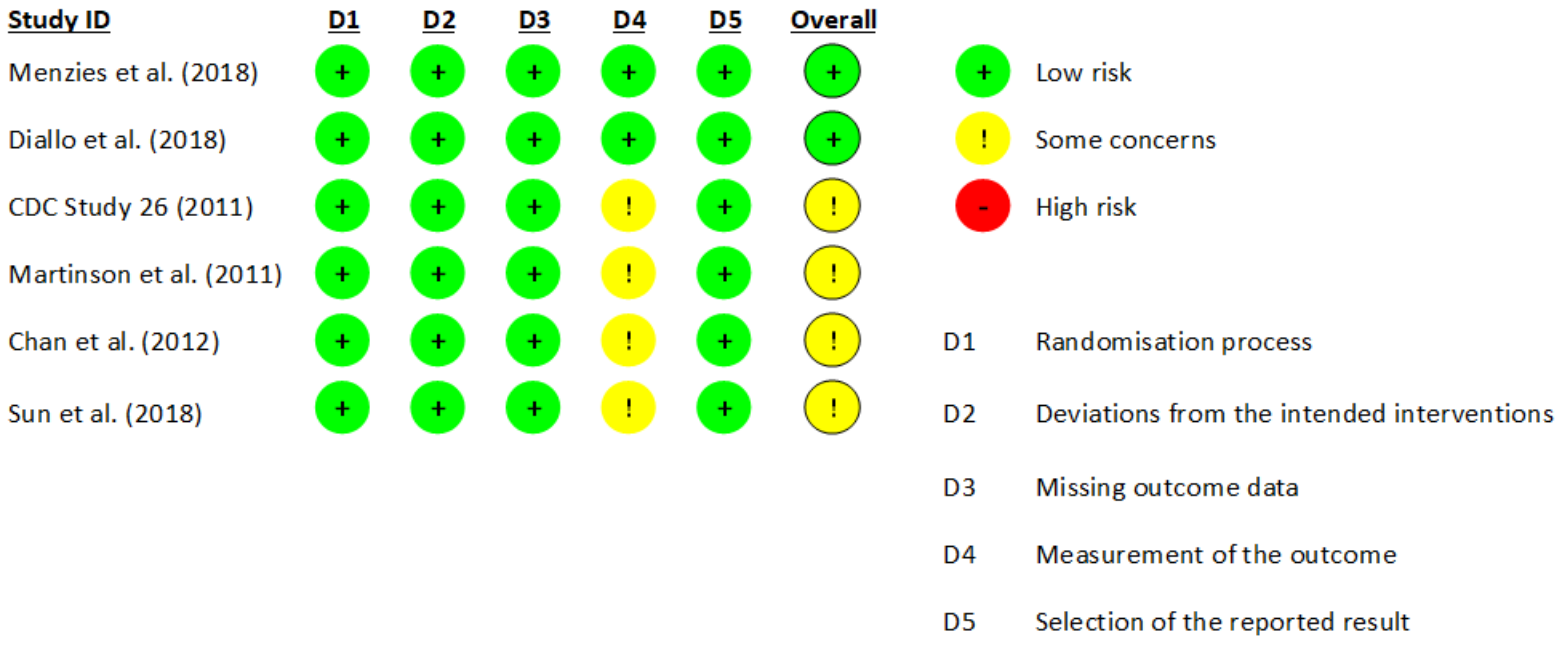
Treatment completion: age (continuous), sex, body mass index category (underweight: <18.5 kg/m²; normal: ≥18.5 and <25 kg/m²; overweight: ≥25 and <30 kg/m²; and obese:≥30 kg/m²), diabetes (yes/no), smoking status (never, current, and ever), HIV (yes/no), and alcohol use (never, current, and ever)

Adverse events: age, sex, body mass index category, and HIV (yes/no)

Efficacy for prevention of tuberculosis disease: age (continuous), sex, body mass category, HIV (yes/no) and TST size category (≤10 mm; >10 to ≤15 mm; >15 to ≤20 mm; and >20 mm).

Missing observations for categorical variables were included as a 'not available' (NA) category.

Figure 4.S1. Risk of Bias (RoB2) Tool assessment for included studies.



Supplemental Table 4.S1. Study-level characteristics of trials included in the IPD.

Author (year)	Country; WHO region	Study drug and sample size	Administration of study drugs	Age mean (SD)	PLHIV (%)	Outcomes assessable in IPD
CDC Study 26 (2011)[14-16,28]	USA, Canada, Brazil, China, Spain; AMR,EUR, WPR	3HP: n = 4437 9H: n = 4156	3HP: DOT 9H: self-administered	35 (15.7)	4.0%	Completion Adverse events Active TB
Martinson (2011)[12]‡	South Africa; AFR	3HP: n = 328 6H: n = 327	3HP: DOT 6H: self-administered	31 (6.7)	100%	Completion Adverse events Active TB
Sun (2018)[27]	Taiwan; WPR	3HP: n = 132 9H: n = 131	3HP: DOT 9H: DOT	32 (15.7)	0%	Completion Adverse events Active TB*
Chan (2012)[33]	Taiwan; WPR	4R: n = 190 6H: n = 183	4R: DOT 6H: DOT	42 (10.8)	0%	Adverse events only†
Menzies (2018)[13]	Australia, Benin, Brazil, Canada, Ghana, Guinea, Indonesia, Saudi Arabia, and South Korea; WPR, AMR,AFR, EMR, SEAR	4R: n = 3443 9H: n = 3416	4R: Self-administered 9H: Self-administered	38 (13.7)	4%	Completion Adverse events Active TB
Diallo (2018)[11]	Australia, Benin, Brazil, Canada, Ghana, Guinea, and Indonesia; WPR, AFR, AMR, SEAR	4R: n = 422 9H: n = 407	4R: Self-administered 9H: Self-administered	9.4 (4.7)	0%	Completion Adverse events Active TB

DOT: directly observed therapy; 3HP: 3 months of rifapentine + isoniazid; 4R: 4 months of rifampicin; H: isoniazid; SD: standard deviation. WHO regions: AFR: African Region; AMR: Regions of the Americas; EMR: Eastern Mediterranean Region; EUR: European Region; SEAR: South-East Asian Region; WPR: Western Pacific Region. ‡ Included 329 subjects on 3 months of rifampicin+isoniazid (3HR) but were not analyzed as no other data was obtained for this regimen. * Authors provided additional data on active TB not reported in their publication. † Authors did not provide individual patient data to harmonize completion outcomes, and thus excluded from completion analysis. In all studies all participants underwent chest radiography to exclude tuberculosis disease, and sputum cultures if necessary

Supplemental Table 4.S2. Study-level characteristics and outcomes of trials excluded from the IPD.

Author (year, reference)	Country	Study drug and sample size	Administration of study drugs	Age mean (SD)	HIV (%)	Treatment outcomes		
						Treatment completion (definition) & results	Adverse events	Active TB
Belknap [iAdhere] (2017)[34]	USA, Spain, Hong Kong, South Africa	3HP DOT: n = 337 3HP SAT: n = 337 3HP SAT with reminders n = 328	3HP: DOT 3HP: SAT 3HP: SAT with reminders	36 [27-49]	1.1%	11 of 12 doses in 16 weeks: 3HP DOT 87.2% 3HP SAT: 74.0% 3HP SA with reminders: 76.4%	**Grade 3 or 4 drug stop:** 3HP DOT: 0.9%; 3HP SAT: 3%; 3HP SAT with reminders 2.1% **Grades 3 or 4 drug related:** 3HP DOT: 2.4%; 3HP SAT: 3%; 3HP SAT with reminders 4.3%	NR
White (2012)[36]	USA	4R: n = 180 9H (twice weekly): n = 184	4R: SAT 9H: SAT	71% <35 years	NR	4R: 100% in 6 months 9H: 100% in 12months 4R: 33%* 9H: 26%	Any Grade 3 or higher 4R: 1.7% 9H: 3.3%	NR
HKCS (1992)[35]	Hong Kong	3HR: n = 167 6H: n = 173	4R: SAT 6H: SAT	76% 45 to 64 years	NR	Not Reported	Any drug stopped: 3HR: 5% 6H: 5%	3HR: 16% 6H: 14%
Whalen (2001)[26]	Uganda	3HR: n = 556 6H: n = 536	3HR: SAT 6H: SAT	3HR: 29 (NR) 6H: 29 (NR)	100%	Not Reported	Moderate AEs 3HR: 1.1% 6H: 0.7%	3HR: 1.32 per 100 PYS 6H: 1.1 per 100 PYS
Rivero (2007)[30]	Spain	3HR: n = 103 6H: n = 108	3HR: SAT 6H: SAT	3HR: 33 [NR] 6H: 31.3 [NR]	100%	≥80% of doses 3HR: 61.0% 6H: 63.9%	Drug stopped: 3HR: 6.7% 6H: 6.4%	3HR: 4.6 per 100 PYS 6H: 3.5 per 100 PYS
Spyridis (2007)[29]	Greece	3HR: n = 474 9H: n = 232	3HR: SAT 9H: SAT	3HR: 8.8 (3.4) 9H: 9.1 (3.7)	NR	Excellent compliance: 3HR 81.8% 9H: 65.5%	No serious adverse events in any group	No participants developed clinical disease

DOT: directly observed therapy; 3HP: 3 months of rifapentine + isoniazid; 4R: 4 months of rifampicin; SD: standard deviation. *In those who remained in jail throughout the study, completion was 79% in 9H and 83% in 4R.

Supplemental Table 4.S3. Entire population: Stratified by age groups, the adjusted risk ratio (aRR) and risk difference (aRD) with their 95% confidence intervals (CI) from both direct and network meta-analysis models for the comparison of treatment completion (80% of expected doses in 120% of allowed time) between 3HP and 4R

	Intervention				Comparator			Direct IPD MA aRR/aRD (95% CI)	Indirect NMA 3HP compared to 4R aRR/aRD (95% CI)**
Study	**n**	**N**	**%**		**n**	**N**	**%**		
Those <18 years of age†									
3HP vs 6-9H	**3HP**				**6-9H**				
CDC Study 26	446	553	81		335	506	66	aRR: 1.24 (1.08, 1.42) aRD: 0.16 (0.11 , 0.21)	aRR: 1.12 (1.01, 1.23) aRD: 0.07 (0.00, 0.15)
Sun	17	17	100		16	20	80		
3HP Total	463	570	81		351	526	67		
4R vs 9H	**4R**				**9H**				
Diallo	352	422	83		305	407	75	aRR: 1.11 (0.95 , 1.29) aRD: 0.08 (0.03 , 0.14)	
4R Total	352	422	83		305	407	75		
Those ≥18 years of age §									
3HP vs 6-9H	**3HP**				**6-9H**				
CDC Study 26	3099	3884	79.8		2273	3649	62.3	aRR: 1.32 (1.25 , 1.39) aRD: 0.19 (0.17 , 0.21)	aRR: 1.05 (1.00, 1.10) aRD: 0.05 (0.02, 0.08)
Martinson	300	328	92		143	327	44		
Sun	101	115	88		88	111	79		
3HP Total	3500	4327	81.0		2504	4087	61.3		
4R vs 9H†	**4R**				**9H**				
Menzies	2476	3443	71.9		1965	3416	57.5	aRR: 1.25 (1.18 , 1.33) aRD: 0.15 (0.12 , 0.17)	
4R Total	2476	3443	71.9		1965	3416	57.5		
Those <35 years of age*									
3HP vs 6-9H	**3HP**				**6-9H**				
CDC Study 26	1861	2305	80.7		1365	2220	61.5	aRR: 1.36 (1.27, 1.45) aRD: 0.22 (0.19, 0.24)	aRR: 1.09 (1.02, 1.15) aRD: 0.07 (0.03, 0.11)
Martinson	232	255	91		107	253	42		
Sun	84	92	91		73	90	81		
3HP Total	2177	2652	82.1		1545	2563	60.3		
4R vs 9H	**4R**				**9H**				
Menzies	1157	1667	69.4		883	1653	53.4	aRR: 1.25 (1.16 , 1.35) aRD: 0.15 (0.12, 0.17)	
Diallo	352	422	83		305	407	75		
4R Total	1509	2089	72.2		1188	2060	57.7		

Supplemental Table 4.S3. Continued.

	Intervention				Comparator			Direct IPD MA aRR/aRD (95% CI)	Indirect NMA 3HP compared to 4R aRR/aRD (95% CI)**
Study	**n**	**N**	**%**		**n**	**N**	**%**		
Those who are ≥35 to ≤65 years of age‡									
3HP vs 6-9H	**3HP**				**6-9H**				
CDC Study 26	1571	1989	79.0		1169	1829	63.9	aRR: 1.25 (1.16 , 1.34) aRD: 0.16 (0.13 , 0.19)	aRR: 1.03 (0.97, 1.09) aRD: 0.03 (-0.01, 0.07)
Martinson	68	73	93		36	74	49		
Sun	33	39	85		28	37	76		
3HP Total	1672	2101	79.6		1233	1940	63.6		
4R vs 9H	**4R**				**9H**				
Menzies	1235	1658	74.5		1027	1655	62.1	aRR: 1.2 (1.11 , 1.31) aRD: 0.13 (0.10 , 0.16)	
4R Total	1235	1658	74.5		1027	1655	62.1		
Those who are >65 years of age‡									
3HP vs 6-9H	**3HP**				**6-9H**				
CDC Study 26	113	143	79		74	106	70	aRR: 1.13 (0.84, 1.52) aRD: 0.09 (-0.02 , 0.20)	aRR: 0.83 (0.62, 1.08) aRD: -0.09 (-0.27, 0.07)
Sun	1	1	100		3	4	75		
3HP Total	114	144	79		77	110	70		
4R vs 9H	**4R**				**9H**				
Menzies	84	118	71		55	108	51	aRR: 1.35 (0.96, 1.93) aRD: 0.18 (0.06 , 0.31)	
4R Total	84	118	71		55	108	51		

† For studies of 3HP, a random intercept for study was included while studies of 4R had no random intercepts, all risk ratios and risk differences adjusted for age, sex, body mass index (BMI), smoking, and alcohol use.

§ For studies of 3HP, a random intercept for study was included while studies of 4R had no random intercepts, all risk ratios and risk differences adjusted for age, sex, BMI, diabetes, HIV status, smoking, and alcohol use.

* All models included random intercept for study and all risk ratios and risk differences were adjusted for age, sex, BMI, diabetes, HIV status, smoking, and alcohol use.

‡ For studies of 3HP, a random intercept for study was included while studies of 4R had no random intercepts, all risk ratios and risk differences adjusted for age, sex, BMI, diabetes, HIV status, smoking, and alcohol use.

**Confidence intervals were estimated with bootstrap resampling methods on 1000 replications and calculated using the 2.5th and 97.5th percentiles of the sampling distribution

Supplemental Table 4.S4. Entire population: Stratified by HIV status, the adjusted risk ratio (aRR) and risk difference (aRD) with their 95% confidence intervals (CI) from both direct and network meta-analysis models for the comparison of treatment completion (80% of expected doses in 120% of allowed time) between 3HP and 4R.

	Intervention				Comparator			Direct IPD MA aRR/aRD (95% CI)	Indirect NMA 3HP compared to 4R aRR/aRD (95% CI)**
Study	**n**	**N**	**%**		**n**	**N**	**%**		
People living with HIV*									
3HP vs 6-9H	**3HP**				**6-9H**				
CDC Study 26	160	182	88		118	164	72	aRR: 1.7 (1.46 , 1.98) aRD: 0.37 (0.32, 0.42)	aRR: 1.69 (1.42, 1.94) aRD: 0.36 (0.25, 0.47)
Martinson	300	328	92		143	327	44		
All	460	510	90		261	491	53		
4R vs 9H	**4R**				**9H**				
Menzies	102	132	77		107	138	78	aRR: 1.00 (0.76 , 1.3) aRD: 0.01 (-0.10, 0.09)	
Total	102	132	77		107	138	78		
People who are HIV negative†									
3HP vs 6-9H	**3HP**				**6-9H**				
CDC Study 26	1642	2072	79.2		1249	2064	60.5	aRR: 1.3 (1.21 , 1.39) aRD: 0.18 (0.16, 0.21)	aRR: 1.04 (0.99, 1.10) aRD: 0.04 (0.01, 0.07)
Sun	118	132	89		104	131	79		
Total	1760	2204	79.9		1353	2195	61.6		
4R vs 9H	**4R**				**9H**				
Menzies	2374	3311	71.7		1858	3278	56.7	aRR: 1.25 (1.18 , 1.32) aRD: 0.14 (0.12, 0.17)	
Diallo	352	422	83		305	407	75		
Total	2726	3733	73.0		2163	3685	58.7		

* For studies of 3HP, a random intercept for study was included while studies of 4R had no random intercepts, all risk ratios and risk differences adjusted for age, sex, and body mass index.
† All models included random intercept for study and risk ratios and risk differences were adjusted for age, sex, body mass index, diabetes, HIV status, smoking, and alcohol use.
**Confidence intervals were estimated with bootstrap resampling methods on 1000 replications and calculated using the 2.5th and 97.5th percentiles of the sampling distribution

Supplemental Table 4.S5. Safety population: Stratified by age groups, the adjusted risk ratio (aRR) and risk difference (aRD) with their 95% confidence intervals (CI) from both direct and network meta-analysis estimates for the comparison of treatment-related adverse events that led to permanent drug discontinuation between 3HP and 4R

	Intervention			Comparator			Direct IPD MA aRR/aRD (95% CI)	Indirect NMA 3HP compared to 4R aRR/aRD (95% CI)**
Study	**n**	**N**	**%**	**n**	**N**	**%**		
Any treatment-related adverse event that led to permanent drug discontinuation								
Those who are < 50 years of age								
3HP vs 6-9H	**3HP**			**6-9H**				
CDC Study 26	183	3526	5.2	120	3349	3.6	aRR: 1.47 (1.17 , 1.84) aRD: 0.02 (0.01 , 0.02)	aRR: 2.91 (2.01, 4.50) aRD: 0.03 (0.02, 0.04)
Martinson	0	324	0	2	324	1		
Sun	8	110	7	5	108	5		
Total	191	3960	4.8	127	3781	3.4		
4R vs 6-9H	**4R**			**6-9H**				
Chan	1	147	1	11	141	8	aRR: 0.5 (0.36 , 0.71) aRD: -0.02 (-0.02 , -0.01)	
Menzies	48	2544	1.9	84	2498	3.4		
Diallo	0	420	0	0	397	0		
Total	49	3111	1.6	95	3036	3.1		
Those who are ≥50 years of age								
3HP vs 6-9H	**3HP**			**6-9H**				
CDC Study 26	64	817	8	50	716	7	aRR: 1.16 (0.81 , 1.67) aRD: 0.01 (-0.02 , 0.04)	aRR: 2.70 (1.56, 5.40) aRD: 0.05 (0.01, 0.08)
Martinson	0	4	0	0	2	0		
Sun	4	22	18	2	23	9		
Total	68	843	8.1	52	741	7.0		
4R vs 6-9H	**4R**			**6-9H**				
Chan	1	43	2	2	42	5	aRR: 0.43 (0.26 , 0.71) aRD: -0.04 (-0.06 , -0.02)	
Menzies	20	737	3	47	733	6		
Total	21	780	3	49	775	6		
Treatment-related grade 3 or 4 adverse events that led to permanent drug discontinuation								
Those who are < 50 years of age								
3HP vs 6-9H	**3HP**			**6-9H**				
CDC Study 26	73	3526	2.1	45	3349	1.3	aRR: 1.45 (1.01, 2.07) aRD: 0.01 (0.00, 0.01)	aRR: 4.14 (2.39, 8.02) aRD: 0.02 (0.01, 0.03)
Martinson	0	324	0	2	324	1		
Sun	1	110	1	3	108	3		
Total	74	3960	1.9	50	3781	1.3		
4R vs 6-9H	**4R**			**6-9H**				
Chan	1	147	1	11	141	8	aRR: 0.35 (0.21 , 0.58) aRD: -0.01 (-0.02, -0.01)	
Menzies	20	2544	0.8	48	2498	1.9		
Diallo	0	420	0	0	397	0		
Total	21	3111	0.7	59	3036	1.9		

Supplemental Table 4.S5. Continued.								
	Intervention			Comparator			Direct IPD MA aRR/aRD (95% CI)	Indirect NMA 3HP compared to 4R aRR/aRD (95% CI)**
Study	n	N	%	n	N	%		
Those who are ≥50 years of age								
3HP vs 6-9H	**3HP**			**6-9H**			aRR: 0.93 (0.57, 1.53) aRD: 0.00 (-0.02, 0.02)	aRR: 2.44 (1.07, 6.21) aRD: 0.02 (-0.01, 0.04)
CDC Study 26	31	817	4	30	716	4		
Martinson	0	4	0	0	2	0		
Sun	1	22	5	1	23	4		
Total	32	843	4	31	741	4		
4R vs 6-9H	**4R**			**6-9H**				
Chan	1	43	2	2	42	5	aRR: 0.38 (0.18 , 0.79) aRD: -0.02 (-0.04 , -0.01)	
Menzies	9	737	1	24	733	3		
Total	10	780	1	26	775	3		
All risk ratios and risk differences are adjusted for age, sex, body mass index, and HIV status. In instances where cell counts are zero, normality assumptions are not met. **Confidence intervals were estimated with bootstrap resampling methods on 1000 replications and calculated using the 2.5th and 97.5th percentiles of the sampling distribution								

Supplemental Table 4.S6. Adjusted risk ratios (aRR) and risk differences (aRD) with their 95% confidence intervals (CI) from direct and network meta-analysis models for the comparison of the incidence of adverse events (various definitions) between 3HP and 4R in the safety population.

	Intervention		Comparator		Direct IPD-MA†	IPD-NMA†
	Events n / N	%	Events n / N	%	aRR & aRD (95% CI)	3HP vs 4R aRR & aRD (95% CI)**
Any adverse event that led to permanent drug discontinuation						
3HP vs 6-9H	**3HP**		**6-9H**			
CDC Study 26	293/4343	6.8	248/4066	6.1	**aRR**: 1.13	
Martinson	4/328	1	6/326	2	(0.96 , 1.34)	
Sun	12/132	9	7/131	5	a**RD**: 0.01	**aRR**: 2.26 (1.76, 3.05)
Total	309/4803	6.4	261/4523	5.8	(0.00 , 0.02)	a**RD**: 0.04 (0.02, 0.05)
4R vs 6-9H	**4R**		**6-9H**			
Chan	11/190	6	26/183	14	**aRR**: 0.50	
Menzies	93/3281	2.8	179/3231	5.5	(0.40 , 0.63)	
Diallo	1/420	0	1/397	0	**aRD**: -0.03	
Total	105/3891	2.7	206/3811	5.4	(-0.04 , -0.02)	
Any Grade 3 or 4 adverse events with any relation to treatment						
3HP vs 6-9H	**3HP**		**6-9H**			
CDC Study 26	239/4343	5.5	244/4066	6.0	**aRR**: 0.90	
Martinson	16/328	5	21/326	6	(0.76 , 1.07)	
Sun	3/132	2	5/131	4	**aRD**: -0.01	**aRR**: 2.19 (1.56, 3.16)
Total	258/4803	5.4	270/4523	6.0	(-0.02 , 0.00)	**aRD**: 0.014 (0.003, 0.025)
4R vs 6-9H	**4R**		**6-9H**			
Chan	2/190	1	13/183	7	**aRR**: 0.41	
Menzies	51/3281	1.6	112/3231	3.5	(0.30 , 0.57)	
Diallo	0/420	0	1/397	0	**aRD**: -0.02	
Total	53/3891	1.4	126/3811	3.3	(-0.03 , -0.01)	
Grade 3 or 4 adverse events that were judged related to treatment*						
3HP vs 6-9H	**3HP**		**6-9H**			
CDC Study 26	126/4343	2.9	99/4066	2.4	**aRR**: 1.1	
Martinson	2/328	1	8/326	3	(0.86 , 1.42)	
Sun	2/132	2	5/131	4	a**RD**: 0.00	**aRR**: 2.95 (1.84, 4.97)
Total	130/4803	2.7	112/4523	2.5	(0.00 , 0.01)	**aRD**: 0.02 (0.01, 0.03)
4R vs 6-9H	**4R**		**6-9H**			
Chan	2/190	1	13/183	7	**aRR**: 0.37	
Menzies	31/3281	0.9	74/3231	2.3	(0.25 , 0.56)	
Diallo	0/420	0	0/397	0	**aRD**: -0.01	
Total	33/3891	0.9	87/3811	2.3	(-0.02 , -0.01)	

† Risk ratios and risk differences adjusted for age, sex, body mass index, and HIV infection. *Judged to be possibly, probably, or definitely related to study drug in primary studies **Confidence intervals were estimated with bootstrap resampling methods on 1000 replications and calculated using the 2.5th and 97.5th percentiles of the sampling distribution

Supplemental Table 4.S7. In the entire and per-protocol populations: the adjusted risk ratio (aRR) and risk difference (aRD) with their 95% confidence intervals (CI) from both direct and network meta-analysis estimates for the comparison of treatment-related grade 3 or 4 adverse events that led to permanent drug discontinuation between 3HP and 4R.

	Intervention			Comparator			Direct IPD MA aRR/aRD (95% CI)	Indirect NMA 3HP compared to 4R aRR/aRD (95% CI)**
Study	**n**	**N**	**%**	**n**	**N**	**%**		
Entire population								
3HP vs 6-9H	**3HP**			**6-9H**				
CDC Study 26	104	4437	2.3	75	4156	1.8	aRR: 1.24 (0.93 , 1.66) aRD: 0.004 (-0.001 , 0.01)*	aRR: 3.44 (2.16 , 5.99) aRD: 0.018 (0.009, 0.025)*
Martinson	0	328	0	2	327	1		
Sun	2	132	2	4	131	3		
Total	106	4897	2.2	81	4614	1.8		
4R vs 6-9H	**4R**			**6-9H**				
Chan	2	190	1	13	183	7	aRR: 0.36 (0.25 , 0.56) aRD: -0.014 (-0.02 , -0.01)	
Menzies	29	3443	0.8	72	3416	2.1		
Diallo	0	422	0	0	407	0		
Total	31	4055	0.8	85	4006	2.1		
Per-protocol population (those who took at least 80% of expected doses)								
3HP vs 6-9H	**3HP**			**6-9H**				
CDC Study 26	104	3772	2.8	75	3122	2.4	aRR: 1.1 (0.83 , 1.47) aRD: 0.003* (-0.00 , 0.01)	aRR: 3.62 (2.27, 6.21) aRD: 0.024 (0.015, 0.035)*
Martinson	0	314	0	2	167	1		
Sun	2	120	2	4	108	4		
Total	106	4206	2.5	81	3397	2.4		
4R vs 6-9H	**4R**			**6-9H**				
Chan	2	169	1	13	155	8	aRR: 0.31 (0.20, 0.46) aRD: -0.02 (-0.03 , -0.02)	
Menzies	29	2757	1.1	72	2235	3.2		
Diallo	0	367	0	0	316	0		
Total	31	3293	0.9	85	2706	3.1		

Risk ratios and risk differences adjusted for age, sex, body mass index, and HIV infection. All models include random intercept for study. * Additional decimal place added for clarity. **Confidence intervals were estimated with bootstrap resampling methods on 1000 replications and calculated using the 2.5th and 97.5th percentiles of the sampling distribution

Supplemental Table 4.S8. Number and rate (%) of adverse events for all definitions, by HIV status in the safety population.

	3HP			4R			6-9H in 3HP studies			6-9H in 4R studies		
HIV status	**n**	**N**	**%**	**n**	**N**	**%**	**n**	**N**	**%**	**n**	**N**	**%**
Treatment-related grade 3 or 4 adverse events that led to permanent drug discontinuation												
People living with HIV	3	509	1	1	129	1	8	485	2	5	138	4
People who are HIV negative	65	2169	3	30	3762	0.8	47	2154	2.2	80	3673	2.2
HIV status unknown	38	2125	1.8				26	1884	1.4			
Total	106	4803	2.2	31	3891	0.8	81	4523	1.8	85	3811	2.2
Any treatment-related adverse event that led to permanent drug discontinuation*												
People living with HIV	5	509	1	2	129	2	9	485	2	5	138	4
People who are HIV negative	143	2169	6.6	68	3762	1.8	92	2154	4.3	139	3673	3.8
HIV status unknown	111	2125	5.2				78	1884	4.1			
Total	259	4803	5.4	70	3891	1.8	179	4523	4	144	3811	3.8
Any adverse event that led to permanent drug discontinuation												
People living with HIV	9	509	2	3	129	2	14	485	3	8	138	6
People who are HIV negative	175	2169	8.1	102	3762	2.7	136	2154	6.3	198	3673	5.4
HIV status unknown	125	2125	5.9				111	1884	5.9			
Total	309	4803	6.4	105	3891	2.7	261	4523	5.8	206	3811	5.4
Any Grade 3 or 4 adverse events with any relation to treatment												
People living with HIV	31	509	6	2	129	2	45	485	9	8	138	6
People who are HIV negative	147	2169	6.8	51	3762	1.4	146	2154	6.8	118	3673	3.2
HIV status unknown	80	2125	3.8				79	1884	4.2			
Total	258	4803	5.4	53	3891	1.4	270	4523	6	126	3811	3.3
Grade 3 or 4 adverse events that were judged related to treatment												
People living with HIV	8	509	2	1	129	1	16	485	3	5	138	4
People who are HIV negative	76	2169	3.5	32	3762	0.9	59	2154	2.7	82	3673	2.2
HIV status unknown	46	2125	2.2				37	1884	2			
Total	130	4803	2.7	33	3891	0.8	112	4523	2.5	87	3811	2.3

REFERENCES

1. World Health Organization. Global tuberculosis report 2022. Geneva, Switzerland: World Health Organization, 2022.
2. World Health Organization. WHO operational handbook on tuberculosis. Module 1: prevention - tuberculosis preventive treatment. Geneva, Switzerland: World Health Organization, 2020.
3. Reid A, Grant AD, White RG, et al. Accelerating progress towards tuberculosis elimination: the need for combination treatment and prevention. *The International Journal of Tuberculosis and Lung Disease* 2015; **19**(1): 5-9.
4. Uplekar M, Weil D, Lonnroth K, et al. WHO's new end TB strategy. *Lancet* 2015; **385**(9979): 1799-801.
5. Tiberi S, du Plessis N, Walzl G, et al. Tuberculosis: progress and advances in development of new drugs, treatment regimens, and host-directed therapies. *The Lancet Infectious diseases* 2018.
6. Njie GJ, Morris SB, Woodruff RY, Moro RN, Vernon AA, Borisov AS. Isoniazid-Rifapentine for Latent Tuberculosis Infection: A Systematic Review and Meta-analysis. *American journal of preventive medicine* 2018.
7. Hirsch-Moverman Y, Daftary A, Franks J, Colson PW. Adherence to treatment for latent tuberculosis infection: systematic review of studies in the US and Canada. *The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease* 2008; **12**(11): 1235-54.
8. Menzies D, Long R, Trajman A, et al. Adverse events with 4 months of rifampin therapy or 9 months of isoniazid therapy for latent tuberculosis infection: a randomized trial. *Annals of internal medicine* 2008; **149**(10): 689-97.
9. World Health Organization. Latent tuberculosis infection: updated and consolidated guidelines for programmatic management. Geneva: World Health Organization, 2018.
10. Borisov AS, Bamrah Morris S, Njie GJ, et al. Update of Recommendations for Use of Once-Weekly Isoniazid-Rifapentine Regimen to Treat Latent Mycobacterium tuberculosis Infection. *MMWR Morb Mortal Wkly Rep* 2018; **67**(25): 723-6.
11. Diallo T, Adjobimey M, Ruslami R, et al. Safety and Side Effects of Rifampin versus Isoniazid in Children. *The New England journal of medicine* 2018; **379**(5): 454-63.
12. Martinson NA, Barnes GL, Moulton LH, et al. New regimens to prevent tuberculosis in adults with HIV infection. *The New England journal of medicine* 2011; **365**(1): 11-20.
13. Menzies D, Adjobimey M, Ruslami R, et al. Four Months of Rifampin or Nine Months of Isoniazid for Latent Tuberculosis in Adults. *The New England journal of medicine* 2018; **379**(5): 440-53.
14. Sterling TR, Villarino ME, Borisov AS, et al. Three months of rifapentine and isoniazid for latent tuberculosis infection. *The New England journal of medicine* 2011; **365**(23): 2155-66.
15. Villarino ME, Scott NA, Weis SE, et al. Treatment for preventing tuberculosis in children and adolescents: a randomized clinical trial of a 3-month, 12-dose regimen of a combination of rifapentine and isoniazid. *JAMA Pediatr* 2015; **169**(3): 247-55.

16. Sterling TR, Moro RN, Borisov AS, et al. Flu-like and Other Systemic Drug Reactions Among Persons Receiving Weekly Rifapentine Plus Isoniazid or Daily Isoniazid for Treatment of Latent Tuberculosis Infection in the PREVENT Tuberculosis Study. *Clin Infect Dis* 2015; **61**(4): 527-35.
17. Zenner D, Beer N, Harris RJ, Lipman MC, Stagg HR, van der Werf MJ. Treatment of Latent Tuberculosis Infection: An Updated Network Meta-analysis. *Annals of internal medicine* 2017; **167**(4): 248-55.
18. National Cancer Institute. Common Terminology Criteria for Adverse Events (CTCAE), v5.0 2017 https://www.eortc.be/services/doc/ctc/CTCAE_4.03_2010-06-14_QuickReference_5x7.pdf.
19. Sterne JAC, Savovic J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *The British Medical Journal* 2019; **366**: l4898.
20. Thorlund K, Mills EJ. Sample size and power considerations in network meta-analysis. *Syst Rev* 2012; **1**: 41.
21. Bagos P, Nikolopoulos G. Mixed-Effects Poisson Regression Models for Meta-Analysis of Follow-Up Studies with Constant or Varying Durations. *The International Journal of Biostatistics* 2009; **5**(1).
22. Veroniki AA, Straus SE, Soobiah C, Elliott MJ, Tricco AC. A scoping review of indirect comparison methods and applications using individual patient data. *BMC Med Res Methodol* 2016; **16**: 47.
23. Morton SC, Murad MH, O'Connor E, et al. Quantitative Synthesis-An Update. Methods Guide for Effectiveness and Comparative Effectiveness Reviews. Rockville (MD); 2008.
24. R Core Team (2020). R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria, URL https://wwwR-projectorg/*.
25. Whalen CC, Johnson JL, Okwera A, et al. A trial of three regimens to prevent tuberculosis in Ugandan adults infected with the human immunodeficiency virus. Uganda-Case Western Reserve University Research Collaboration. *The New England journal of medicine* 1997; **337**(12): 801-8.
26. Sun HY, Huang YW, Huang WC, et al. Twelve-dose weekly rifapentine plus isoniazid for latent tuberculosis infection: A multicentre randomised controlled trial in Taiwan. *Tuberculosis (Edinb)* 2018; **111**: 121-6.
27. Sterling TR, Scott NA, Miro JM, et al. Three months of weekly rifapentine and isoniazid for treatment of Mycobacterium tuberculosis infection in HIV-coinfected persons. *AIDS* 2016; **30**(10): 1607-15.
28. Spyridis NP, Spyridis PG, Gelesme A, et al. The effectiveness of a 9-month regimen of isoniazid alone versus 3- and 4-month regimens of isoniazid plus rifampin for treatment of latent tuberculosis infection in children: results of an 11-year randomized study. *Clin Infect Dis* 2007; **45**(6): 715-22.
29. Rivero A, Lopez-Cortes L, Castillo R, et al. [Randomized clinical trial investigating three chemoprophylaxis regimens for latent tuberculosis infection in HIV-infected patients]. *Enferm Infecc Microbiol Clin* 2007; **25**(5): 305-10.
30. Moro RN, Borisov AS, Saukkonen J, et al. Factors Associated With Noncompletion of Latent Tuberculosis Infection Treatment: Experience From the PREVENT TB Trial in the United States and Canada. *Clin Infect Dis* 2016; **62**(11): 1390-400. doi: 10.093/cid/ciw126. Epub 2016 Mar 6.

31. Johnson JL, Okwera A, Hom DL, et al. Duration of efficacy of treatment of latent tuberculosis infection in HIV-infected adults. *AIDS* 2001; **15**(16): 2137-47.
32. Chan PC, Yang CH, Chang LY, et al. Latent tuberculosis infection treatment for prison inmates: a randomised controlled trial. *The International Journal of Tuberculosis and Lung Disease* 2012; **16**(5): 633-8.
33. Belknap R, Holland D, Feng PJ, et al. Self-administered Versus Directly Observed Once-Weekly Isoniazid and Rifapentine Treatment of Latent Tuberculosis Infection: A Randomized Trial. *Annals of internal medicine* 2017; **167**(10): 689-97.
34. HKCS. A double-blind placebo-controlled clinical trial of three antituberculosis chemoprophylaxis regimens in patients with silicosis in Hong Kong. Hong Kong Chest Service/Tuberculosis Research Centre, Madras/British Medical Research Council. *Am Rev Respir Dis* 1992; **145**(1): 36-41.
35. White MC, Tulsky JP, Lee JR, et al. Isoniazid vs. rifampin for latent tuberculosis infection in jail inmates: toxicity and adherence. *Journal of correctional health care*, 2012. (accessed.

Chapter 5 – Efficacy of adding clofazimine to WHO group A drugs for treatment of rifampicin- and multidrug-resistant tuberculosis: an emulated target trial.

Section 5.1 Preface

In this second manuscript I shift focus towards treatment of multidrug-resistant tuberculosis (MDR-TB). Although the WHO recommends many drugs for treatment of MDR-TB with advanced disease and/or extensive resistance, there is a lack of evidence regarding the effect that each additional group B drug (clofazimine and cycloserine/terizidone) has on treatment success when added to the group A drugs (bedaquiline, linezolid, and a fluoroquinolone), and specifically the effect that adding clofazimine has on treatment success. This study aims to address that gap in the literature. Additionally, treatment for MDR-TB is long and uses drugs with poor efficacy and high risk of adverse events, thus there is concern for bias due to time-varying confounding arising during the course of treatment.

This study was used to compare treatment success between MDR-TB patients who initially received, in addition to other drugs being concurrently prescribed for MDR-TB treatment, all three WHO group A drugs but not clofazimine to patients who received the same three group A drugs with the addition of clofazimine. I attempted to identify the causal effect on treatment success between these regimens by emulating a target trial using observational data from the EndTB observational study and applying methods to account for time-varying confounding.

Section 5.2 Manuscript 2

Efficacy of adding clofazimine to WHO group A drugs in treatment of rifampicin and multidrug-resistant tuberculosis: an emulated target trial.

Nicholas Winters[1], Mireille Schnitzer[1,2,3], Yan Liu[2], and Dick Menzies[1]

Affiliations:

1. Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Canada
2. Faculty of Pharmacy, Université de Montréal, Canada.
3. Department of Social and Preventive Medicine, Université de Montréal, Canada.

ABSTRACT

Objective: For treatment of rifampicin- and multidrug-resistant tuberculosis (MDR-TB) the World Health Organization (WHO) recommend all three group A drugs (bedaquiline, linezolid, and a fluoroquinolone) supplemented with group B drugs (clofazimine and cycloserine or terizidone). Both the effect of adding clofazimine to the group A drugs and the impact that time-varying confounding may have on treatment success is uncertain.

Methods: We conducted a target trial emulation using observational data to compare treatment success and safety between two core regimens received in addition to other drugs being concurrently prescribed for MDR-TB treatment: all three WHO group A drugs without clofazimine (control group) and all three WHO group A drugs with clofazimine (clofazimine group). We conducted an intention-to-treat (ITT) analysis using inverse probability of treatment weights (IPTW), and four per-protocol analyses accounting for treatment strategy censoring with time-varying inverse probability of censoring weights (IPCW). Average treatment effects (ATE) and their 95% confidence intervals (CIs) were estimated for each analysis.

Results: We included 791 patients: 197 in the control group and 594 in the clofazimine group. In the ITT analysis the IPTW-ATE comparing the clofazimine group to the control group for treatment success was 0.07 (95% CI: -0.01, 0.15). Adverse events leading to permanent discontinuation of clofazimine were rare. In the first per-protocol analysis, the IPCW-ATE for treatment success was substantially attenuated from the ITT (-0.01 [95% CI: -0.10, 0.07]). In the second per-protocol analysis IPCW estimates were similar to the IPTW estimates in the ITT analysis. In the third per-protocol analysis the IPCW-ATE was also attenuated from the ITT (0.02 [95% CI: -0.05, 0.09]). The IPCW estimates in the fourth per-protocol analysis were similar to the IPTW estimates in the ITT analysis.

Conclusion: In this target trial emulation, we found some evidence that adding clofazimine the three WHO group A drugs increased treatment success in an ITT analysis and in the second and fourth per-protocol analysis, without evidence of added toxicity. However, we did not find evidence that adding clofazimine to the WHO group A drugs increases treatment success in the remaining per-protocol analyses.

INTRODUCTION

Rifampicin-resistant (RR) and multidrug-resistant tuberculosis (MDR-TB), defined as tuberculosis with resistance to both rifampicin and isoniazid, are serious global health issues with an estimated 450,000 cases in 2021.[1,2] Although shorter treatments are recommended by the World Health Organization (WHO) for some forms of drug-resistant and MDR-TB,[3-8] patients with more extensive resistance and advanced disease require longer treatment with at least four likely effective drugs.[9] Regimens for more advanced forms of MDR-TB with additional drug resistance should include all three WHO group A drugs (bedaquiline, linezolid, and a fluoroquinolone) and at least one group B drug (clofazimine and cycloserine or terizidone). However, the comparative efficacy between regimens that include the addition of clofazimine to all three group A drugs and those with only group A drugs without clofazimine, has not been established.

Studying treatment efficacy of MDR-TB treatments is complicated due to the long treatment duration and use of drugs with poor safety profiles,[10] which make adherence to treatment difficult. Conventionally, analyses of MDR-TB treatment have assessed efficacy of individual drugs given at treatment initiation[11-14] as ever or never received during treatment. However, factors throughout treatment (such as adverse events, bacterial culture results, and acquired resistance) may affect treatment changes and can also be predictive of outcomes.[15,16] These time-varying confounders could potentially result in biased estimates of treatment effects but can be accounted for using appropriate methods such as marginal structural models, which can be fit using inverse probability weights to adjust for measured confounders of treatment and censoring.[17-19] An additional method that aims to further reduce bias in observational studies is use of target trial emulation.[20,21] Emulation of a target

trial involves specifying a hypothetical randomized trial by outlining a protocol that includes eligibility, treatment strategies, follow-up, outcomes, causal contrasts, and an analysis plan, which is then emulated using observational data.

In this study, we emulate a target trial[20,21] to compare efficacy and safety between two core MDR-TB treatment regimens received at treatment initiation, in addition to other drugs being concurrently received for MDR-TB treatment which were given at the discretion of the provider, that consisted of: bedaquiline, linezolid, and a fluoroquinolone without clofazimine; or bedaquiline, linezolid, and a fluoroquinolone plus clofazimine. In an intention-to-treat (ITT) analysis we compared treatment effects between standard conditional adjustment for confounders and baseline inverse probability of treatment weighted (IPTW) estimates. The occurrence of adverse events between the two treatment groups were also described in the ITT population to assess safety. Per-protocol analyses were conducted to assess the sustained treatment effects between the two core regimens under different censoring criteria, which accounted for time-varying confounding using inverse probability of censoring weights (IPCW). We assessed four per-protocol effects which were interpreted as follows: 1) the effect of including clofazimine at treatment initiation compared to those who never received clofazimine at any time during treatment, but allowing for stoppages of any drug in the core regimen except for 'other' reasons for stopping; 2) the effect of including clofazimine from treatment initiation compared to those who did not receive clofazimine at treatment initiation (allowing for clofazimine starting post-baseline), and allowing for stoppages of any drug in the core regimen except for 'other' reasons for stopping; 3) the effect of including clofazimine at treatment initiation compared to those who never received clofazimine at any time during treatment in those remaining on

their assigned therapy until their outcome; and 4) the effect of including clofazimine from treatment initiation compared to those who did not receive clofazimine at treatment initiation (allowing for clofazimine starting post-baseline) in those who remained on assigned therapy until their treatment outcome.

METHODS

To emulate our target trial we used data from the EndTB observational study conducted between 2015 and 2018 across 17 countries that included 2789 patients with RR/MDR-TB, described in detail previously.[22,23] Patients in the EndTB study were treated according to WHO and national guidelines and received 18 to 24-month regimens including bedaquiline and/or delamanid. Follow-up of patients was from treatment initiation until the end of their treatment; end of treatment outcomes were defined according to WHO 2013.[24] Detailed information was recorded on start and stop dates of prescribed drugs as well as reasons for drug stoppages. Adverse events were graded by reporting physician according to Medicine Sans Frontières (MSF) grading system,[25] and the MSF Pharmacovigilance unit received reports of any serious adverse events within 24 hours of physician knowledge.

We emulated a target trial using the framework outlined by Hernan et al.,[20] which is summarised in Supplemental Table 5.S1 and described as follows:

Target trial design

Eligibility criteria: Patients of any age initiating MDR-TB treatment who are eligible at baseline to receive all drugs in each treatment regimen, disregarding drug susceptibility testing results.

Treatment strategies: Initiation at baseline, in addition to other drugs being concurrently prescribed for MDR-TB treatment at the discretion of the provider, of the control intervention of bedaquiline, linezolid, and a fluoroquinolone (FQ), i.e. the three group A drugs as defined by WHO,[9] without clofazimine (control group) or the same three group A drugs (bedaquiline, linezolid, and a FQ) plus clofazimine (clofazimine group). Patients will also be receiving other supplementary MDR-TB drugs in addition to the drugs in their treatment strategy, but these were not considered in the selection of patients or adjusted for in the analyses as these would be prescribed and managed at the discretion of the treating physician. We refer to the drugs being compared in the control group and the clofazimine group (the three group A drugs bedaquiline, linezolid, and an FQ) as the *core regimens* to distinguish them from the entire treatment regimen that includes all concurrent drugs used in treatment. For the first per-protocol analysis, the treatment strategy involved remaining on the assigned treatment but allowing for permanent drug stoppages for these reasons: adverse events, planned treatment changes, reintroduction or replacement of stopped drug, resistance to drug, drug supply or drug administration issues, or pregnancy but censoring patients for any 'other' reason a drug was permanently stopped as well as censoring patients in the control group who start clofazimine. This is interpreted as the effect of including clofazimine at treatment initiation compared to those who never received clofazimine at any time during treatment and allowing for stoppages of any drug in the core regimen except for 'other' reasons. Although no additional detail on what constitutes the 'other' reasons were available in the data, this reason is unique from the rest of the drug stoppages as it includes reasons that are not due to expected issues arising during treatment and may include such things as a patient decision to stop a drug or defaults. In the second per-protocol analysis,

the treatment strategy was the same as in the first per-protocol analysis but allowed patients in the control group to start clofazimine without being censored and is interpreted as the effect of including clofazimine from treatment initiation compared to those who did not receive clofazimine at treatment initiation, allowing for clofazimine starting post-baseline and for stoppages of any drug in the core regimen except for 'other' reasons. In the third per-protocol analysis, the treatment strategy involved remaining on the assigned treatment but censoring for any of the reasons for permanent drug stoppages mentioned above as well as censoring patients in the control group who start clofazimine. This is interpreted as the effect of including clofazimine at treatment initiation compared to those who never received clofazimine at any time during treatment in all patients who remained on their assigned core regimen until their outcome. In the fourth per-protocol analysis, the treatment strategy was the same as in the third per-protocol analysis but allowed patients in the control group to start clofazimine without being censored and is interpreted as the effect of including clofazimine from treatment initiation compared to those who did not receive clofazimine at treatment initiation (allowing for clofazimine starting post-baseline), in all patients who remained on their assigned core regimen until their treatment outcome.

Treatment assignment: MDR-TB patients would be randomized to either the control group or the clofazimine group.

Outcomes: Treatment success defined as cure or completion of treatment, compared to all other negative outcomes (death, treatment failure [defined as microbiologic failure or a change in any two drugs received due to adverse events including the supplemental,

concurrent MDR-TB drugs given in addition to the core regimens], or lost to follow-up) as defined in WHO 2013.[24]

Causal contrasts of interest: We will estimate the intention-to-treat (ITT) effect and four per-protocol (sustained treatment) effects.

Analysis plan: The ITT analysis involves direct comparison of the proportion of patients with treatment success among those assigned to each treatment. The per-protocol analyses will censor any patient deviating from their respective treatment strategies described above. Any patient randomized to the control group will be censored if they received clofazimine for >30 days any time after baseline treatment initiation in the first and third per-protocol analyses. All per-protocol analyses will adjust for both baseline and post-baseline confounders of drug stoppages and treatment censoring (and for initiating clofazimine for those in the control group in the first and third per-protocol analyses) using time-varying inverse probability of censoring weights (IPCW).

Emulation using observational data:

Eligibility, treatment strategies, outcome, and causal contrast are the same as target trial.

Treatment assignment: Patients who at baseline initiated MDR-TB therapy with the core regimen of the control group (all three WHO group A drugs without clofazimine) or with the core regimen of the clofazimine group (all three WHO group A drugs plus clofazimine), in addition to other drugs being concurrently received for MDR-TB treatment at baseline.

Analysis plan:

Descriptive statistics: Baseline patient characteristics were described using n (%) for categorical variables and mean (standard deviation [SD]) or medians (interquartile range [IQR]) in the available, non-imputed data. Adverse events resulting in a drug stoppage for each drug in the core regimens, as well as any drug concurrently received in addition to the core regimens, were summarized in a descriptive analysis as n (%) in the overall (ITT) population.

Missing data: For statistical analyses, all missing observations for confounders were imputed using multivariate imputation by chained equations (MICE), and twenty data sets were generated with 25 Gibb's sampling iterations.[26] For drug susceptibility testing (DST), we assumed bedaquiline, linezolid, clofazimine were susceptible unless resistance was proven, while levofloxacin and moxifloxacin DST were combined into one composite variable; if moxifloxacin DST was missing it was replaced with levofloxacin DST, and if DST for both FQs were missing they were considered likely effective if there was no history of past treatment use with either of these drugs. If the DST results were missing for both FQs, and information regarding prior use of both of these FQs were also missing, then the missing DST observations for this FQ variable were imputed.

Statistical analysis:

To estimate ITT effects, we used three linear regression models: (i) unadjusted, (ii) conditionally adjusted for baseline covariates, and (iii) weighted by IPW using a propensity score predicting initial treatment assignment. The conditional linear regression was adjusted for the following baseline covariate set (W): age (continuous), sex, diabetes, HIV infection

(positive or negative), renal insufficiency, current alcohol use, current smoker, intravenous drug use, bilateral disease, cavitation on chest radiography, sputum acid-fast bacillus (AFB) smear status (positive or negative), culture result (positive or negative), body mass index category (underweight: <18.5 kg/m²; normal: ≥18.5 and <25 kg/m²; and overweight or obese: ≥25 kg/m^2), Eastern Cooperative Oncology Group (ECOG) functional status[27] at baseline (fully active, restricted in physically strenuous activity but ambulatory, ambulatory with limited self-care, and completely disabled), past tuberculosis drug use, MDR category (only MDR-TB, MDR-TB plus resistance to second-line injectable (SLI) but FQ sensitive, MDR-TB plus resistance to FQ but SLI sensitive, and MDR-TB plus resistance to both FQ and SLI), and drug susceptibility testing for linezolid, clofazimine, and the composite variable for likely effective FQ (described previously). All dichotomous variables were included as yes or no while continuous variables were included as linear functions. For the baseline IPTW, we first estimated the probability of a subject receiving treatment using a logistic regression model predicting treatment assignment (A), adjusted for the baseline covariate set (W) described above, defined as P(A=1 | W). Patients were then assigned weights based on the intervention received: those in the clofazimine group were assigned a weight equal to 1/P(A=1 | W); while those in the control group were weighted as 1/[1-P(A=1 | W)], weights were not stabilized or truncated. Our target parameters were the average treatment effects (ATE), i.e. the difference in probability of successful outcome, with their 95% confidence interval (CI) calculated using a robust sandwich variance estimator.

For each per-protocol analysis we first estimated the crude effects, comparing outcomes in those who were uncensored between the two treatment groups, and then the baseline IPTW

estimates in the same population, calculated using the same baseline covariate set (W) outlined for the ITT analysis above.

Per-protocol effects were estimated while adjusting for time-dependent predictors of censoring using longitudinal IPCW in addition to the baseline treatment weights described above. For drug stoppages violating the treatment strategy, subjects were considered censored if the drug was stopped at least 14 days before their last observed study day (to distinguish regimen censoring from those who were likely failing treatment), while stopping bedaquiline was considered censored if occurring at least 14 days before 6 months of bedaquiline use, any stoppage of bedaquiline after was not considered censored. Time periods were discretized into each month a patient was on treatment (to reflect the minimum frequency of follow-up visits in the EndTB observational study), and updated monthly with their treatment status, drug stoppage and starts, and the following time-varying covariate set (L_t): body mass index category, culture result, ECOG functional status, highest count of adverse events, and highest count of serious adverse events. A pooled logistic regression model for censoring at any month was fit conditional on W and time-updated variables L_t, dependent on being uncensored in the previous month, and stratified by initial treatment group. There was a maximum of 36 months of treatment, however no new censoring occurred after 22 months, thus the censoring models were pooled over month 2 to month 22. Each uncensored person-month was assigned the weight $1/P(Censor_t=1 \mid W, L_t, Censor_{t-1} = 0)$. The IPCW were not stabilized or truncated. Finally, a linear regression model with treatment success as outcome and treatment group as a covariate, fit using only uncensored participants, was weighted by the cumulative product of the IPCWs multiplied by the baseline IPTW (as calculated in the ITT analysis) to estimate the per-protocol ATE

(or difference in probability of successful outcome). We also conducted a post-hoc, sensitivity analysis to assess effect measure modification for FQ susceptibility on DST by stratifying analyses by susceptible or resistant using the composite variable for FQ DST that includes likely effective if no past treatment with an FQ described previously, in the ITT population. All analyses were conducted in R Version 4.1.2.[28]

RESULTS

Of the 2789 subjects enrolled in the EndTB cohort, we included 791 patients (197 were in the control group and 594 were in the clofazimine group) from 16 countries. The baseline characteristics prior to imputation, stratified by initial treatment group, are summarized in Table 5.1. The average age of all subjects was 37 (SD 13) years, 39.6% were female, and the mean body mass index was 20 (SD 4) kg/m². Patients in both regimens were similar, with exception of past first-line drug use being more frequent in the control group, while past second-line drug use and MDR-TB with resistance to both an FQ and an SLI were more frequent in the clofazimine group. Drug susceptibility testing was rarely performed for bedaquiline, linezolid, and clofazimine, but testing for levofloxacin and moxifloxacin were performed for the majority of patients for which the presence of resistance was similar. The number of drugs used and number of effective drugs used were similar in both treatment groups. For concurrent drugs received in addition to the core regimens, more patients in the control group received pyrazinamide, amikacin, ethionamide or prothionamide, and cycloserine/terizidone, while more patients in the clofazimine group received cilastatin/imipenem/meropenem and amoxicillin-clavulanic acid (Table 5.2).

The characteristics of censored and uncensored patients are summarized in Supplemental Table 5.S2A and 5.S2B for each per-protocol analysis. In the first per-protocol analysis, 115 patients were censored: 90 were in the control group and 25 were in the clofazimine group. Among the censored patients, alcohol and smoking were more common, there were fewer patients with fully active ECOG functional status at baseline, but with less resistance to FQ compared to uncensored patients. In the second per-protocol analysis, 34 patients were censored: 9 in the control group and 25 in the clofazimine group and the difference between censored and uncensored patients were similar to the patients in the first per-protocol analysis. In the third per-protocol analysis, there were 276 patients censored: 103 in the control group and 173 in the clofazimine group. The uncensored patients were similar to the overall population included in the ITT analysis; however the censored patients had a higher prevalence of HIV and a lower prevalence of fully active ECOG functional status than those who were not censored. In the fourth per-protocol analysis, 233 patients were censored: 60 in the control group and 173 in the clofazimine group, and differences between the two groups were similar to patients in the third per-protocol analysis.

Intention-to-treat analysis

In the ITT analysis (Table 5.3), the number of patients with successful treatment outcomes was 144/197 (73.1%) for those in the control group (all three group A drugs without clofazimine) and 466/594 (78.5%) for those in the clofazimine group (all three group A drugs plus clofazimine). The unadjusted estimate was 0.05 (95% CI: -0.02, 0.12) while the fully adjusted estimate was 0.03 (95% CI: -0.04, 0.11). For the model calculated using baseline IPTW, the estimate was 0.07 (95% CI: -0.01, 0.15). In the sensitivity analysis assessing effect measure modification by the composite variable for DST results for FQs

(which includes likely effectiveness), the IPTW estimate in those who were susceptible to FQs at baseline was 0.13 (95% CI: 0.01, 0.24) and -0.03 (95% CI: -0.15, 0.10) for those who were resistant to FQs at baseline (Supplemental Table 5.S3).

The rate of adverse events leading to drug stoppage for drugs prescribed in addition to those in the core regimens were similar between treatment groups in the ITT (Table 5.4). Few patients (2.9%) in the clofazimine group stopped clofazimine for an adverse event. Of the 85 in the control group who started clofazimine later only 2 patients (2.3%) stopped clofazimine due to an adverse event.

Per-protocol analyses

In all per-protocol analyses, patients were censored uniformly over time (Figure 5.1), with no substantial difference between treatment groups. In the first per-protocol analysis, 43.1% patients in the control group were censored for starting clofazimine while 2.5% were censored for other reasons (Table 5.5). All patients in the clofazimine group for the first per-protocol analysis were censored for other reasons. In the second per-protocol analysis, patients were only censored for other reasons for drug stoppages. In the third per-protocol analysis the most common reason for censoring in the control group was because of starting clofazimine (36.5% of the total control group at baseline) and for planned changes to treatment (7.1%). On the other hand, in the clofazimine group the most common reason for censoring was because of an adverse event (11.3%) and for planned changes to treatment (8.4%). In the fourth per-protocol analysis, the most common reasons for censoring was due to adverse events followed by planned treatment changes. The patients in the control group who started clofazimine at any time during treatment were similar to those in the control group who never started clofazimine at any time during treatment, except both

moxifloxacin susceptibility was more prevalent and more effective drugs were used in those who never started clofazimine (Supplemental Table 5.S4). There were more drug stoppages due to adverse events in the patients in the control group who started clofazimine compared to controls who never started clofazimine at any time (Supplemental Table 5.S5). Timing of when patients in the control group started clofazimine is shown in Supplemental Figure 5.S1.

In the first per-protocol analysis, 115 patients were censored. The number of uncensored patients with successful treatment outcomes was 77/107 (72%) for those in the control group and 453/569 (79.6%) in the clofazimine group (Table 5.3). The unadjusted estimate and the baseline IPTW estimate contrasting the outcomes of uncensored patients were 0.08 (95% CI: -0.02, 0.17) and 0.09 (95% CI: -0.02, 0.19) respectively. The IPCW estimate was -0.01 (95% CI: -0.10, 0.07), with censoring weights for the control group and clofazimine group having a median value of 5.3 (IQR 3.7 to 8.5) and 1.3 (IQR 1.2 to 1.4) respectively. For patients in the second per-protocol analysis, where patients starting clofazimine were not censored, 34 patients were censored and the IPCW estimates were similar to the IPTW estimates in the ITT analysis.

In the third per-protocol analysis, 276 patients were censored. The number of patients with successful treatment outcomes was 67/94 (71.3%) for those in the control group and 336/421 (79.8%) in the clofazimine group (Table 5.3). The unadjusted estimate and the baseline IPTW estimate contrasting the outcomes of uncensored patients were 0.09 (95% CI: -0.01, 0.19) and 0.10 (95% CI: -0.02, to 0.22) respectively. The IPCW estimate was 0.02 (95% CI: -0.05, 0.09), with censoring weights for the control group and clofazimine group having a median value of 5.4 (IQR 4.0 to 8.1) and 1.8 (IQR 1.5 to 2.1) respectively. In the

fourth per-protocol analysis, where patients starting clofazimine were not censored, 233 patients were censored and the IPCW estimates were similar to the IPTW estimates in the ITT analysis.

DISCUSSION

In this target trial emulation we compared two core regimens in addition to other drugs being concurrently received for treatment of MDR-TB: all three WHO group A drugs (bedaquiline, linezolid, and a fluoroquinolone) but without clofazimine (control group) to the same three group A drugs but with the addition of clofazimine (clofazimine group). Although confidence intervals included the null, our ITT point-estimates indicated some evidence that the addition of clofazimine increases successful treatment outcomes by about 7% compared to the control group. The occurrence of adverse events that led to drug stoppage were similar between treatment groups in the ITT analysis, and clofazimine was stopped due to an adverse event in only about 3% of those receiving clofazimine from treatment initiation. In the second per-protocol analysis results were similar to the IPTW estimates in the ITT analysis. However, in the first and third per-protocol analyses, there was no effect on treatment success when clofazimine was added to the core regimen of the three group A drugs once accounting for time-varying confounding with the IPCW that included censoring patients in the control who started clofazimine, with a similar attenuation of the IPCW estimate in the fourth per-protocol analysis.

Interpreting the estimates for each analysis requires some considerations. The ITT analysis included all subjects who initiated their respective regimen disregarding drug stoppages

within their initial core regimen and not accounting for subjects initially in the control group who eventually started taking clofazimine. However, the first and third per-protocol analyses censored patients in the control group who started clofazimine and adjusted for factors predicting this censoring with use of IPCW. In the first per-protocol analysis, the censoring due to only other reasons to stop a drug in their regimen, as well as censoring those in the control group who started clofazimine, resulted in substantial changes to the ATE compared to their respective unadjusted and baseline IPTW estimates in the same population as well as the estimates in the ITT analysis. Additionally, once accounting for the additional censoring in the third per-protocol analysis, there was no effect of adding clofazimine. In the first and third per-protocol analyses, uncensored patients in the control groups were those who remained on their initial core regimen for their entire course of treatment without the need to start clofazimine. In these weighted analyses it is not expected to see a difference in treatment success, as subjects who are able to adhere to all drugs initiated will likely have higher chances of successful treatment outcomes regardless of initial regimen.[29] Results of the TB PRACTECAL randomized controlled trial, investigating shorter treatments of bedaquiline, linezolid, and pretomanid regimens (supplemented with either moxifloxacin or clofazimine) compared to standard WHO regimens for treatment of RR/MDR-TB, indicated a similar substantial ITT effect for each intervention but an attenuated effect between regimens in per-protocol analyses of protocol-adherent patients.[30] Our results also indicate that regimens containing all three group A drugs are already effective regimens as treatment success was over 70% in each treatment group. It is also important to note that few patients (2.9%) in the clofazimine group stopped taking clofazimine for an adverse event.

The use of censoring weights with adjustment for time-varying confounders in the first and third per-protocol analyses substantially attenuated the estimate of the ATE compared to their respective unadjusted and baseline IPTW estimates of the same patients as well as compared to estimates in the ITT analysis. These results indicate that there is an important impact of IPCW and time-varying confounding under these types of censoring. Clofazimine is relatively safe compared to other drugs given for treatment of MDR-TB with lower risk of adverse events.[10,31] It is possible patients in the control group were receiving other concurrent drugs in addition to their core regimens that were stopped due to adverse events, which resulted in clofazimine being added to their treatment regimen. In this situation, patients that remained uncensored who were similar to censored patients would be given more weight due to their lower probability of remaining uncensored, and their favourable outcomes would replace the unfavourable outcomes of those who were censored (indeed the final censoring weights in the first and third per-protocol analyses were larger for those in the control groups than the clofazimine groups). Another possible explanation for the discrepancy between estimates in the ITT analysis and per-protocol analyses, is that patients who required the addition of clofazimine may have been failing treatment, and including their outcomes in the control group for the ITT analysis made the ITT effect seem larger in the clofazimine group, which was not observed in the first and third per-protocol analyses once these patients were censored.

Our study had some limitations. We had a small sample size of patients (especially in the control groups and even fewer after censoring) initiating the two core regimens, and power

to detect significant differences between the treatment groups was limited. As the lower bound of the confidence intervals for the ITT analysis were near the null, a more powered analysis may find a significant treatment effect. Our small sample size also precluded our ability to assess ATEs in subgroup analyses for important clinical populations (such as DST results for FQ, people living with HIV, and patients with extensive disease[32]). Additionally, we were not able to restrict our analysis to those who were susceptible to all drugs in each treatment group as drug susceptibility testing of all drugs in the core regimens was not performed for most patients. Attempts were made to account for this, including adjustment for baseline resistance for all drugs in each regimen and the third and fourth per-protocol analyses accounted for censoring due to acquired drug resistance. We also assessed effect measure modification due to susceptibility or resistance to FQ on DST in the ITT population and found evidence that treatment outcomes are modified by DST results for FQ. Accounting for FQ resistance is important in treatment of MDR-TB, as resistance to FQ is a major predictor of poor treatment outcomes,[14,33] and can affect the efficacy of a treatment regimen. This effect modification should be confirmed in studies with larger sample sizes. We also did not have reasons for patients starting clofazimine in the control group, so a detailed understanding of why a patient started is not possible (although possibly added to replace the drugs concurrently received in addition to their core regimens that were stopped due to adverse events or prescribed to patients who were failing treatment). Additionally, as censoring for starting of clofazimine in the control group resulted in the largest attenuation of the ATE, there is a possibility that the time-varying covariates we used to calculate censoring weights were not adequate to predict censoring due to the initiation of clofazimine. The estimates should be interpreted cautiously, and future studies should

consider recording data on reasons for starting drugs after baseline. Of course, the lack of randomization still allows for unmeasured confounding, which is unavoidable with observational data. However, the target trial approach we used aimed to reduce other biases that are common in observational research by specifically outlining our hypothetical target trial so that focus could be on controlling for measured confounding,[20,34,35] and we were able to adjust for important baseline and time-varying prognostic factors.

This analysis also had several strengths. The data used had detailed records of drugs used in treatment with their start and stop dates (including reason for stoppage), and the data were collected with the purpose of accounting for time-varying confounders.[23] The use of target trial emulation and the IPCW methods to address time-varying confounding helped to account for the biases that usually affect analyses of drugs given at treatment initiation alone.[15,16] In our target trial protocol we had explicit eligibility criteria including timing of treatment initiation, and we only included patients who received their core treatment intervention at baseline which reduced potential immortal-time bias.[35] Additionally, with our well-defined per-protocol analyses we able to identify potential sources of non-adherence, and account for such deviations from our treatment strategies with use of censoring weights, which is especially important when studying the sustained treatment strategies inherent in MDR-TB treatment.[36,37]

When randomized trials cannot be conducted due to cost, time, or feasibility, an emulation of a target trial is a useful alternative to obtain a more valid causal effect estimate for a research question using observational data.[34] Considering the substantial reduction of the

unadjusted and baseline IPTW effect estimates observed in the first and third per-protocol analyses when censoring weights were applied, researchers in future studies may want to consider applying methods to account for the time-varying confounding these censoring factors appear to create.

With this target trial emulation we found some evidence in the ITT analysis and the second per-protocol analysis that adding clofazimine to the three WHO group A drugs provides some benefit for treatment success without evidence of added toxicity. However, once accounting for censoring from treatment strategies for controls who started clofazimine and for time-varying confounding using censoring probabilities in the first and third per-protocol analyses, the added benefit of clofazimine was no longer observed. The results of this study should be confirmed in future research.

ACKNOWLEDGMENTS

We thank Dr. Molly F. Franke and Dr. Letizia Trevisi for their assistance with data queries as well as their guidance with the time-varying analyses and target trial design. We also want to acknowledge Unitaid and the EndTB consortium for providing the data for this study.

Table 5.1. Baseline characteristics of the overall study population by treatment group (intention-to-treat analysis).

	Control group	Clofazimine group	Overall
	n = 197	n = 594	n = 791
Sex = Female (%)	83 (42.1)	230 (38.7)	313 (39.6)
Age (mean (SD))	37 (12)	37 (13)	37 (13)
Body mass index kg/m² (mean (SD))	19 (4)	21 (4)	20 (4)
Diabetes (%)			
Yes	28 (14.2)	81 (13.6)	109 (13.8)
No	169 (85.8)	507 (85.4)	676 (85.5)
Missing	0 (0.0)	6 (1.0)	6 (0.8)
Renal insufficiency (%)			
Yes	3 (1.5)	39 (6.6)	42 (5.3)
No	180 (91.4)	503 (84.7)	683 (86.3)
Missing	14 (7.1)	52 (8.8)	66 (8.3)
HIV infection status (%)			
Positive	11 (5.6)	48 (8.1)	59 (7.5)
Negative	186 (94.4)	546 (91.9)	732 (92.5)
Missing	0 (0.0)	0 (0.0)	0 (0.0)
Alcohol use ever (%)			
Yes	21 (10.7)	67 (11.3)	88 (11.1)
No	169 (85.8)	517 (87.0)	686 (86.7)
Missing	7 (3.6)	10 (1.7)	17 (2.1)
Smoking ever (%)			
Yes	45 (22.8)	136 (22.9)	181 (22.9)
No	148 (75.1)	446 (75.1)	594 (75.1)
Missing	4 (2.0)	12 (2.0)	16 (2.0)
Intravenous Drug use ever (%)			
Yes	1 (0.5)	7 (1.2)	8 (1.0)
No	191 (97.0)	569 (95.8)	760 (96.1)
Missing	5 (2.5)	18 (3.0)	23 (2.9)
Bilateral disease on x-ray (%)			
Yes	129 (65.5)	369 (62.1)	498 (63.0)
No	58 (29.4)	191 (32.2)	249 (31.5)
Missing	10 (5.1)	34 (5.7)	44 (5.6)
Cavitation on x-ray (%)			
Yes	117 (59.4)	350 (58.9)	467 (59.0)
No	70 (35.5)	196 (33.0)	266 (33.6)
Missing	10 (5.1)	48 (8.1)	58 (7.3)
Culture smear status (%)			
Positive	105 (53.3)	277 (46.6)	382 (48.3)
Negative	79 (40.1)	277 (46.6)	356 (45.0)
Missing	13 (6.6)	40 (6.7)	53 (6.7)

Table 5.1 continued.

	Control group	Clofazimine group	Overall
	n = 197	n = 594	n = 791
ECOG functional status (%)			
Fully active	82 (41.6)	256 (43.1)	338 (42.7)
Limited self care	5 (2.5)	15 (2.5)	20 (2.5)
Ambulatory	88 (44.7)	247 (41.6)	335 (42.4)
Completely disabled	3 (1.5)	8 (1.3)	11 (1.4)
Missing	19 (9.6)	68 (11.4)	87 (11.0)
Past tuberculosis treatment (%)			
Yes	167 (84.8)	553 (93.1)	720 (91.0)
No	28 (14.2)	30 (5.1)	58 (7.3)
Missing	2 (1.0)	11 (1.9)	13 (1.6)
Past tuberculosis treatment category (%)			
None	28 (14.2)	30 (5.1)	58 (7.3)
Prior treatment only with first-line drugs	36 (18.3)	50 (8.4)	86 (10.9)
Prior treatment with second-line drugs	131 (66.5)	503 (84.7)	634 (80.2)
Missing	2 (1.0)	11 (1.9)	13 (1.6)
Past first-line tuberculosis drug use (%)			
Yes	36 (18.3)	50 (8.4)	86 (10.9)
No	159 (80.7)	533 (89.7)	692 (87.5)
Missing	2 (1.0)	11 (1.9)	13 (1.6)
Past second-line tuberculosis drug use (%)			
Yes	131 (66.5)	503 (84.7)	634 (80.2)
No	64 (32.5)	80 (13.5)	144 (18.2)
Missing	2 (1.0)	11 (1.9)	13 (1.6)
MDR-TB resistance category (%)			
MDR-TB FQ &SLI sensitive	62 (31.5)	167 (28.1)	229 (29.0)
MDR-TB + SLI resistant & FQ sensitive	21 (10.7)	172 (29.0)	193 (24.4)
MDR-TB + FQ resistant & SLI sensitive	68 (34.5)	91 (15.3)	159 (20.1)
MDR-TB + SLI & FQ resistance	22 (11.2)	139 (23.4)	161 (20.4)
Missing	24 (12.2)	25 (4.2)	49 (6.2)
DST for bedaquiline = Susceptible (%)	2 (1.0)	16 (2.7)	18 (2.3)
DST for linezolid (%)			
Not tested	194 (98.5)	569 (96.6)	763 (97.1)
Resistant	1 (0.5)	0 (0.0)	1 (0.1)
Susceptible	2 (1.0)	20 (3.4)	22 (2.8)
DST for clofazimine (%)			
Not tested	195 (99.0)	574 (97.5)	769 (97.8)
Resistant	2 (1.0)	3 (0.5)	5 (0.6)
Susceptible	0 (0.0)	12 (2.0)	12 (1.5)

Table 5.1. Continued.

	Control group	Clofazimine group	Overall
	n = 197	n = 594	n = 791
DST for levofloxacin (%)			
Not tested	105 (53.3)	237 (40.2)	342 (43.5)
Resistant	38 (19.3)	125 (21.2)	163 (20.7)
Susceptible	54 (27.4)	227 (38.5)	281 (35.8)
DST for moxifloxacin (%)			
Not tested	91 (46.2)	289 (49.1)	380 (48.3)
Resistant	34 (17.3)	85 (14.4)	119 (15.1)
Susceptible	72 (36.5)	215 (36.5)	287 (36.5)
Likely effective for levofloxacin (%)			
Likely effective	73 (37.1)	177 (29.8)	250 (31.6)
Not effective	37 (18.8)	110 (18.5)	147 (18.6)
Missing	87 (44.2)	307 (51.7)	394 (49.8)
Likely effective for moxifloxacin (%)			
Likely effective	53 (26.9)	210 (35.4)	263 (33.2)
Not effective	35 (17.8)	98 (16.5)	133 (16.8)
Missing	109 (55.3)	286 (48.1)	395 (49.9)
DST for fluoroquinolone (with likely effectiveness) (%)†			
Not tested	35 (17.8)	95 (16.0)	130 (16.4)
Resistant	42 (21.3)	132 (22.2)	174 (22.0)
Susceptible	120 (60.9)	367 (61.8)	487 (61.6)
Number of effective drugs used at baseline (median [IQR])	4 [4, 5]	4 [4, 5]	4 [4, 5]
Number of drugs used at baseline (median [IQR])	6 [6, 7]	6 [6, 7]	6 [6, 7]
Treatment duration in months (mean (SD))	18.8 (5.9)	18.6 (5.9)	18.6 (5.9)
Country treated (%)			
Armenia	5 (2.5)	5 (0.8)	10 (1.3)
Bangladesh	23 (11.7)	50 (8.4)	73 (9.2)
Belarus	0 (0.0)	28 (4.7)	28 (3.5)
Ethiopia	4 (2.0)	17 (2.9)	21 (2.7)
Georgia	15 (7.6)	41 (6.9)	56 (7.1)
Haiti	1 (0.5)	4 (0.7)	5 (0.6)
Indonesia	8 (4.1)	6 (1.0)	14 (1.8)
Kazakhstan	80 (40.6)	232 (39.1)	312 (39.4)
Kenya	0 (0.0)	1 (0.2)	1 (0.1)
Kyrgyzstan	1 (0.5)	7 (1.2)	8 (1.0)
Lesotho	6 (3.0)	24 (4.0)	30 (3.8)
Myanmar	0 (0.0)	4 (0.7)	4 (0.5)
Pakistan	43 (21.8)	62 (10.4)	105 (13.3)
Peru	4 (2.0)	90 (15.2)	94 (11.9)
South Africa	3 (1.5)	21 (3.5)	24 (3.0)
Vietnam	4 (2.0)	2 (0.3)	6 (0.8)

† Levofloxacin and moxifloxacin DST were combined into one variable; if moxifloxacin DST was missing it was replaced with levofloxacin DST, and if both fluoroquinolones DST were missing, they were considered likely effective if there was no history of past treatment use with either of these drugs. If the DST was missing for both FQ, and information regarding prior use of both of these FQ was also missing, then the missing DST for this FQ variable were imputed.

Table 5.2. Drugs received in addition to drugs in the core regimens, by treatment group in the overall population.

	Received at baseline			Received at any time during treatment		
	Control group	Clofazimine group	Overall	Control group	Clofazimine group	Overall
n	**197**	**594**	**791**	**197**	**594**	**791**
Ethambutol	32 (16.2)	36 (6.1)	68 (8.6)	33 (16.8)	43 (7.2)	76 (9.6)
Pyrazinamide	151 (76.6)	333 (56.1)	484 (61.2)	154 (78.2)	342 (57.6)	496 (62.7)
Kanamycin	7 (3.6)	29 (4.9)	36 (4.6)	8 (4.1)	35 (5.9)	43 (5.4)
Capreomycin	51 (25.9)	131 (22.1)	182 (23.0)	54 (27.4)	147 (24.7)	201 (25.4)
Amikacin	61 (31.0)	64 (10.8)	125 (15.8)	63 (32.0)	75 (12.6)	138 (17.4)
Prothionamide / Ethionamide	109 (55.3)	152 (25.6)	261 (33.0)	109 (55.3)	152 (25.6)	261 (33.0)
Cycloserine / Terizidone	155 (78.7)	306 (51.5)	461 (58.3)	158 (80.2)	325 (54.7)	483 (61.1)
Para-aminosalicylic acid	62 (31.5)	138 (23.2)	200 (25.3)	70 (35.5)	152 (25.6)	222 (28.1)
Cilastatin/Imipenem/Meropenem	7 (3.6)	68 (11.4)	75 (9.5)	14 (7.1)	86 (14.7)	100 (12.6)
Amoxicillin-Clavulanic Acid	9 (4.6)	91 (15.3)	100 (12.6)	16 (8.1)	113 (19.0)	129 (16.3)
Delamanid	14 (7.1)	40 (6.7)	54 (6.8)	29 (14.7)	99 (16.7)	128 (16.2)

Note only 2 patients received streptomycin at treatment initiation, while 2 patients received clarithromycin later in treatment (all were in the clofazimine group).

Table 5.3. Average treatment effects and 95% confidence intervals (CI) between the clofazimine group and the control group among each analysis population and by method of analysis.

	Treatment success / total n (%)			ATE (95%CI)
	Control group	Clofazimine group	n total	Clofazimine group compared to control group
Intention-to-treat analysis				
Crude/unadjusted				0.05 (-0.02, 0.12)
Fully adjusted	144 / 197 (73.1)	466 / 594 (78.5)	791	0.03 (-0.04, 0.11)
IPTW baseline				0.07 (-0.01, 0.15)
1st Per-Protocol analysis				
Crude/unadjusted				0.08 (-0.02, 0.17)
IPTW baseline	77 / 107 (72)	453 / 569 (79.6)	676	0.09 (-0.02, 0.19)
IPCW time-varying				-0.01 (-0.10, 0.07)
2nd Per-Protocol analysis				
Crude/unadjusted				0.06 (-0.01, 0.13)
IPTW baseline	139 / 188 (73.9)	453 / 569 (79.6)	757	0.07 (-0.01, 0.15)
IPCW time-varying				0.07 (-0.01, 0.15)
3rd Per-Protocol analysis				
Crude/unadjusted				0.09 (-0.01, 0.19)
IPTW baseline	67 / 94 (71.3)	336 / 421 (79.8)	515	0.10 (-0.02, 0.22)
IPCW time-varying				0.02 (-0.05, 0.09)
4th Per-Protocol analysis				
Crude/unadjusted				0.05 (-0.04, 0.13)
IPTW baseline	103 / 137 (75.2)	336 / 421 (79.8)	558	0.06 (-0.04, 0.15)
IPCW time-varying				0.04 (-0.03, 0.11)

ATE: Average treatment effect; CI: confidence interval; IPTW inverse probability of treatment weight; IPCW: inverse probability of censor weights. Fully adjusted models adjusted for the same variables included in baseline IPTW. Treatment success defined as cure or completed compared to all negative outcomes. Note: IPCW time-varying estimates were calculated with pooling and multiplied by baseline IPTW for final weights. Fully adjusted models and IPTW included the following variables: age, sex, diabetes, HIV infection, body mass index category, renal insufficiency, current alcohol use, current smoker, intravenous drug use, bilateral disease, cavitation on chest radiography, smear status, culture result, Eastern Cooperative Oncology Group (ECOG) functional status at baseline, past tuberculosis drug use, MDR category (only MDR-TB, MDR-TB plus resistance to second-line injectable (SLI) but FQ sensitive, MDR-TB plus resistance to FQ but SLI sensitive, and MDR-TB plus resistance to both FQ and SLI), drug susceptibility testing for linezolid, clofazimine, and likely effective fluoroquinolone. IPCW estimates included were adjusted for the same baseline variables above but with these additional time-varying covariates: body mass index category, culture result, ECOG status, highest count of adverse events, and highest count of serious adverse events.

Table 5.4. Drugs received that were permanently stopped due to an adverse event at any time during treatment.

	Intention-to-treat population	
	Control group n = 197	**Clofazimine group n = 594**
	n (%)	**n (%)**
Drugs in the core regimens		
Linezolid	18 (9.1)	39 (6.6)
Bedaquiline	0	13 (2.2)
Levofloxacin	1 (0.5)	14 (2.4)
Moxifloxacin	4 (2.0)	23 (3.9)
Clofazimine	2 (1.0)†	17 (2.9)
Drugs received in addition to the core regimens		
Ethambutol	2 (6.1)	8 (18.6)
Pyrazinamide	33 (21.4)	63 (18.4)
Kanamycin	2 (25.0)	11 (31.4)
Capreomycin	8 (14.8)	19 (12.9)
Amikacin	11 (17.5)	14 (18.7)
Clarithromycin	0	1 (50.0)
Prothionamide	16 (34.8)	18 (30.5)
Ethionamide	10 (14.1)	20 (20.4)
Cycloserine	18 (11.4)	35 (10.7)
Terizidone	0	3 (15.8)
Para-aminosalicylic acid	13 (18.6)	32 (21.1)
Imipenem-Cilastatin	0	21 (24.7)
Amoxicillin-Clavulanic acid	0	21 (18.6)
Delamanid	1 (3.4)	1 (1.0)
Number of drugs stopped for AEs per patient		
One drug stopped	57 (28.9)	159 (26.8)
More than one drug stopped	35 (17.8)	84 (14.1)
No drugs stopped	105 (53.3)	351 (59.1)

† Received clofazimine post-baseline assignment to control group. ITT: Intention-to-treat; AE: adverse event.

Table 5.5. Reasons for earliest censoring event in all per-protocol analyses.

	1st Per-protocol analysis		2nd Per-protocol analysis		3rd Per-protocol analysis		4th Per-protocol analysis	
	Control group	**Clofazimine group**	**Control group**	**Clofazimine group**	**Control group**	**Clofazimine group**	**Control group**	**Clofazimine group**
	n = 197	**n = 594**	**n = 197**	**n = 594**	**n = 197**	**n = 594**	**n = 197**	**n = 594**
Reason for censoring, n (%)*								
Adverse event					10 (5.1)	67 (11.3)	22 (11.2)	67 (11.3)
Drug supply or drug administration issue					1 (0.5)	2 (0.3)	2 (1)	2 (0.3)
Planned change					14 (7.1)	50 (8.4)	17 (8.6)	50 (8.4)
Pregnancy					0 (0)	1 (0.2)	0 (0)	1 (0.2)
Reintroduction/replacement of stopped drug					1 (0.5)	4 (0.7)	3 (1.5)	4 (0.7)
Resistance to drug					0 (0)	27 (4.5)	8 (4.1)	27 (4.5)
Other reasons	5 (2.5)	25 (4.2)	9 (4.6)	25 (4.2)	5 (2.5)	22 (3.7)	8 (4.1)	22 (3.7)
Clofazimine started**	85 (43.1)				72 (36.5)			
Not censored	107 (54.3)	569 (95.8)	188 (95.4)	569 (95.8)	94 (47.7)	421 (70.9)	137 (69.5)	421 (70.9)

*Note some patients had multiple reasons for censoring across per-protocol populations and censoring was defined differently in each population with the earliest censor reason being reported, therefore the 'other' category used to define one per-protocol analysis may not equal the other per-protocol analyses. **Note: 85 subjects in the control group started clofazimine for at least 30 days at some point in their treatment and were censored (the subject's earliest censoring event is reported).

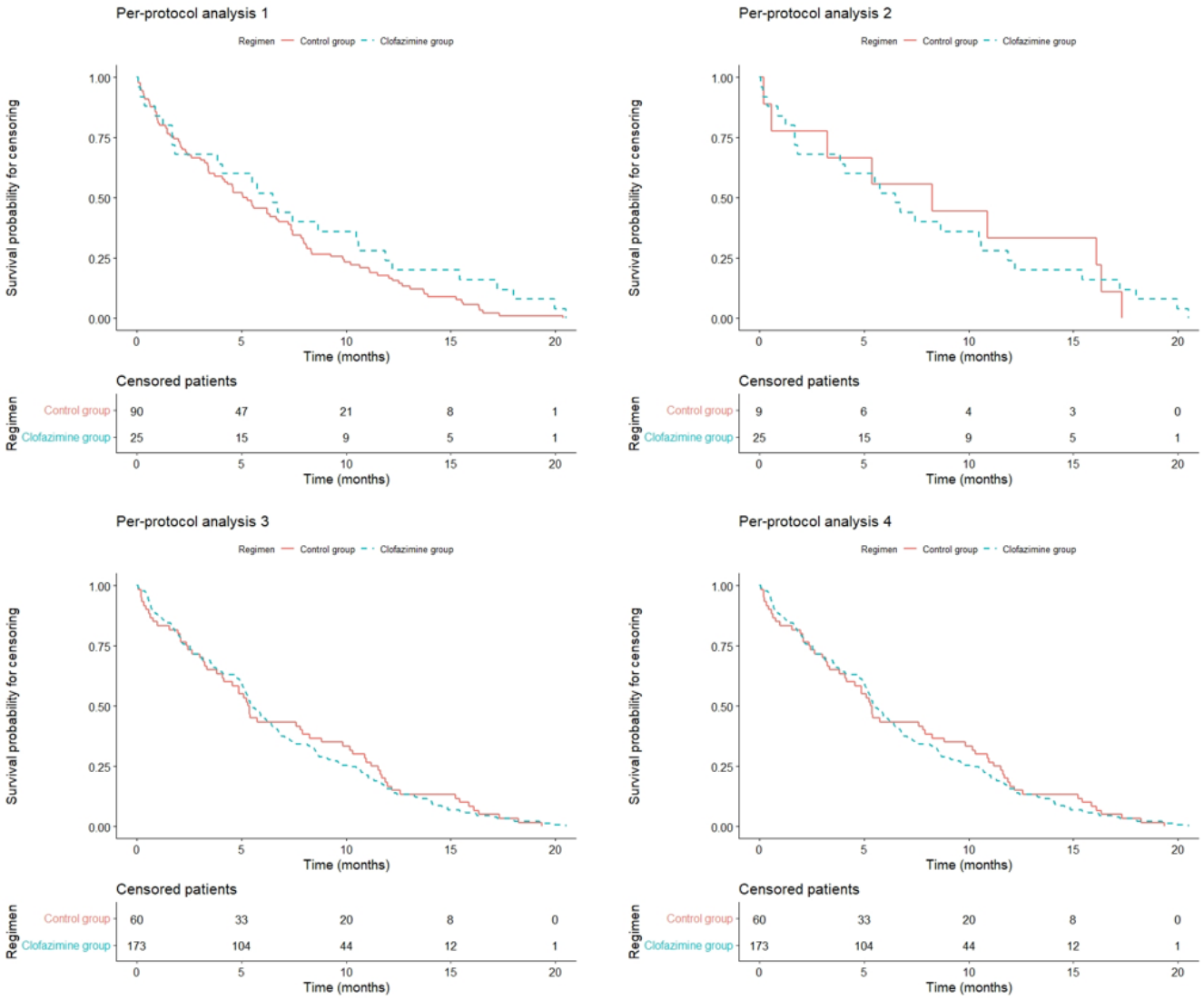


Figure 5.1. Survival probability curves for those who were censored in all per-protocol analyses. The 1st per-protocol analysis: the treatment strategy involved remaining on the assigned treatment but allowing for permanent drug stoppages for these reasons: adverse events, planned treatment changes, reintroduction or replacement of stopped drug, resistance to drug, drug supply or drug administration issues, or pregnancy but censoring for any other reason a patient stopped a drug as well as censoring patients in the control group who start clofazimine. 2nd Per-protocol analysis: the treatment strategy was the same as in the 1st per-protocol analysis but allowed patients in the control to start clofazimine without being censored. 3rd Per-protocol analysis: the treatment strategy involved remaining on the assigned treatment but censoring for any of the above-mentioned reasons for drug stoppages, as well as censoring patients in the control group who start clofazimine. 4th Per-protocol analysis, the treatment strategy was the same as in the 3rd per-protocol analysis but allowed patients in the control to start clofazimine without being censored.

Section 5.3 Supplemental material

Supplemental Table 5.S1. Target trial specification and emulation using the EndTB cohort.

Protocol component	Target trial	Emulation in EndTB data
Eligibility	Patients of any age initiating MDR-TB treatment who are eligible at baseline to receive all drugs in each treatment regimen.	Same
Treatment strategies	Initiation at baseline, in addition to other drugs being concurrently received for MDR-TB treatment at the discretion of the provider, of the core regimens of all three WHO group A drugs (bedaquiline, linezolid, and a FQ) without clofazimine (control group) or the same three WHO group A drugs (bedaquiline, linezolid, and a FQ) plus the addition of clofazimine (clofazimine group). 1st per-protocol analysis: the treatment strategy involved remaining on the assigned treatment but allowing for permanent drug stoppages for these reasons: adverse events, planned treatment changes, reintroduction or replacement of stopped drug, resistance to drug, drug supply or drug administration issues, or pregnancy but censoring for any other reason a patient stopped a drug as well as censoring patients in the control group who start clofazimine. 2nd Per-protocol analysis: the treatment strategy was the same as in the 1st per-protocol analysis but allowed patients in the control to start clofazimine without being censored. 3rd Per-protocol analysis: the treatment strategy involved remaining on the assigned treatment but censoring for any of the above-mentioned reasons for drug stoppages, as well as censoring patients in the control group who start clofazimine. 4th Per-protocol analysis, the treatment strategy was the same as in the 3rd per-protocol analysis but allowed patients in the control to start clofazimine without being censored.	Same
Assignment procedures	MDR-TB patients would be randomized to the control group or the clofazimine group	Individuals are assigned to each treatment group at baseline
Follow-up period	Patients are followed from treatment initiation until their end of treatment outcome.	Patients are followed until their end of treatment outcome or censor
Outcome	Treatment success was defined as cure or completion of treatment, compared to all other negative outcomes (death, treatment failure [defined as microbiologic failure or a change in any two drugs received due to adverse events including the supplemental, concurrent MDR-TB drugs in addition to core regimens], or lost to follow up) as defined in WHO 2013.	Same
Causal contrast	We will estimate the intention-to-treat (ITT) effect and the per-protocol effect	Observational analogue of the ITT and per-protocol effects.
Analysis plan	The ITT analysis involves direct comparison of the proportion of patients with treatment success among those assigned to each treatment. The per-protocol analyses will censor any patient deviating from the respective treatment strategies described above. The per-protocol analyses will adjust for both baseline and post-baseline confounders of treatment censoring.	Same

MDR-TB: multidrug-resistant tuberculosis; ITT: intention-to-treat; FQ: fluoroquinolone; IPTW: inverse probability of treatment weight; IPCW: inverse probability of censor weight.

Supplemental Table 5.S2A. Baseline characteristics of patients in the per-protocol analyses by censor status.

	1st Per-protocol analysis				2nd Per-protocol analysis			
	Uncensored		Censored		Uncensored		Censored	
	Control group	Clofazimine group	Control group	Clofazimine group	Control group	Clofazimine group	Control group	Clofazimine group
n	**107**	**569**	**90**	**25**	**188**	**569**	**9**	**25**
Sex = Female (%)	50 (46.7)	224 (39.4)	33 (36.7)	6 (24.0)	81 (43.1)	224 (39.4)	2 (22.2)	6 (24.0)
Age (mean (SD))	35 (11)	36(13)	38 (13)	39 (12)	36 (12)	36 (13)	44 (12)	39 (12)
Body mass index kg/m² (mean (SD))	19 (4)	21 (4)	20 (4)	20 (3)	19.4 (4.1)	20.5 (4.4)	20.7 (4.2)	20.1 (3.0)
Diabetes (%)								
Yes	12 (11.2)	78 (13.7)	16 (17.8)	3 (12.0)	25 (13.3)	78 (13.7)	3 (33.3)	3 (12.0)
No	95 (88.8)	486 (85.4)	74 (82.2)	21 (84.0)	163 (86.7)	486 (85.4)	6 (66.7)	21 (84.0)
Missing	0 (0.0)	5 (0.9)	0 (0.0)	1 (4.0)	0 (0)	5 (0.9)	0 (0)	1 (4.0)
Renal insufficiency (%)								
Yes	2 (1.9)	39 (6.9)	1 (1.1)	0 (0.0)	3 (1.6)	39 (6.9)	0 (0)	0 (0)
No	98 (91.6)	479 (84.2)	82 (91.1)	24 (96.0)	171 (91.0)	479 (84.2)	9 (100)	24 (96.0)
Missing	7 (6.5)	51 (9.0)	7 (7.8)	1 (4.0)	14 (7.4)	51 (9.0)	0 (0)	1 (4.0)
HIV infection status (%)								
Positive	6 (5.6)	43 (7.6)	5 (5.6)	5 (20.0)	11 (5.9)	43 (7.6)	0 (0)	5 (20.0)
Negative	101 (94.4)	526 (92.4)	85 (94.4)	20 (80.0)	177 (94.1)	526 (92.4)	9 (100)	20 (80.0)
Missing	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0)	0 (0)	0 (0)	0 (0)
Alcohol use ever (%)								
Yes	13 (12.1)	63 (11.1)	8 (8.9)	4 (16.0)	19 (10.1)	63 (11.1)	2 (22.2)	4 (16.0)
No	93 (86.9)	496 (87.2)	76 (84.4)	21 (84.0)	164 (87.2)	496 (87.2)	5 (55.6)	21 (84.0)
Missing	1 (0.9)	10 (1.8)	6 (6.7)	0 (0.0)	5 (2.7)	10 (1.8)	2 (22.2)	0 (0)
Smoking ever (%)								
Yes	23 (21.5)	127 (22.3)	22 (24.4)	9 (36.0)	40 (21.3)	127 (22.3)	5 (55.6)	9 (36.0)
No	83 (77.6)	432 (75.9)	65 (72.2)	14 (56.0)	144 (76.6)	432 (75.9)	4 (44.4)	14 (56.0)
Missing	1 (0.9)	10 (1.8)	3 (3.3)	2 (8.0)	4 (2.1)	10 (1.8)	0 (0)	2 (8.0)
Intravenous Drug use ever (%)								
Yes	1 (0.9)	6 (1.1)	0 (0.0)	1 (4.0)	1 (0.5)	6 (1.1)	0 (0)	1 (4.0)
No	105 (98.1)	546 (96.0)	86 (95.6)	23 (92.0)	182 (96.8)	546 (96.0)	9 (100)	23 (92.0)
Missing	1 (0.9)	17 (3.0)	4 (4.4)	1 (4.0)	5 (2.7)	17 (3.0)	0 (0)	1 (4.0)
Bilateral disease on x-ray (%)								
Yes	70 (65.4)	353 (62.0)	59 (65.6)	16 (64.0)	123 (65.4)	353 (62.0)	6 (66.7)	16 (64.0)
No	29 (27.1)	184 (32.3)	29 (32.2)	7 (28.0)	55 (29.3)	184 (32.3)	3 (33.3)	7 (28.0)
Missing	8 (7.5)	32 (5.6)	2 (2.2)	2 (8.0)	10 (5.3)	32 (5.6)	0 (0)	2 (8.0)
Cavitation on x-ray (%)								
Yes	56 (52.3)	333 (58.5)	61 (67.8)	17 (68.0)	110 (58.5)	333 (58.5)	7 (77.8)	17 (68.0)
No	43 (40.2)	190 (33.4)	27 (30.0)	6 (24.0)	68 (36.2)	190 (33.4)	2 (22.2)	6 (24.0)
Missing	8 (7.5)	46 (8.1)	2 (2.2)	2 (8.0)	10 (5.3)	46 (8.1)	0 (0)	2 (8.0)
Culture smear status (%)								
Positive	59 (55.1)	266 (46.7)	46 (51.1)	11 (44.0)	101 (53.7)	266 (46.7)	4 (44.4)	11 (44.0)
Negative	41 (38.3)	263 (46.2)	38 (42.2)	14 (56.0)	74 (39.4)	263 (46.2)	5 (55.6)	14 (56.0)
Missing	7 (6.5)	40 (7.0)	6 (6.7)	0 (0.0)	13 (6.9)	40 (7.0)	0 (0)	0 (0)

Supplemental Table 5.S2A. Baseline characteristics of patients in the per-protocol analyses by censor status.

	1st Per-protocol analysis				2nd Per-protocol analysis			
	Uncensored		Censored		Uncensored		Censored	
	Control group	Clofazimine group	Control group	Clofazimine group	Control group	Clofazimine group	Control group	Clofazimine group
n	**107**	**569**	**90**	**25**	**188**	**569**	**9**	**25**
Functional status (ECOG) (%)								
Fully active	49 (45.8)	253 (44.5)	33 (36.7)	3 (12.0)	79 (42.0)	253 (44.5)	3 (33.3)	3 (12.0)
Limited self care	2 (1.9)	14 (2.5)	3 (3.3)	1 (4.0)	5 (2.7)	14 (2.5)	0 (0)	1 (4.0)
Ambulatory	43 (40.2)	234 (41.1)	45 (50.0)	13 (52.0)	83 (44.1)	234 (41.1)	5 (55.6)	13 (52.0)
Completely disabled	2 (1.9)	7 (1.2)	1 (1.1)	1 (4.0)	3 (1.6)	7 (1.2)	0 (0)	1 (4.0)
Missing	11 (10.3)	61 (10.7)	8 (8.9)	7 (28.0)	18 (9.6)	61 (10.7)	1 (11.1)	7 (28.0)
Past TB treatment (%)								
Yes	88 (82.2)	529 (93.0)	79 (87.8)	24 (96.0)	158 (84.0)	529 (93.0)	9 (100)	24 (96.0)
No	18 (16.8)	30 (5.3)	10 (11.1)	0 (0.0)	28 (14.9)	30 (5.3)	0 (0)	0 (0)
Missing	1 (0.9)	10 (1.8)	1 (1.1)	1 (4.0)	2 (1.1)	10 (1.8)	0 (0)	1 (4.0)
Past treatment category (%)								
None	18 (16.8)	30 (5.3)	10 (11.1)	0 (0.0)	28 (14.9)	30 (5.3)	0 (0)	0 (0)
Prior treatment only with first line drugs	19 (17.8)	48 (8.4)	17 (18.9)	2 (8.0)	35 (18.6)	48 (8.4)	1 (11.1)	2 (8.0)
Prior treatment with second line drugs	69 (64.5)	481 (84.5)	62 (68.9)	22 (88.0)	123 (65.4)	481 (84.5)	8 (88.9)	22 (88.0)
Missing	1 (0.9)	10 (1.8)	1 (1.1)	1 (4.0)	2 (1.1)	10 (1.8)	0 (0)	1 (4.0)
Past first-line TB drug use (%)								
Yes	19 (17.8)	48 (8.4)	17 (18.9)	2 (8.0)	35 (18.6)	48 (8.4)	1 (11.1)	2 (8.0)
No	87 (81.3)	511 (89.8)	72 (80.0)	22 (88.0)	151 (80.3)	511 (89.8)	8 (88.9)	22 (88.0)
Missing	1 (0.9)	10 (1.8)	1 (1.1)	1 (4.0)	2 (1.1)	10 (1.8)	0 (0)	1 (4.0)
Past second-line TB drug use (%)								
Yes	69 (64.5)	481 (84.5)	62 (68.9)	22 (88.0)	123 (65.4)	481 (84.5)	8 (88.9)	22 (88.0)
No	37 (34.6)	78 (13.7)	27 (30.0)	2 (8.0)	63 (33.5)	78 (13.7)	1 (11.1)	2 (8.0)
Missing	1 (0.9)	10 (1.8)	1 (1.1)	1 (4.0)	2 (1.1)	10 (1.8)	0 (0)	1 (4.0)
MDR-TB resistance category (%)								
MDR-TB FQ &SLI sensitive	32 (29.9)	161 (28.3)	30 (33.3)	6 (24.0)	59 (31.4)	161 (28.3)	3 (33.3)	6 (24.0)
MDR-TB + SLI resistant & FQ sensitive	10 (9.3)	163 (28.6)	11 (12.2)	9 (36.0)	18 (9.6)	163 (28.6)	3 (33.3)	9 (36.0)
MDR-TB + FQ resistant & SLI sensitive	39 (36.4)	88 (15.5)	29 (32.2)	3 (12.0)	65 (34.6)	88 (15.5)	3 (33.3)	3 (12.0)
MDR-TB + SLI & FQ resistance	11 (10.3)	133 (23.4)	11 (12.2)	6 (24.0)	22 (11.7)	133 (23.4)	0 (0)	6 (24.0)
Missing	15 (14.0)	24 (4.2)	9 (10.0)	1 (4.0)	24 (12.8)	24 (4.2)	0 (0)	1 (4.0)
DST for bedaquiline = Susceptible (%)	2 (1.9)	15 (2.7)	0 (0)	1 (4.2)	2 (1.1)	15 (2.7)	0 (0)	1 (4.2)
DST for linezolid (%)								
Not tested	105 (98.1)	546 (96.6)	90 (100)	24 (96)	185 (98.4)	546 (96.6)	9 (100)	24 (96.8)
Resistant	1 (0.9)	0 (0.0)	0	0	1 (0.5)	0 (0)	0	0
Susceptible	1 (0.9)	19 (3.4)	0 (0)	1 (4.2)	2 (1.1)	19 (3.4)	0 (0)	1 (4.2)
DST for clofazimine (%)								
Not tested	105 (98.1)	551 (97.5)	90 (100)	24 (96)	186 (98.9)	551 (97.5)	9 (100)	24 (96.8)
Resistant	2 (1.9)	3 (0.5)	0	0	2 (1.1)	3 (0.5)	0	0
Susceptible	0 (0.0)	11 (1.9)	0	1 (4.2)	0 (0)	11 (1.9)	0 (0)	1 (4.2)

Supplemental Table 5.S2A. Baseline characteristics of patients in the per-protocol analyses by censor status.

	1st Per-protocol analysis				2nd Per-protocol analysis			
	Uncensored		Censored		Uncensored		Censored	
	Control group	Clofazimine group	Control group	Clofazimine group	Control group	Clofazimine group	Control group	Clofazimine group
n	**107**	**569**	**90**	**25**	**188**	**569**	**9**	**25**
DST for levofloxacin (%)								
Not tested	50 (46.7)	223 (39.5)	55 (61.1)	14 (58.3)	99 (52.7)	223 (39.5)	6 (66.7)	14 (58.3)
Resistant	24 (22.4)	123 (21.8)	14 (15.6)	2 (8.3)	36 (19.1)	123 (21.8)	2 (22.2)	2 (8.3)
Susceptible	33 (30.8)	219 (38.8)	21 (23.3)	8 (33.3)	53 (28.2)	219 (38.8)	1 (11.1)	8 (33.3)
DST for moxifloxacin (%)								
Not tested	40 (37.4)	276 (48.8)	51 (56.7)	13 (54.2)	86 (45.7)	276 (48.8)	5 (55.6)	13 (54.2)
Resistant	21 (19.6)	83 (14.7)	13 (14.4)	2 (8.3)	32 (17.0)	83 (14.7)	2 (22.2)	2 (8.3)
Susceptible	46 (43.0)	206 (36.5)	26 (28.9)	9 (37.5)	70 (37.2)	206 (36.5)	2 (22.2)	9 (37.5)
Likely effective for levofloxacin (%)								
Likely effective	42 (39.3)	170 (29.9)	31 (34.4)	7 (28.0)	72 (38.3)	170 (29.9)	1 (11.1)	7 (28.0)
Not effective	19 (17.8)	105 (18.5)	18 (20.0)	5 (20.0)	33 (17.6)	105 (18.5)	4 (44.4)	5 (20.0)
Missing	46 (43.0)	294 (51.7)	41 (45.6)	13 (52.0)	83 (44.1)	294 (51.7)	4 (44.4)	13 (52.0)
Likely effective for moxifloxacin (%)								
Likely effective	31 (29.0)	201 (35.3)	22 (24.4)	9 (36.0)	52 (27.7)	201 (35.3)	1 (11.1)	9 (36.0)
Not effective	16 (15.0)	94 (16.5)	19 (21.1)	4 (16.0)	32 (17.0)	94 (16.5)	3 (33.3)	4 (16.0)
Missing	60 (56.1)	274 (48.2)	49 (54.4)	12 (48.0)	104 (55.3)	274 (48.2)	5 (55.6)	12 (48.0)
DST for fluoroquinolone (with likely effectiveness) (%)†								
Not tested	14 (13.1)	89 (15.6)	21 (23.3)	6 (24.0)	31 (16.5)	89 (15.6)	4 (44.4)	6 (24.0)
Resistant	26 (24.3)	130 (22.8)	16 (17.8)	2 (8.0)	40 (21.3)	130 (22.8)	2 (22.2)	2 (8.0)
Susceptible	67 (62.6)	350 (61.5)	53 (58.9)	17 (68.0)	117 (62.2)	350 (61.5)	3 (33.3)	17 (68.0)
Number of effective drugs used at baseline (median [IQR])	5 [4, 6]	4 [4, 5]	4 [3, 5]	4 [4, 5]	4 [4, 5]	4 [4, 5]	3 [2, 4]	4 [4, 5]
Number of drugs used at baseline (median [IQR])	6 [6, 7]	6 [6, 7]	6 [6, 7]	6 [5, 7]	6 [6, 7]	6 [6, 7]	6 [6, 7]	6 [5, 7]
Treatment duration in months (mean (SD))	18.06 (6.9)	18.68 (5.8)	19.7 (4.5)	16.3 (7.6)	18.9 (5.8)	18.7 (5.8)	16.0 (7.7)	16.3 (7.7)
Country treated (%)								
Armenia	4 (3.7)	3 (0.5)	1 (1.1)	2 (8.0)	5 (2.7)	3 (0.5)	0 (0)	2 (8.0)
Bangladesh	10 (9.3)	50 (8.8)	13 (14.4)	0 (0.0)	22 (11.7)	50 (8.8)	1 (11.1)	0 (0)
Belarus	0 (0.0)	28 (4.9)	0	0	0 (0)	28 (4.9)	0	0
Ethiopia	3 (2.8)	17 (3.0)	1 (1.1)	0 (0.0)	4 (2.1)	17 (3.0)	0	0
Georgia	6 (5.6)	38 (6.7)	9 (10.0)	3 (12.0)	15 (8.0)	38 (6.7)	0 (0)	3 (12.0)
Haiti	1 (0.9)	4 (0.7)	0	0	1 (0.5)	4 (0.7)	0	0
Indonesia	4 (3.7)	6 (1.1)	4 (4.4)	0 (0.0)	8 (4.3)	6 (1.1)	0	0

Supplemental Table 5.S2A. Baseline characteristics of patients in the per-protocol analyses by censor status.

	1st Per-protocol analysis				2nd Per-protocol analysis			
	Uncensored		Censored		Uncensored		Censored	
	Control group	Clofazimine group	Control group	Clofazimine group	Control group	Clofazimine group	Control group	Clofazimine group
n	**107**	**569**	**90**	**25**	**188**	**569**	**9**	**25**
Country (continued)								
Kazakhstan	39 (36.4)	220 (38.7)	41 (45.6)	12 (48.0)	74 (39.4)	220 (38.7)	6 (66.7)	12 (48.0)
Kenya	0 (0.0)	1 (0.2)	0	0	0 (0)	1 (0.2)	0	0
Kyrgyzstan	0 (0.0)	7 (1.2)	1 (1.1)	0 (0.0)	1 (0.5)	7 (1.2)	0	0
Lesotho	2 (1.9)	22 (3.9)	4 (4.4)	2 (8.0)	6 (3.2)	22 (3.9)	0 (0)	2 (8.0)
Myanmar	0 (0.0)	4 (0.7)	0	0	0 (0)	4 (0.7)	0	0
Pakistan	30 (28.0)	59 (10.4)	13 (14.4)	3 (12.0)	41 (21.8)	59 (10.4)	2 (22.2)	3 (12.0)
Peru	4 (3.7)	89 (15.6)	0 (0.0)	1 (4.0)	4 (2.1)	89 (15.6)	0 (0)	1 (4.0)
South Africa	1 (0.9)	19 (3.3)	2 (2.2)	2 (8.0)	3 (1.6)	19 (3.3)	0 (0)	2 (8.0)
Vietnam	3 (2.8)	2 (0.4)	1 (1.1)	0 (0.0)	4 (2.1)	2 (0.4)	0	0

† Levofloxacin and moxifloxacin DST were combined into one variable; if moxifloxacin DST was missing it was replaced with levofloxacin DST, and if both fluoroquinolones DST were missing, they were considered likely effective if there was no history of past treatment use with either of these drugs. If the DST was missing for both FQ, and information regarding prior use of both of these FQ was also missing, then the missing DST for this FQ variable were imputed.

Supplemental Table 5.S2B. Baseline characteristics of patients in the per-protocol analyses by censor status.

	3rd Per-protocol analysis				4th Per-protocol analysis			
	Uncensored		Censored		Uncensored		Censored	
	Control group	Clofazimine group	Control group	Clofazimine group	Control group	Clofazimine group	Control group	Clofazimine group
n	94	421	103	173	137	421	60	173
Sex = Female (%)	45 (47.9)	161 (38.2)	38 (36.9)	69 (39.9)	60 (43.8)	161 (38.2)	23 (38.3)	69 (39.9)
Age (mean (SD))	35 (12)	35 (13)	38 (13)	39 (13)	36 (12)	35 (13)	38 (13)	39 (13)
Body mass index kg/m² (mean (SD))	19 (4)	20 (4)	20 (4)	21 (4)	19.2 (4.0)	20.4 (4.4)	20.0 (4.2)	20.9 (4.4)
Diabetes (%)								
Yes	12 (12.8)	58 (13.8)	16 (15.5)	23 (13.3)	18 (13.1)	58 (13.8)	10 (16.7)	23 (13.3)
No	82 (87.2)	359 (85.3)	87 (84.5)	148 (85.5)	119 (86.9)	359 (85.3)	50 (83.3)	148 (85.5)
Missing	0 (0.0)	4 (1.0)	0 (0.0)	2 (1.2)	0 (0)	4 (1.0)	0 (0)	2 (1.2)
Renal insufficiency (%)								
Yes	2 (2.1)	28 (6.7)	1 (1.0)	11 (6.4)	2 (1.5)	28 (6.7)	1 (1.7)	11 (6.4)
No	85 (90.4)	353 (83.8)	95 (92.2)	150 (86.7)	127 (92.7)	353 (83.8)	53 (88.3)	150 (86.7)
Missing	7 (7.4)	40 (9.5)	7 (6.8)	12 (6.9)	8 (5.8)	40 (9.5)	6 (10.0)	12 (6.9)
HIV infection status (%)								
Positive	3 (3.2)	18 (4.3)	8 (7.8)	30 (17.3)	6 (4.4)	18 (4.3)	5 (8.3)	30 (17.3)
Negative	91 (96.8)	403 (95.7)	95 (92.2)	143 (82.7)	131 (95.6)	403 (95.7)	55 (91.7)	143 (82.7)
Missing	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0)	0 (0)	0 (0)	0 (0)
Alcohol use ever (%)								
Yes	10 (10.6)	46 (10.9)	11 (10.7)	21 (12.1)	14 (10.2)	46 (10.9)	7 (11.7)	21 (12.1)
No	83 (88.3)	366 (86.9)	86 (83.5)	151 (87.3)	120 (87.6)	366 (86.9)	49 (81.7)	151 (87.3)
Missing	1 (1.1)	9 (2.1)	6 (5.8)	1 (0.6)	3 (2.2)	9 (2.1)	4 (6.7)	1 (0.6)
Smoking ever (%)								
Yes	22 (23.4)	94 (22.3)	23 (22.3)	42 (24.3)	33 (24.1)	94 (22.3)	12 (20.0)	42 (24.3)
No	71 (75.5)	321 (76.2)	77 (74.8)	125 (72.3)	101 (73.7)	321 (76.2)	47 (78.3)	125 (72.3)
Missing	1 (1.1)	6 (1.4)	3 (2.9)	6 (3.5)	3 (2.2)	6 (1.4)	1 (1.7)	6 (3.5)
Intravenous Drug use ever (%)								
Yes	1 (1.1)	3 (0.7)	0 (0.0)	4 (2.3)	1 (0.7)	3 (0.7)	0 (0)	4 (2.3)
No	92 (97.9)	407 (96.7)	99 (96.1)	162 (93.6)	133 (97.1)	407 (96.7)	58 (96.7)	162 (93.6)
Missing	1 (1.1)	11 (2.6)	4 (3.9)	7 (4.0)	3 (2.2)	11 (2.6)	2 (3.3)	7 (4.0)
Bilateral disease on x-ray (%)								
Yes	63 (67.0)	262 (62.2)	66 (64.1)	107 (61.8)	93 (67.9)	262 (62.2)	36 (60.0)	107 (61.8)
No	24 (25.5)	139 (33.0)	34 (33.0)	52 (30.1)	36 (26.3)	139 (33.0)	22 (36.7)	52 (30.1)
Missing	7 (7.4)	20 (4.8)	3 (2.9)	14 (8.1)	8 (5.8)	20 (4.8)	2 (3.3)	14 (8.1)
Cavitation on x-ray (%)								
Yes	51 (54.3)	255 (60.6)	66 (64.1)	95 (54.9)	82 (59.9)	255 (60.6)	35 (58.3)	95 (54.9)
No	36 (38.3)	137 (32.5)	34 (33.0)	59 (34.1)	47 (34.3)	137 (32.5)	23 (38.3)	59 (34.1)
Missing	7 (7.4)	29 (6.9)	3 (2.9)	19 (11.0)	8 (5.8)	29 (6.9)	2 (3.3)	19 (11.0)
Culture smear status (%)								
Positive	56 (59.6)	192 (45.6)	49 (47.6)	85 (49.1)	79 (57.7)	192 (45.6)	26 (43.3)	85 (49.1)
Negative	36 (38.3)	201 (47.7)	43 (41.7)	76 (43.9)	53 (38.7)	201 (47.7)	26 (43.3)	76 (43.9)
Missing	2 (2.1)	28 (6.7)	11 (10.7)	12 (6.9)	5 (3.6)	28 (6.7)	8 (13.3)	12 (6.9)

Supplemental Table 5.S2B. Baseline characteristics of patients in the per-protocol analyses by censor status.

	3rd Per-protocol analysis				4th Per-protocol analysis			
	Uncensored		Censored		Uncensored		Censored	
	Control group	Clofazimine group	Control group	Clofazimine group	Control group	Clofazimine group	Control group	Clofazimine group
n	94	421	103	173	137	421	60	173
Functional status (ECOG) (%)								
Fully active	45 (47.9)	206 (48.9)	37 (35.9)	50 (28.9)	64 (46.7)	206 (48.9)	18 (30.0)	50 (28.9)
Limited self care	1 (1.1)	12 (2.9)	4 (3.9)	3 (1.7)	2 (1.5)	12 (2.9)	3 (5.0)	3 (1.7)
Ambulatory	38 (40.4)	168 (39.9)	50 (48.5)	79 (45.7)	57 (41.6)	168 (39.9)	31 (51.7)	79 (45.7)
Completely disabled	1 (1.1)	5 (1.2)	2 (1.9)	3 (1.7)	2 (1.5)	5 (1.2)	1 (1.7)	3 (1.7)
Missing	9 (9.6)	30 (7.1)	10 (9.7)	38 (22.0)	12 (8.8)	30 (7.1)	7 (11.7)	38 (22.0)
Past TB treatment (%)								
Yes	78 (83.0)	397 (94.3)	89 (86.4)	156 (90.2)	115 (83.9)	397 (94.3)	52 (86.7)	156 (90.2)
No	15 (16.0)	17 (4.0)	13 (12.6)	13 (7.5)	20 (14.6)	17 (4.0)	8 (13.3)	13 (7.5)
Missing	1 (1.1)	7 (1.7)	1 (1.0)	4 (2.3)	2 (1.5)	7 (1.7)	0 (0)	4 (2.3)
Past treatment category (%)								
None	15 (16.0)	17 (4.0)	13 (12.6)	13 (7.5)	20 (14.6)	17 (4.0)	8 (13.3)	13 (7.5)
Prior treatment only with first line drugs	16 (17.0)	35 (8.3)	20 (19.4)	15 (8.7)	23 (16.8)	35 (8.3)	13 (21.7)	15 (8.7)
Prior treatment with second line drugs	62 (66.0)	362 (86.0)	69 (67.0)	141 (81.5)	92 (67.2)	362 (86.0)	39 (65.0)	141 (81.5)
Missing	1 (1.1)	7 (1.7)	1 (1.0)	4 (2.3)	2 (1.5)	7 (1.7)	0 (0)	4 (2.3)
Past first-line TB drug use (%)								
Yes	16 (17.0)	35 (8.3)	20 (19.4)	15 (8.7)	23 (16.8)	35 (8.3)	13 (21.7)	15 (8.7)
No	77 (81.9)	379 (90.0)	82 (79.6)	154 (89.0)	112 (81.8)	379 (90.0)	47 (78.3)	154 (89.0)
Missing	1 (1.1)	7 (1.7)	1 (1.0)	4 (2.3)	2 (1.5)	7 (1.7)	0 (0)	4 (2.3)
Past second-line TB drug use (%)								
Yes	62 (66.0)	362 (86.0)	69 (67.0)	141 (81.5)	92 (67.2)	362 (86.0)	39 (65.0)	141 (81.5)
No	31 (33.0)	52 (12.4)	33 (32.0)	28 (16.2)	43 (31.4)	52 (12.4)	21 (35.0)	28 (16.2)
Missing	1 (1.1)	7 (1.7)	1 (1.0)	4 (2.3)	2 (1.5)	7 (1.7)	0 (0)	4 (2.3)
MDR-TB resistance category (%)								
MDR-TB FQ &SLI sensitive	28 (29.8)	118 (28.0)	34 (33.0)	49 (28.3)	42 (30.7)	118 (28.0)	20 (33.3)	49 (28.3)
MDR-TB + SLI resistant & FQ sensitive	9 (9.6)	125 (29.7)	12 (11.7)	47 (27.2)	13 (9.5)	125 (29.7)	8 (13.3)	47 (27.2)
MDR-TB + FQ resistant & SLI sensitive	35 (37.2)	70 (16.6)	33 (32.0)	21 (12.1)	52 (38.0)	70 (16.6)	16 (26.7)	21 (12.1)
MDR-TB + SLI & FQ resistance	9 (9.6)	91 (21.6)	13 (12.6)	48 (27.7)	14 (10.2)	91 (21.6)	8 (13.3)	48 (27.7)
Missing	13 (13.8)	17 (4.0)	11 (10.7)	8 (4.6)	16 (11.7)	17 (4.0)	8 (13.3)	8 (4.6)
DST for bedaquiline Susceptible (%)	2 (2.1)	7 (1.7)	0 (0.0)	9 (5.3)	2 (1.5)	7 (1.7)	0 (0)	9 (5.3)
DST for linezolid (%)								
Not tested	92 (97.9)	407 (97.4)	102 (99)	164 (95)	134 (97.8)	407 (97.4)	9 (100)	24 (96.0)
Resistant	1 (1.1)	0 (0.0)	0	0	1 (0.7)	0 (0)	0	0
Susceptible	1 (1.1)	11 (2.6)	1 (1.0)	9 (5)	2 (1.5)	11 (2.6)	0 (0)	9 (5.3)

Supplemental Table 5.S2B. Baseline characteristics of patients in the per-protocol analyses by censor status.

	3rd Per-protocol analysis				4th Per-protocol analysis			
	Uncensored		Censored		Uncensored		Censored	
	Control group	Clofazimine group	Control group	Clofazimine group	Control group	Clofazimine group	Control group	Clofazimine group
n	**94**	**421**	**103**	**173**	**137**	**421**	**60**	**173**
DST for clofazimine (%)								
Not tested	92 (97.9)	411 (98.3)	103 (100)	163 (95.3)	135 (98.5)	411 (98.3)	60 (100)	163 (95.3)
Resistant	2 (2.1)	1 (0.2)	0 (0.0)	2 (1.2)	2 (1.5)	1 (0.2)	0 (0)	2 (1.2)
Susceptible	0 (0.0)	6 (1.4)	0 (0.0)	6 (3.5)	0 (0)	6 (1.4)	0 (0)	6 (3.5)
DST for levofloxacin (%)								
Not tested	44 (46.8)	156 (37.3)	61 (59.2)	81 (47.4)	73 (53.3)	156 (37.3)	32 (53.3)	81 (47.4)
Resistant	20 (21.3)	88 (21.1)	18 (17.5)	37 (21.6)	25 (18.2)	88 (21.1)	13 (21.7)	37 (21.6)
Susceptible	30 (31.9)	174 (41.6)	24 (23.3)	53 (31.0)	39 (28.5)	174 (41.6)	15 (25.0)	53 (31.0)
DST for moxifloxacin (%)								
Not tested	36 (38.3)	199 (47.6)	55 (53.4)	90 (52.6)	60 (43.8)	199 (47.6)	31 (51.7)	90 (52.6)
Resistant	18 (19.1)	61 (14.6)	16 (15.5)	24 (14.0)	23 (16.8)	61 (14.6)	11 (18.3)	24 (14.0)
Susceptible	40 (42.6)	158 (37.8)	32 (31.1)	57 (33.3)	54 (39.4)	158 (37.8)	18 (30.0)	57 (33.3)
Likely effective for levofloxacin (%)								
Likely effective	35 (37.2)	116 (27.6)	38 (36.9)	61 (35.3)	44 (32.1)	116 (27.6)	29 (48.3)	61 (35.3)
Not effective	15 (16.0)	62 (14.7)	22 (21.4)	48 (27.7)	23 (16.8)	62 (14.7)	14 (23.3)	48 (27.7)
Missing	44 (46.8)	243 (57.7)	43 (41.7)	64 (37.0)	70 (51.1)	243 (57.7)	17 (28.3)	64 (37.0)
Likely effective for moxifloxacin (%)								
Likely effective	30 (31.9)	165 (39.2)	23 (22.3)	45 (26.0)	47 (34.3)	165 (39.2)	6 (10.0)	45 (26.0)
Not effective	15 (16.0)	78 (18.5)	20 (19.4)	20 (11.6)	24 (17.5)	78 (18.5)	11 (18.3)	20 (11.6)
Missing	49 (52.1)	178 (42.3)	60 (58.3)	108 (62.4)	66 (48.2)	178 (42.3)	43 (71.7)	108 (62.4)
DST for fluoroquinolone (with likely effectiveness) (%)†								
Not tested	13 (13.8)	65 (15.4)	22 (21.4)	30 (17.3)	25 (18.2)	65 (15.4)	10 (16.7)	30 (17.3)
Resistant	22 (23.4)	91 (21.6)	20 (19.4)	41 (23.7)	28 (20.4)	91 (21.6)	14 (23.3)	41 (23.7)
Susceptible	59 (62.8)	265 (62.9)	61 (59.2)	102 (59.0)	84 (61.3)	265 (62.9)	36 (60.0)	102 (59.0)
Number of effective drugs used at baseline (median [IQR])	5 [4, 6]	5 [4, 5]	4 [3, 5]	4 [4, 5]	5 [4, 5]	5 [4, 5]	4 [3, 5]	4 [4, 5]
Number of drugs used at baseline (median [IQR])	6 [6, 7]	6 [6, 7]	6 [6, 7]	6 [6, 7]	6 [6, 7]	6 [6, 7]	6 [6, 7]	6 [6, 7]
Treatment duration in months (mean (SD))	17.9 (7.1)	18.6 (6.1)	19.6 (4.5)	18.5 (5.4)	18.8 (6.4)	18.6 (6.1)	18.8 (4.8)	18.5 (5.4)
Country treated (%)								
Armenia	4 (4.3)	2 (0.5)	1 (1.0)	3 (1.7)	4 (2.9)	2 (0.5)	1 (1.7)	3 (1.7)
Bangladesh	10 (10.6)	44 (10.5)	13 (12.6)	6 (3.5)	19 (13.9)	44 (10.5)	4 (6.7)	6 (3.5)
Belarus	0 (0.0)	23 (5.5)	0 (0.0)	5 (2.9)	0 (0)	23 (5.5)	0 (0)	5 (2.9)
Ethiopia	2 (2.1)	14 (3.3)	2 (1.9)	3 (1.7)	3 (2.2)	14 (3.3)	1 (1.7)	3 (1.7)
Georgia	5 (5.3)	12 (2.9)	10 (9.7)	29 (16.8)	8 (5.8)	12 (2.9)	7 (11.7)	29 (16.8)
Haiti	0 (0.0)	4 (1.0)	1 (1.0)	0 (0.0)	0 (0)	4 (1.0)	1 (1.7)	0 (0)
Indonesia	4 (4.3)	4 (1.0)	4 (3.9)	2 (1.2)	4 (2.9)	4 (1.0)	4 (6.7)	2 (1.2)

Supplemental Table 5.S2B. Baseline characteristics of patients in the per-protocol analyses by censor status.

	3rd Per-protocol analysis				4th Per-protocol analysis			
	Uncensored		Censored		Uncensored		Censored	
	Control group	Clofazimine group	Control group	Clofazimine group	Control group	Clofazimine group	Control group	Clofazimine group
n	**94**	**421**	**103**	**173**	**137**	**421**	**60**	**173**
Country (continued)								
Kazakhstan	34 (36.2)	164 (39.0)	46 (44.7)	68 (39.3)	53 (38.7)	164 (39.0)	27 (45.0)	68 (39.3)
Kenya	0 (0.0)	1 (0.2)	0	0	0 (0)	1 (0.2)		
Kyrgyzstan	0 (0.0)	6 (1.4)	1 (1.0)	1 (0.6)	0 (0)	6 (1.4)	1 (1.7)	1 (0.6)
Lesotho	0 (0.0)	9 (2.1)	6 (5.8)	15 (8.7)	1 (0.7)	9 (2.1)	5 (8.3)	15 (8.7)
Myanmar	0 (0.0)	2 (0.5)	0 (0.0)	2 (1.2)	0 (0)	2 (0.5)	0 (0)	2 (1.2)
Pakistan	28 (29.8)	55 (13.1)	15 (14.6)	7 (4.0)	38 (27.7)	55 (13.1)	5 (8.3)	7 (4.0)
Peru	4 (4.3)	75 (17.8)	0 (0.0)	15 (8.7)	4 (2.9)	75 (17.8)	0 (0)	15 (8.7)
South Africa	0 (0.0)	4 (1.0)	3 (2.9)	17 (9.8)	0 (0)	4 (1.0)	3 (5.0)	17 (9.8)
Vietnam	3 (3.2)	2 (0.5)	1 (1.0)	0 (0.0)	3 (2.2)	2 (0.5)	1 (1.7)	0 (0)

† Levofloxacin and moxifloxacin DST were combined into one variable; if moxifloxacin DST was missing it was replaced with levofloxacin DST, and if both fluoroquinolones DST were missing, they were considered likely effective if there was no history of past treatment use with either of these drugs. If the DST was missing for both FQ, and information regarding prior use of both of these FQ was also missing, then the missing DST for this FQ variable were imputed.

Supplemental Table 5.S3. Stratified analyses for effect measure modification by drug susceptibility results for fluoroquinolones† in the intention-to-treat analysis comparing treatment success between the clofazimine group and control group.

	Susceptible to fluoroquinolones on DST	Resistant to fluoroquinolones on DST
	ATE (95% CI)	**ATE (95% CI)**
Intention-to-treat		
Crude/unadjusted	0.10 (0.00, 0.20)	-0.02 (-0.12, 0.09)
Fully adjusted	0.10 (-0.01, 0.20)	-0.02 (-0.14, 0.10)
IPTW baseline	0.13 (0.01, 0.24)	-0.03 (-0.15, 0.10)

IPTW: Inverse probability of treatment weight; ATE: average treatment effect; DST: drug susceptibility testing; CI: confidence interval † Levofloxacin and moxifloxacin DST were combined into one variable; if moxifloxacin DST was missing it was replaced with levofloxacin DST, and if both fluoroquinolones DST were missing, they were considered likely effective if there was no history of past treatment use with either of these drugs. If the DST was missing for both FQ, and information regarding prior use of both of these FQ was also missing, then the missing DST for this FQ variable were imputed. Note: analyses were stratified using the imputed variable for likely effective fluoroquinolone resistance and as such n per strata differ between data sets, and pooled results are presented.

Supplemental Table 5.S4. Baseline characteristics of the patients in the control group who started clofazimine at any time during treatment and those who did not.

	Did not start clofazimine	Started clofazimine	
	n = 112	n = 85	p value
Sex = Female (%)	52 (46.4)	31 (36.5)	0.209
Age (mean (SD))	35.4 (12)	38.3 (13)	0.105
Body mass index (mean (SD))	19.0 (4)	20.0 (4)	0.079
Diabetes (%)			0.559
Yes	14 (12.5)	14 (16.5)	
No	98 (87.5)	71 (83.5)	
Renal insufficiency (%)			0.82
Yes	2 (1.8)	1 (1.2)	
No	103 (92.0)	77 (90.6)	
Missing	7 (6.2)	7 (8.2)	
HIV infection status (%)			1
Positive	6 (5.4)	5 (5.9)	
Negative	106 (94.6)	80 (94.1)	
Alcohol use ever (%)			0.03
Yes	15 (13.4)	6 (7.1)	
No	96 (85.7)	73 (85.9)	
Missing	1 (0.9)	6 (7.1)	
Smoking ever (%)			0.411
Yes	25 (22.3)	20 (23.5)	
No	86 (76.8)	62 (72.9)	
Missing	1 (0.9)	3 (3.5)	
Intravenous Drug use ever (%)			0.168
Yes	1 (0.9)	0 (0.0)	
No	110 (98.2)	81 (95.3)	
Missing	1 (0.9)	4 (4.7)	
Bilateral disease on x-ray (%)			0.292
Yes	73 (65.2)	56 (65.9)	
No	31 (27.7)	27 (31.8)	
Missing	8 (7.1)	2 (2.4)	
Cavitation on x-ray (%)			0.096
Yes	60 (53.6)	57 (67.1)	
No	44 (39.3)	26 (30.6)	
Missing	8 (7.1)	2 (2.4)	
Culture smear status (%)			0.924
Positive	61 (54.5)	44 (51.8)	
Negative	44 (39.3)	35 (41.2)	
Missing	7 (6.2)	6 (7.1)	

Supplemental Table 5.S4. Baseline characteristics of the patients in the control group who started clofazimine at any time during treatment and those who did not.

	Did not start clofazimine	Started clofazimine	
	n = 112	n = 85	p value
Functional status (ECOG) (%)			0.676
Fully active	50 (44.6)	32 (37.6)	
Limited self care	2 (1.8)	3 (3.5)	
Ambulatory	46 (41.1)	42 (49.4)	
Completely disabled	2 (1.8)	1 (1.2)	
Missing	12 (10.7)	7 (8.2)	
Past TB treatment (%)			0.683
Yes	93 (83.0)	74 (87.1)	
No	18 (16.1)	10 (11.8)	
Missing	1 (0.9)	1 (1.2)	
Past treatment category (%)			0.821
None	18 (16.1)	10 (11.8)	
Prior treatment only with first line drugs	19 (17.0)	17 (20.0)	
Prior treatment with second line drugs	74 (66.1)	57 (67.1)	
Missing	1 (0.9)	1 (1.2)	
Past first-line TB drug use (%)			0.84
Yes	19 (17.0)	17 (20.0)	
No	92 (82.1)	67 (78.8)	
Missing	1 (0.9)	1 (1.2)	
Past second-line TB drug use (%)			0.966
Yes	74 (66.1)	57 (67.1)	
No	37 (33.0)	27 (31.8)	
Missing	1 (0.9)	1 (1.2)	
MDR-TB resistance category (%)			0.855
MDR-TB FQ &SLI sensitive	34 (30.4)	28 (32.9)	
MDR-TB + SLI resistant & FQ sensitive	11 (9.8)	10 (11.8)	
MDR-TB + FQ resistant & SLI sensitive	41 (36.6)	27 (31.8)	
MDR-TB + SLI & FQ resistance	11 (9.8)	11 (12.9)	
Missing	15 (13.4)	9 (10.6)	
DST for bedaquiline = Susceptible (%)	2 (1.8)	0 (0.0)	0.602
DST for linezolid (%)			0.671
Not tested	110 (98.2)	84 (98.8)	
Resistant	1 (0.9)	0 (0.0)	
Susceptible	1 (0.9)	1 (1.2)	
DST for clofazimine = Resistant (%)	2 (1.8)	0 (0.0)	0.602
DST for levofloxacin (%)			0.074
Not tested	52 (46.4)	53 (62.4)	
Resistant	26 (23.2)	12 (14.1)	
Susceptible	34 (30.4)	20 (23.5)	

Supplemental Table 5.S4. Baseline characteristics of the patients in the control group who started clofazimine at any time during treatment and those who did not.

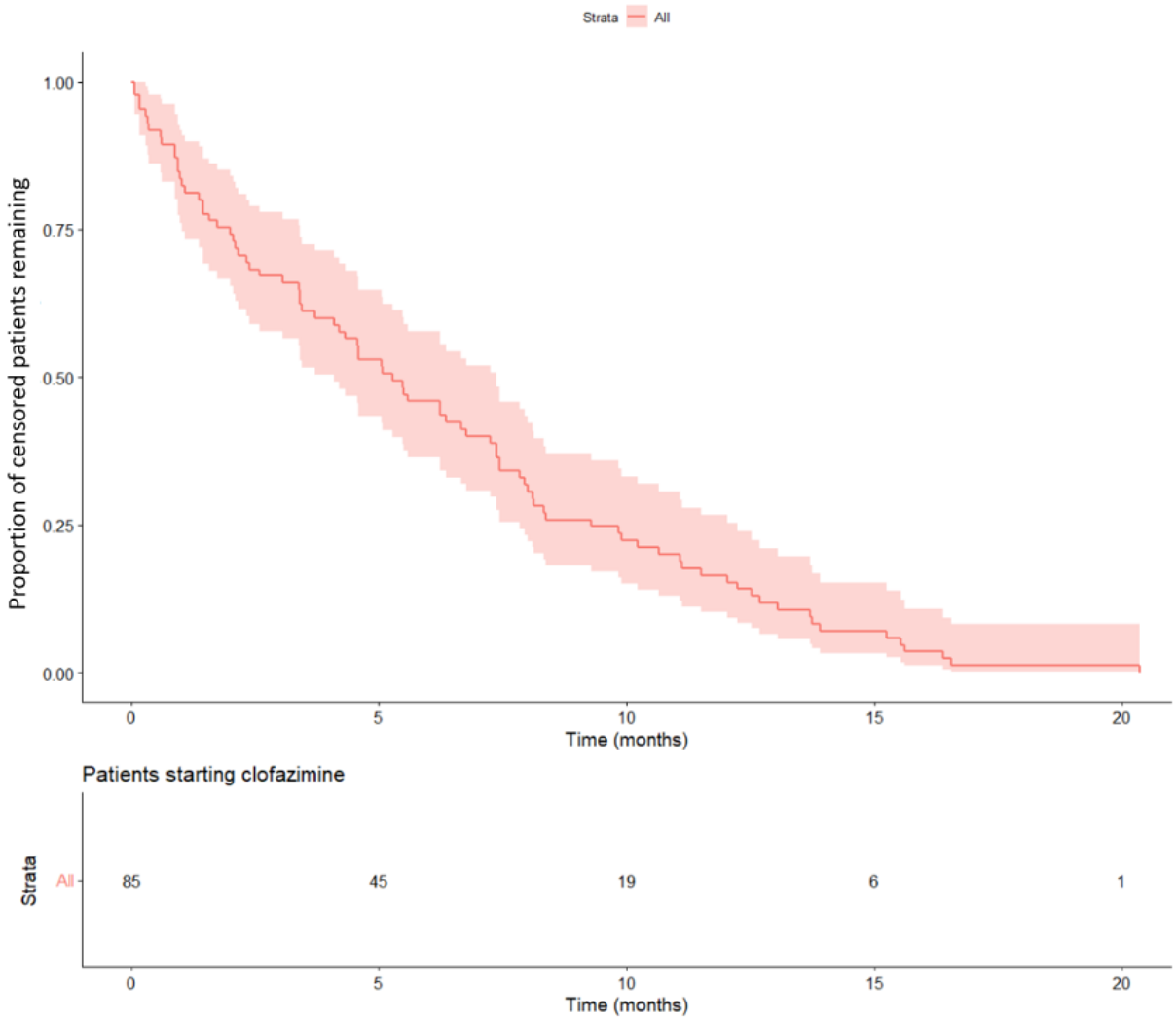
	Did not start clofazimine	Started clofazimine	
	n = 112	n = 85	p value
DST for moxifloxacin (%)			0.019
Not tested	42 (37.5)	49 (57.6)	
Resistant	23 (20.5)	11 (12.9)	
Susceptible	47 (42.0)	25 (29.4)	
Likely effective for levofloxacin (%)			0.893
Likely effective	43 (38.4)	30 (35.3)	
Not effective	21 (18.8)	16 (18.8)	
Missing	48 (42.9)	39 (45.9)	
Likely effective for moxifloxacin (%)			0.772
Likely effective	31 (27.7)	22 (25.9)	
Not effective	18 (16.1)	17 (20.0)	
Missing	63 (56.2)	46 (54.1)	
DST for fluoroquinolone (with likely effective) (%)			0.181
Not tested	16 (14.3)	19 (22.4)	
Resistant	28 (25.0)	14 (16.5)	
Susceptible	68 (60.7)	52 (61.2)	
Number of effective drugs used at baseline (median [IQR])	5 [4, 6]	4 [3, 5]	0.023†
Number of drugs used at baseline (median [IQR])	6 [6, 7]	6 [6, 7]	0.107†
Treatment duration in months (mean (SD))	17.9 (7)	20.0 (4)	0.011
Country treated (%)			0.163
Armenia	4 (3.6)	1 (1.2)	
Bangladesh	11 (9.8)	12 (14.1)	
Ethiopia	3 (2.7)	1 (1.2)	
Georgia	6 (5.4)	9 (10.6)	
Haiti	1 (0.9)	0 (0.0)	
Indonesia	4 (3.6)	4 (4.7)	
Kazakhstan	42 (37.5)	38 (44.7)	
Kyrgyzstan	0 (0.0)	1 (1.2)	
Lesotho	2 (1.8)	4 (4.7)	
Pakistan	31 (27.7)	12 (14.1)	
Peru	4 (3.6)	0 (0.0)	
South Africa	1 (0.9)	2 (2.4)	
Vietnam	3 (2.7)	1 (1.2)	
Treatment success (%)	80 (71.4)	64 (75.3)	0.547

All p-values are for Chi-Square tests except stated otherwise. †Kruskal-Wallis rank sum test

Supplemental Table 5.S5. Adverse events leading to permanent drug stoppage for patients in the control group who started clofazimine at any time during treatment and those who did not.

	Did not start clofazimine n = 112	Started clofazimine n = 85	Overall n = 197	
	n (%)	n (%)	n (%)	p value
Drugs in the core regimens				
Linezolid	4 (3.6)	14 (16.5)	18 (9.1)	0.004
Levofloxacin	1 (1.5)	0 (0.0)	1 (0.9)	1
Moxifloxacin	1 (1.8)	3 (6.1)	4 (3.8)	0.53
Clofazimine	0	2 (2.4)	2 (2.4)	NE
Drugs received in addition to the core regimens				
Ethambutol	1 (4.8)	1 (8.3)	2 (6.1)	1
Pyrazinamide	10 (11.2)	23 (35.4)	33 (21.4)	0.001
Kanamycin	1 (33.3)	1 (20.0)	2 (25.0)	1
Capreomycin	2 (6.9)	6 (24.0)	8 (14.8)	0.168
Amikacin	4 (10.5)	7 (28.0)	11 (17.5)	0.148
Prothionamide	9 (36.0)	7 (33.3)	16 (34.8)	1
Ethionamide	1 (2.3)	9 (33.3)	10 (14.1)	0.001
Cycloserine	5 (5.4)	13 (19.7)	18 (11.4)	0.011
Para-aminosalicylic acid	3 (8.8)	10 (27.8)	13 (18.6)	0.084
Delamanid	0 (0.0)	1 (7.1)	1 (3.4)	0.972
Number of drugs stopped for AEs per patient				
One drug stopped	26 (23.2)	31 (36.5)	57 (28.9)	
More than one drug stopped	7 (6.2)	28 (32.9)	35 (17.8)	
No drugs stopped	79 (70.5)	26 (30.6)	105 (53.3)	<0.001

ITT: Intention-to-treat; AE: adverse event, NE: not estimable. All p-values are for Chi-Square tests. No patients stopped bedaquiline, terizidone, imipenem-cilastatin, or amoxicillin-clavulanic acid for an adverse event.



Supplemental Figure 5.S1. Timing of clofazimine start for patients in the control group who received clofazimine after baseline treatment initiation.

REFERENCES:

1. World Health Organization. Global tuberculosis report 2022. Geneva, Switzerland: World Health Organization, 2022.
2. Dheda K, Gumbo T, Maartens G, et al. The Lancet Respiratory Medicine Commission: 2019 update: epidemiology, pathogenesis, transmission, diagnosis, and management of multidrug-resistant and incurable tuberculosis. *Lancet Respir Med* 2019; **7**(9): 820-6.
3. Van Deun A, Maug AK, Salim MA, et al. Short, highly effective, and inexpensive standardized treatment of multidrug-resistant tuberculosis. *Am J Respir Crit Care Med* 2010; **182**(5): 684-92.
4. Nunn AJ, Phillips PPJ, Meredith SK, et al. A Trial of a Shorter Regimen for Rifampin-Resistant Tuberculosis. *N Engl J Med* 2019; **380**(13): 1201-13.
5. Mok J, Lee M, Kim DK, et al. 9 months of delamanid, linezolid, levofloxacin, and pyrazinamide versus conventional therapy for treatment of fluoroquinolone-sensitive multidrug-resistant tuberculosis (MDR-END): a multicentre, randomised, open-label phase 2/3 non-inferiority trial in South Korea. *Lancet* 2022; **400**(10362): 1522-30.
6. Goodall RL, Meredith SK, Nunn AJ, et al. Evaluation of two short standardised regimens for the treatment of rifampicin-resistant tuberculosis (STREAM stage 2): an open-label, multicentre, randomised, non-inferiority trial. *Lancet* 2022; **400**(10366): 1858-68.
7. Conradie F, Diacon AH, Ngubane N, et al. Treatment of Highly Drug-Resistant Pulmonary Tuberculosis. *N Engl J Med* 2020; **382**(10): 893-902.
8. Conradie F, Bagdasaryan TR, Borisov S, et al. Bedaquiline-Pretomanid-Linezolid Regimens for Drug-Resistant Tuberculosis. *N Engl J Med* 2022; **387**(9): 810-23.
9. World Health Organization. WHO operational handbook on tuberculosis. Module 4: treatment - drug-resistant tuberculosis treatment. Geneva: World Health Organization, 2022.
10. Lan Z, Ahmad N, Baghaei P, et al. Drug-associated adverse events in the treatment of multidrug-resistant tuberculosis: an individual patient data meta-analysis. *Lancet Respir Med* 2020; **8**(4): 383-94.
11. Ahmad Khan F, Salim MAH, du Cros P, et al. Effectiveness and safety of standardised shorter regimens for multidrug-resistant tuberculosis: individual patient data and aggregate data meta-analyses. *Eur Respir J* 2017; **50**(1).
12. Ahuja SD, Ashkin D, Avendano M, et al. Multidrug resistant pulmonary tuberculosis treatment regimens and patient outcomes: an individual patient data meta-analysis of 9,153 patients. *PLoS medicine* 2012; **9**(8): e1001300.
13. Bisson GP, Bastos M, Campbell JR, et al. Mortality in adults with multidrug-resistant tuberculosis and HIV by antiretroviral therapy and tuberculosis drug use: an individual patient data meta-analysis. *Lancet* 2020; **396**(10248): 402-11.
14. Collaborative Group for the Meta-Analysis of Individual Patient Data in MDRTBt, Ahmad N, Ahuja SD, et al. Treatment correlates of successful outcomes in pulmonary multidrug-resistant tuberculosis: an individual patient data meta-analysis. *Lancet* 2018; **392**(10150): 821-34.

15. Rodriguez CA, Sy KTL, Mitnick CD, Franke MF. Time-Dependent Confounding in Tuberculosis Treatment Outcome Analyses: A Review of a Source of Bias. *Am J Respir Crit Care Med* 2020; **202**(9): 1311-4.
16. Franke MF, Mitnick CD. Time for a change: considering regimen changes in analyses of observational drug-resistant TB treatment cohort data. *The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease* 2020; **24**(11): 1151-5.
17. Shinozaki T, Suzuki E. Understanding Marginal Structural Models for Time-Varying Exposures: Pitfalls and Tips. *J Epidemiol* 2020; **30**(9): 377-89.
18. Clare PJ, Dobbins TA, Mattick RP. Causal models adjusting for time-varying confounding-a systematic review of the literature. *Int J Epidemiol* 2019; **48**(1): 254-65.
19. Mansournia MA, Etminan M, Danaei G, Kaufman JS, Collins G. Handling time varying confounding in observational research. *Bmj* 2017; **359**: j4587.
20. Hernan MA, Wang W, Leaf DE. Target Trial Emulation: A Framework for Causal Inference From Observational Data. *JAMA* 2022; **328**(24): 2446-7.
21. Matthews AA, Danaei G, Islam N, Kurth T. Target trial emulation: applying principles of randomised trials to observational studies. *Bmj* 2022; **378**: e071108.
22. Trevisi L, Hernan MA, Mitnick CD, et al. Effectiveness of Bedaquiline Use Beyond Six Months in Patients with Multidrug-Resistant Tuberculosis. *Am J Respir Crit Care Med* 2023.
23. Khan U, Huerga H, Khan AJ, et al. The endTB observational study protocol: treatment of MDR-TB with bedaquiline or delamanid containing regimens. *BMC Infect Dis* 2019; **19**(1): 733.
24. World Health Organization. Definitions and reporting framework for tuberculosis—2013 revision. Geneva: World Health Organization, 2013.
25. Médecins Sans Frontières. Clinical guidelines - Diagnosis and treatment manual, 2023.
26. van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software* 2011; **45**(3): 1 - 67.
27. Zubrod CG, Schneiderman M, Frei E, et al. Appraisal of methods for the study of chemotherapy of cancer in man: Comparative therapeutic trial of nitrogen mustard and triethylene thiophosphoramide. *Journal of Chronic Diseases* 1960; **11**(1): 7-33.
28. Team RC. (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, URL https://wwwR-projectorg/.
29. Kliiman K, Altraja A. Predictors of poor treatment outcome in multi- and extensively drug-resistant pulmonary TB. *Eur Respir J* 2009; **33**(5): 1085-94.
30. Nyang'wa BT, Berry C, Kazounis E, et al. A 24-Week, All-Oral Regimen for Rifampin-Resistant Tuberculosis. *N Engl J Med* 2022; **387**(25): 2331-43.
31. Stadler JAM, Maartens G, Meintjes G, Wasserman S. Clofazimine for the treatment of tuberculosis. *Front Pharmacol* 2023; **14**: 1100488.
32. World Health Organization. WHO consolidated guidelines on tuberculosis. Module 4: treatment - drug-resistant tuberculosis treatment. Geneva: World Health Organization, 2020.
33. Singh N, Singh PK, Singh U, Garg R, Jain A. Fluroquinolone drug resistance among MDR-TB patients increases the risk of unfavourable interim microbiological treatment outcome: An observational study. *J Glob Antimicrob Resist* 2021; **24**: 40-4.

34. Hernan MA, Robins JM. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. *Am J Epidemiol* 2016; **183**(8): 758-64.
35. Hernan MA, Sauer BC, Hernandez-Diaz S, Platt R, Shrier I. Specifying a target trial prevents immortal time bias and other self-inflicted injuries in observational analyses. *J Clin Epidemiol* 2016; **79**: 70-5.
36. Rudolph JE, Benkeser D, Kennedy EH, Schisterman EF, Naimi AI. Estimation of the Average Causal Effect in Longitudinal Data With Time-Varying Exposures: The Challenge of Nonpositivity and the Impact of Model Flexibility. *Am J Epidemiol* 2022; **191**(11): 1962-9.
37. Danaei G, Garcia Rodriguez LA, Cantero OF, Logan RW, Hernan MA. Electronic medical records can be used to emulate target trials of sustained treatment strategies. *J Clin Epidemiol* 2018; **96**: 12-22.

Chapter 6 – Identifying patients with multidrug-resistant tuberculosis who may benefit from shorter durations of treatment.

Section 6.1 Preface

This manuscript continues in the area of observational studies of MDR-TB but focuses on treatment duration. As mentioned in the literature review, several trials have been conducted to assess the efficacy of shorter treatments in patients with MDR-TB. However, the conditions of these trials and the populations included do not reflect treatment in programmatic settings. There is uncertainty whether patients with more advanced disease, extensive resistance, and other comorbidities could also benefit from these shorter regimens.

Studying treatment duration in MDR-TB patients is complicated by the highly individualized approach of treatment and because treatment durations are also determined by negative outcomes such as loss to follow-up, failure, and death. For those remaining on treatment, the regimens and duration are highly individualized and vary by provider and patient presentation, which is not measured and recorded in collected data.

In this manuscript I attempt to address these methodological challenges by including only patients with successful outcomes and using their deviation from their site-specific mean treatment duration (in patients with treatment success) as their outcome. With this analysis, I attempted to identify clinical and treatment factors that are associated with shorter treatment duration to identify MDR-TB patients who may benefit from shorter treatment.

This manuscript has been published in *PloS one* 2023; **18**(10): e0292106.[131]

Section 6.2 Manuscript 3

Identifying patients with multidrug-resistant tuberculosis who may benefit from shorter durations of treatment.

Nicholas Winters[1], Mireille Schnitzer[1,2,3], Jonathon R. Campbell[4,5,6], Susannah Ripley[1], Carla Winston[7], Rada Savic[8,9], Nafees Ahmad[10], Gregory Bisson[11], Keertan Dheda[12], Ali Esmail[12], Medea Gegia[13], Ignacio Monedero[14], Margareth Pretti Dalcolmo[15], Denise Rodrigues[16], Rupak Singla[17], Jae-Joon Yim[18], and Dick Menzies[1]

Affiliations:
1. Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Canada
2. Faculty of Pharmacy, Université de Montréal, Canada.
3. Department of Social and Preventive Medicine, Université de Montréal, Canada.
4. Department of Medicine & Department of Global and Public Health, Faculty of Medicine and Health Sciences, McGill University, Montreal, Canada
5. McGill International TB Centre, Montreal, Canada
6. Respiratory Epidemiology and Clinical Research Unit, Centre for Outcomes Research & Evaluation, Research Institute of the McGill University Health Centre, Montreal, Canada
7. US Centers for Disease Control and Prevention, Atlanta, Georgia
8. Department of Bioengineering and Therapeutic Sciences, University of California San Francisco Schools of Pharmacy and Medicine, San Francisco, California, USA.
9. UCSF Center for Tuberculosis, University of California San Francisco, San Francisco, California, USA.
10. Faculty of Pharmacy and Health Sciences, University of Baluchistan, Quetta, Pakistan
11. Department of Medicine and Department of Biostatistics, Epidemiology, and Informatics, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA, USA
12. Centre for Lung Infection and Immunity, Department of Medicine & UCT Lung Institute, University of Cape Town, Cape Town, South Africa
13. Global Tuberculosis Program, World Health Organization, Geneva, Switzerland
14. TB-HIV Department, International Union against Tuberculosis and Lung Diseases, Paris, France
15. Centro de Referência Helio Fraga, Fiocruz, Brazil
16. Instituto Clemente Ferreira, Sao Paulo, Brazil
17. National Institute of Tuberculosis & Respiratory Diseases, New Delhi, India
18. Department of Internal Medicine, Division of Pulmonary and Critical Care Medicine, Seoul National University College of Medicine, Seoul, South Korea

ABSTRACT

Objective: Studying treatment duration for multidrug-resistant tuberculosis (MDR-TB) using observational data is methodologically challenging. We aim to present a hypothesis generating approach to identify factors associated with shorter duration of treatment.

Study Design and Setting: We conducted an individual patient data meta-analysis among MDR-TB patients restricted to only those with successful treatment outcomes. Using multivariable linear regression, we estimated associations and their 95% confidence intervals (CI) between the outcome of individual deviation in treatment duration (in months) from the mean duration of their treatment site and patient characteristics, drug resistance, and treatments used.

Results: Overall, 6702 patients from 84 treatment sites were included. We found that factors commonly associated with poor treatment outcomes were also associated with longer treatment durations, relative to the site mean duration. Use of bedaquiline was associated with a 0.51 (95% CI: 0.15, 0.87) month decrease in duration of treatment, which was consistent across subgroups, while MDR-TB with fluoroquinolone resistance was associated with 0.78 (95% CI: 0.36, 1.21) months increase.

Conclusion: We describe a method to assess associations between clinical factors and treatment duration in observational studies of MDR-TB patients, that may help identify patients who can benefit from shorter treatment.

Funding: None.

INTRODUCTION

Multidrug-resistant tuberculosis (MDR-TB), defined as tuberculosis with resistance to both rifampicin and isoniazid, is a major global health burden.[1] Although treatment success has increased over time to 60-70%,[1,2] the estimated number of MDR-TB cases has increased from previous years to 450,000 in 2021.[1] Current recommended treatment for advanced and extensive MDR-TB is as long as 18-20 months[3] and entails a high patient burden. There is no doubt that shorter regimens are attractive for patients, health systems, and providers, as they reduce the burden of treatment.[4-7] In the past 10 years, several studies[8-12] have investigated shorter regimens for treatment of MDR-TB in randomized controlled trials (RCTs), but these may not reflect treatment in programmatic settings.

Assessing the effect of MDR-TB treatment duration in non-randomized studies has several potential limitations. Individuals' treatment durations are determined by the outcomes of loss to follow-up, failure, and death. For those remaining on treatment, the regimens and duration are highly individualized and vary by provider and patient presentation, which entail methodological challenges. Despite these challenges, investigators have used individual duration as an outcome,[13] but inferences were limited, and the evidence is considered by the WHO to be of very low quality.[3]

Based on previous analyses using individual patient data (IPD)[13,14] treatment duration varies widely between treatment sites and each site typically has a 'usual' duration of treatment targeted for patients, which may be based on local guidelines, experience, patient population, and availability of anti-tuberculosis drugs. However, there is substantial individual variation around that usual duration at each site. We hypothesize that analyzing

individual differences from the site-specific mean treatment duration, among patients with successful treatment outcomes, may help address these methodologic challenges.

Our aim was to describe associations with site-specific average treatment durations and to use deviations from these site-specific average treatment durations to identify clinical and treatment factors associated with shorter duration of treatment among individual rifampicin-resistant and MDR-TB patients with successful treatment outcomes.

METHODS

We conducted this study using a dataset of the 2019 IPD in MDR-TB described in detail previously.[14] Briefly, the database included data from studies conducted between January 1, 2009, and April 15, 2016 that were identified in a systematic review.[15] In addition, the IPD were updated with data contributed by authors of a 2010 IPD meta-analysis[16] and data from two public calls by the WHO in 2018[17] and 2019.[18] For comprehensive details on search strategy, study eligibility, and quality assessment see Supplement 6.1.

The 2019 IPD in MDR-TB contains records from 55 studies and 13,272 patients who initiated treatment between 1993 and 2019 in 38 countries and regions. The characteristics of studies included in the IPD have been described previously[19,20] and the quality and completeness of all studies in the IPD are described in Supplement 6.2.

Study population

We included studies reporting individual treatment duration and excluded studies which did not provide information on duration, or only provided planned durations. From the included studies, we included only patients that had successful (cured or completed) treatment outcomes, as defined elsewhere,[21,22] and who had their individual treatment

duration recorded. In those with death, failure, or loss to follow-up, their treatment duration is determined by their outcome, which may bias associations between characteristics and treatment duration, and were thus excluded. Any patients for which their individual treatment duration was missing were excluded from our primary analyses.

Outcomes

We assessed two outcomes among patients with successful treatment outcomes: i) the mean treatment duration at each treatment site, which was used in an ecological level analysis to explore potential associations with site-level factors; and ii) the difference between each individual's treatment duration and the mean treatment duration of all patients with treatment success at their site. The latter is our primary outcome in this analysis, which is the individual deviation from the site-specific mean treatment duration; this is referred to as *deviation in treatment duration* throughout the text and interpreted in terms of shorter (negative value) or longer (positive value) duration of treatment in months.

Statistical analysis

Ecological analysis of mean treatment duration of site

We first conducted an ecological analysis of the site-specific mean treatment duration in patients with successful outcomes where the unit of analysis was the treatment site, rather than the individual patient. Using available (non-imputed) data, we computed site-level proportions of categorical variables and means of continuous variables and described all using mean and standard deviation (SD), median and interquartile range [IQR], and range (minimum to maximum). We then performed univariable and multivariable linear

regression in imputed data (described below) to examine associations between site-level characteristics and the mean treatment duration of the site (see Supplement 6.3 for details).

Analysis of individual deviation from mean treatment duration of site

In our primary analysis, our approach was to construct an exploratory, hypothesis generating, multivariable model to identify factors conditionally associated with a change in deviation in treatment duration, while controlling for all variables selected into the model.

For clinical characteristics, drug susceptibility testing results, and treatments used we described categorical variables as n (%) while continuous variables were described using mean and standard deviation (SD) or median and interquartile range (IQR) using the available data (for detail on all variable specifications see Supplement 6.4). We also presented the regression coefficients (in months) and their 95% CI for age- and sex-adjusted univariable associations between deviation in treatment duration and each variable listed previously.

All regression analyses were conducted using data imputed with multivariate imputation by chained equations (MICE) with the assumption that data were missing at random (see Supplement 6.5 for detail). The deviation in treatment duration was imputed for those with either only planned or missing deviation in treatment duration for our sensitivity analyses, along with the other variables, however we only included subjects with non-missing duration in our primary analysis. Twenty data sets were generated with 25 Gibb's sampling iterations.[23]

To construct our exploratory model, we included variables known to be associated with treatment success in the published literature.[13,14,19,20] Additionally, we ran adaptive Lasso

regression,[24] using each imputed data set, on the previously listed characteristics to identify other potentially important predictors of treatment duration that were not a priori identified. Pearson coefficients were used to assess correlation between variables to be included. When highly correlated variables were present, we chose the more clinically relevant variable. We then used multivariable linear mixed-effects models with a random intercept for study to estimate regression coefficients (95% CI) for each selected covariate, controlling for the others.

In subgroup analyses, we assessed the final model stratified by subpopulations of patients: i) with MDR-TB plus resistance to both FQ and SLI and all others with MDR-TB (including resistance to FQ or SLI but not both); ii) with or without extensive disease (defined as yes if acid-fast bacilli (AFB) smear positive at baseline, and if AFB smear status was missing then the presence of radiographic findings of cavitation or bilateral disease); and iii) with or without previous tuberculosis treatment. We also did additional exploratory analyses in subgroups of those with: i) extensive disease with only MDR-TB and those without extensive disease with MDR-TB plus any additional resistance; and ii) those with past tuberculosis treatment with MDR-TB only and those without any past treatment with MDR-TB plus any additional resistance. Additionally, we explored the possible effect of selection bias on our population by analyzing our final model adjusted with inverse probability of selection weights for inclusion into the study population (see Supplement 6.6 for detail). We also performed an analysis that included subjects with missing treatment durations whose durations were imputed in the MICE procedure. Finally, we explored the impact that unmeasured confounding may have on the largest associations estimated from

our primary analysis by calculating E-values as described by VanderWeele et al.[25] (see Supplement 6.7). All analyses were conducted using R version 4.1.2.[26]

RESULTS

Of the 13,272 patients from 55 studies in the entire IPD, we included 6,702 from 49 studies that included 84 treatment sites in 34 countries (Figure 6.1). We excluded 6,570 patients in total. Six entire studies were excluded (2,235 patients) as they provided only planned duration or did not provide duration data (excluded and included studies were similar, see Supplement 6.2). Of the included studies, 4,335 patients were excluded: 44 had success but no duration data and 4,291 did not have treatment success. The characteristics of patients excluded from our analysis are presented in Supplemental Table 6.S1.

Ecological analysis of mean treatment duration of site

Descriptions of the site-level characteristics are presented in Table 6.1. The mean treatment duration of all sites was 22.8 and ranged from 12 to 36 months (see Supplemental Table 6.S2 for mean treatment duration of each site).

In univariable analysis, the proportion of patients at the site with past first-line drug use, MDR-TB plus resistance to both FQ and SLI (MDR-FQ+SLI), or resistance to pyrazinamide were associated with longer mean treatment duration at the site. However, in multivariable analysis, only the proportion of patients with MDR-FQ+SLI was associated with longer mean treatment duration of site (Table 6.1).

Analysis of individual deviation from mean treatment duration of site

The patients included in this analysis are described in Table 6.2A and 6.2B. The average total treatment duration was 22.0 months with SD of 4.6 (median 22 [IQR: 19, 24]). In univariable analyses, lower body mass index, past first- and second-line drug use, cavitation or bilateral disease on X-ray, and AFB smear positivity were all associated with longer treatment duration. Resistance to each drug, if tested (except linezolid, which was rarely tested) was associated with longer treatment duration. Longer treatment duration was associated with MDR-TB plus resistance to SLI but FQ sensitive (MDR-SLI), or MDR-TB plus resistance to FQ but SLI sensitive (MDR-FQ), or MDR-FQ+SLI. Use of capreomycin, kanamycin, moxifloxacin, levofloxacin, PAS, linezolid, clofazimine, Amx-Clv, clarithromycin, or bedaquiline, as well as greater number of drugs, were all associated with longer treatment duration.

In the final multivariable model (see Supplement 6.8 for detail on variable selection due to correlation of variables), longer treatment duration was associated with presence of cavitation, AFB smear positivity, HIV infection, past first-line drug use, and MDR-TB with all types of additional resistance (Figure 6.2). Individual deviation from mean duration of site was also associated with several treatment factors. Use of bedaquiline was associated with shorter treatment duration by -0.51 (95% CI -0.87 to -0.15) months. Longer treatment duration was associated with use of clarithromycin (1.12 months; 95% CI 0.71, 1.53), and with greater number of drugs used, or use of moxifloxacin, kanamycin, capreomycin, or Amx-Clv.

Results were similar when using inverse probability weights for selection into our study population from the entire IPD (Supplemental Table 6.S3). However, in our sensitivity

analysis including patients whose treatment durations were imputed, results were substantially different (Supplemental Figure 6.S1).

E-values for the largest regression coefficients from our primary analysis are presented in Supplemental Table 6.S4. For bedaquiline, an unmeasured confounder would need to have a risk ratio associated with both use of bedaquiline and treatment duration of 1.50 to completely explain away the association we observed with bedaquiline. The largest E-value required of an unmeasured confounder to explain away our estimated associations was for use of clarithromycin, while the smallest was for cavitation.

Subgroup analyses:

In subgroup analyses (Table 6.3) the direction of associations between shorter treatment duration and use of bedaquiline remained consistent across all subgroups (except in those with MDR-FQ+SLI), and regardless of disease extent. Associations between bedaquiline and duration were similar between those with or without past treatment. Additionally, use of Amx-Clv and clarithromycin were consistently associated with longer treatment duration in all subgroups. Body mass index was not associated with treatment duration in any subgroup while HIV was associated with longer duration in those with extensive disease and past tuberculosis treatment.

In other exploratory analyses (Supplemental Table 6.S5) results were similar for bedaquiline, Amx-Clv, clarithromycin, and body mass index. However, HIV was not associated with treatment duration in any exploratory subgroup.

DISCUSSION

With this IPD meta-analysis of 6,702 MDR-TB patients with treatment success, we have applied a novel approach to identify patients who may benefit from shorter MDR-TB treatment. In ecological analysis of site-level factors, the only clinical or treatment characteristic associated with average treatment duration of a site was the proportion of MDR patients with added resistance to FQ and SLI. The lack of associations between mean treatment duration of site with many clinical factors (such as age, HIV infection, past treatment, or other patterns of drug resistance) may indicate that unmeasured factors like physician beliefs, site conventions, or access to medications are more important determinants of treatment duration. In contrast, several clinical and treatment factors were associated with individual treatment duration in our analysis, which have shown to be associated with treatment outcomes in several prior studies.[14,19,20,27] Hence, our novel approach of using individual deviation from the site-specific mean treatment duration may provide a better method to assess clinical and treatment characteristics association with treatment duration.

By accounting for the mean treatment duration of a site in the duration outcome and by restriction to patients with successful treatment outcomes we aimed to create an outcome variable that accounts for the site-level variation and outcome-dependent complexities inherent in studying duration for treatment of MDR-TB. The finding that factors predicting poor treatment outcomes such as MDR with additional resistance to FQ and/or SLI,[1,22] HIV infection,[19] or cavitation[11,28] were associated with longer treatment duration provides support for the use of this method. Our finding that treatment duration is shorter when bedaquiline was used, is supported by several studies that have established the efficacy of

bedaquiline.[14,22,27] Additionally, as this was observational data, we included patient populations in our analysis that were excluded from trials of shorter treatment, such as those with additional resistance to SLIs or FQs,[9,10] low body mass index,[11,28] low HIV CD4 cell counts,[9-11,28] "any comorbidity likely to compromise protocol assessments",[11,28] or extensive disease and past treatment (the last two groups are not eligible for the 9-month all-oral regimen in WHO guidelines[22]). Use of bedaquiline was associated with shorter treatment duration across the majority of subgroups, suggesting that inclusion of patients previously excluded from RCTs[9-11,28] or ineligible in guidelines[22] could be considered in future trials of shorter bedaquiline containing regimens. Additionally, our results indicate that certain patients with more complicated clinical profiles, such as MDR-TB patients without additional resistance who also have extensive disease but either no past treatment or no HIV, may benefit from shorter treatment.

Although our analysis indicated that use of bedaquiline was associated with shorter treatment duration, these results require cautious interpretation as our models were constructed for the purposes of hypothesis generation. The association of shorter duration with use of bedaquiline may reflect the preferred use of the drug in regimens with planned shorter durations. However, this was not observed with linezolid or FQs, which are also used in regimens with shorter planned durations. Some characteristics and drugs that were associated with longer duration may reflect clinical conventions. For instance, use of low efficacy drugs (e.g. clarithromycin, Amx-Clv, and injectables[14]) may reflect use of drugs in desperation for patients with more complicated disease with longer planned duration. Similar conventions apply to associations with cavitation.

Our study has limitations. Primarily, we conducted this analysis in a population treated between 1993 and 2019, and treatment practises, including drug prioritization and advancement in ARTs and their uptake, have changed substantially in the last five years.[22] We also did not have data on the number of cavities, only the presence or absence, and were unable to assess what effect this had on duration. Additionally, there are site-level differences in treatment protocols that may affect treatment outcomes (availability of drugs), that may not be captured in variables we included in our models. However, as we used the average duration of treatment at a site in our duration outcome, we believe this may account for site-level heterogeneity in clinical practice. There is still potential for indication bias affecting duration of treatment in patients with complex profiles which may not be accounted for with our outcome. Although we conducted a 'new user' subgroup analysis[29] in those without previous treatment, this only addresses one aspect that may create indication bias for treatment duration. Because our population included only those with treatment success, our findings may not be generalizable to all patients with MDR-TB. Finally, our results do not reflect causal relationships and should be interpreted with caution, with additional consideration that variance of regression estimates may be underestimated due to the statistical selection of variables.[30]

Despite that, our study has several strengths. We included a large population of patients who had detailed information on important clinical characteristics and treatment. Additionally, we conducted several subgroup analyses of important patient groups that were previously excluded from trials on shorter treatment.[9-11,28] We also used E-values to assess unmeasured confounding (where a larger E-values implies a more robust observed estimate). Although plausible that an unmeasured confounder could account for the

observed association with bedaquiline, we feel it is unlikely such an important predictor would not have been included in our data. Although RCTs can provide clearer evidence on optimal duration, these are expensive, time consuming, often lack generalizability, and can test only a limited number of durations and/or regimens at once. Our use of observational data from a large population of MDR-TB patients from 84 treatment sites and 34 countries provided evidence that should be more generalizable. We also describe an analysis of characteristics associated with duration at the level of the treatment site. We interpret that the lack of associations between patient or treatment characteristics and our outcome may reflect the impact that provider belief and site convention have on duration, a problem which has not been previously described.

Our results produced correlates of individual treatment duration in MDR-TB patients that may help identify patients who would benefit from shorter treatment. We found evidence that certain patients with more extensive disease and drug resistance may benefit from shorter treatment and could be included in future treatment shortening trials.

Figure 6.1. PRISMA diagram for studies and patients included and excluded from the study population.

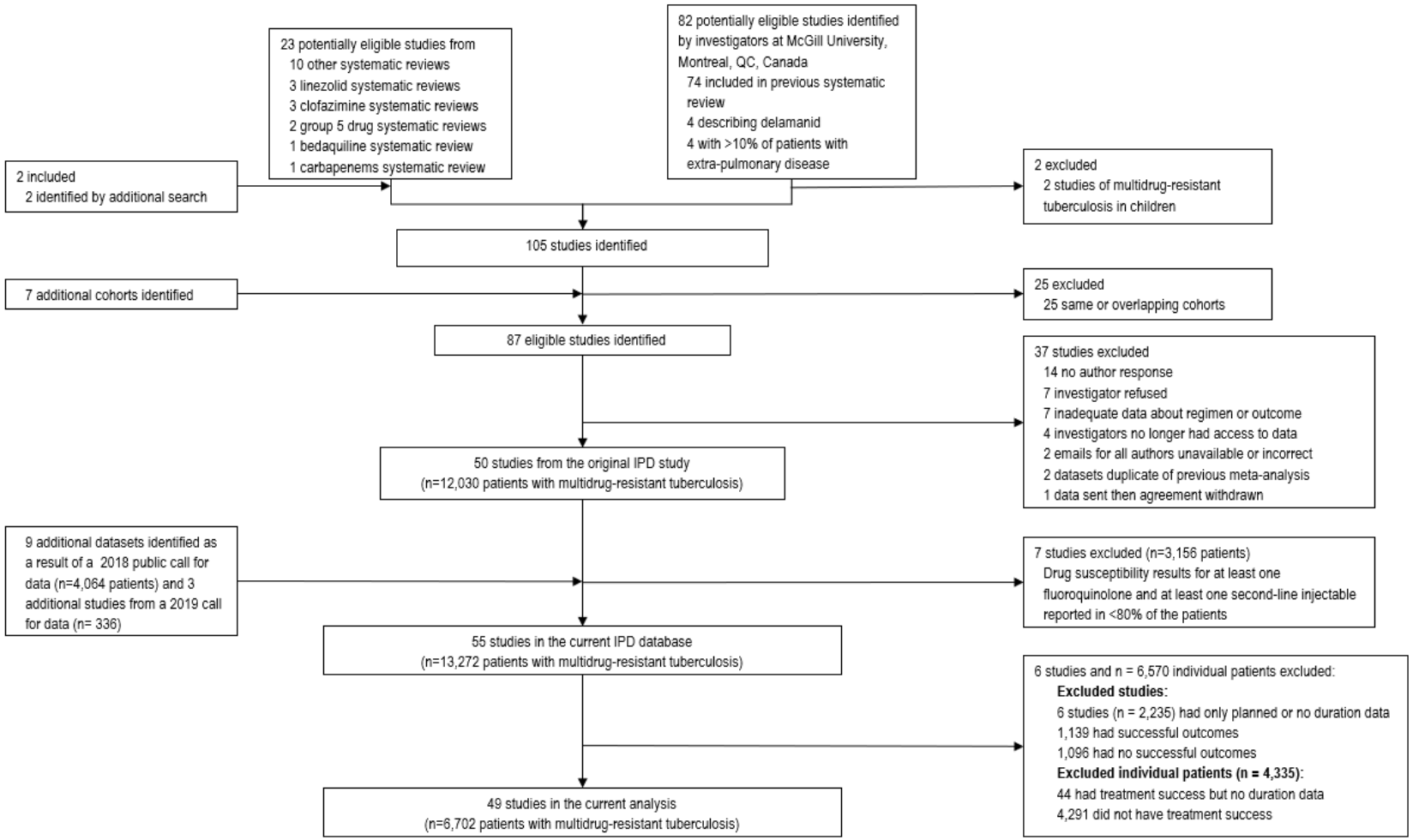
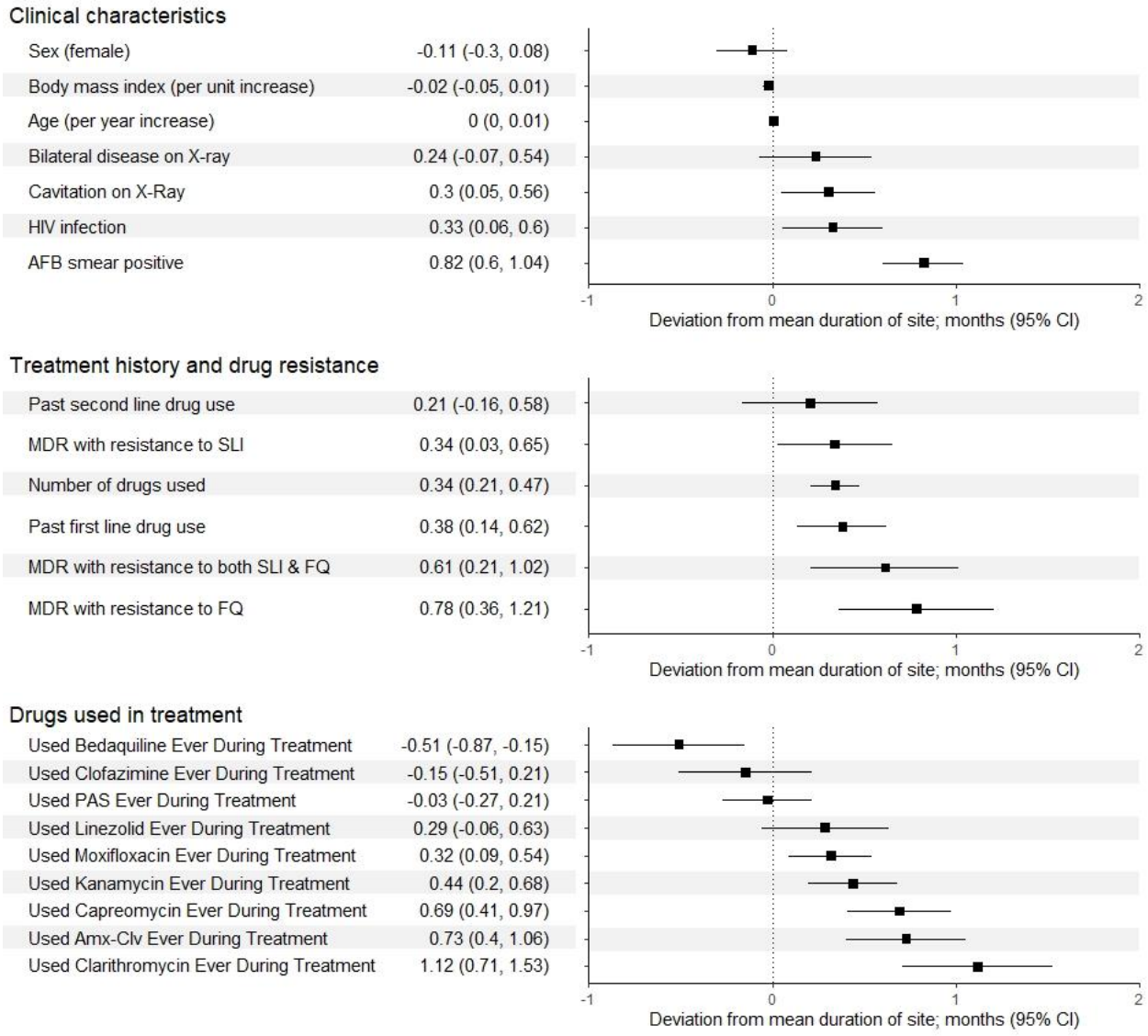


Figure 6.2. Forest plot of associations between deviation in treatment duration (in months) from site mean and patient characteristics, resistance categories. Estimates and 95% confidence intervals (CI) from a multivariable linear mixed model including all variables shown.



* Conditional R^2 for model: 0.08

Table 6.1. Site-level characteristics and their univariable and multivariable associations with the site-specific mean treatment duration in patients with successful treatment outcomes. Estimates and 95% confidence interval (CI) from linear regression models.

Variable: proportion at site unless stated otherwise (n=84)	Mean (SD)	Median [IQR]	Range	Mean site-specific treatment duration	
				Univariable Months (95% CI)	Multivariable Months (95% CI)
Clinical characteristics					
Age (mean years)	37.5 (6.1)	37.6 [33.4, 41.5]	21 to 55	0.01 (-0.1, 0.2)	0.1 (-0.1, 0.3)
Sex (Female)	0.37 (0.2)	0.38 [0.27, 0.48]	0 to 1	-3.5 (-7.8, 0.9)	-2.7 (-7.9, 2.5)
HIV infection	0.09 (0.2)	0 [0, 0.09]	0 to 0.73	3.0 (-1.9, 7.9)	3.0 (-2.9,.8.8)
2018 World Bank income category					
Low/lower-middle income - n (%)	13 (15.5)	NE	NE	Ref	Ref
Upper-middle income - n (%)	31 (36.9)	NE	NE	-1.4 (-4.1, 1.2)	-1.7 (-4.7, 1.2)
High income - n (%)	40 (47.6)	NE	NE	-2.1 (-4.7, 0.4)	-1.5 (-4.9, 1.9)
Extensive disease*	0.72 (0.24)	0.74 [0.52, 0.95]	0 to 1	1.4 (-2.3, 5.2)	1.0 (-3.4, 5.4)
Treatment history and drug resistance					
Past first-line TB drugs	0.70 (0.29)	0.79 [0.47, 0.98]	0 to 1	3.3 (0.2, 6.5)	1.3 (-4.1, 6.7)
Past second-line TB drugs	0.25 (0.32)	0.08 [0, 0.5]	0 to 1	2.3 (-0.6, 5.1)	-0.2 (-5.4, 4.9)
Number of effective drugs used	4.41 (0.78)	4.27 [4, 4.9]	2.4 to 7.08	-0.8 (-2.0, 0.3)	-0.2 (-1.6, 1.2)
MDR-TB + FQ &SLI sensitive	0.51 (0.36)	0.41 [0.18, 0.9]	0 to 1	3.0 (0.5, 5.5)	
MDR-TB + FQ resistant & SLI sensitive	0.14 (0.20)	0.05 [0, 0.19]	0 to 1	1.2 (-3.2, 5.6)	0.1 (-5.2, 5.5)
MDR-TB + SLI resistant & FQ sensitive	0.13 (0.17)	0.08 [0, 0.18]	0 to 1	-2.4 (-7.4, 2.7)	-0.2 (-6.3, 6.0)
MDR-TB + SLI & FQ resistance	0.24 (0.31)	0.06 [0, 0.42]	0 to 1	4.1 (1.4, 6.9)	4.6 (0.2, 9.0)
MDR-TB + Pyrazinamide resistance	0.45 (0.30)	0.44 [0.21, 0.65]	0 to 1	4.3 (0.7, 7.9)	
Drugs used in treatment					
Patients received Bedaquiline	0.32 (0.44)	0 [0, 1]	0 to 1	1.5 (-0.6, 3.5)	
Patients received Linezolid	0.37 (0.40)	0.23 [0, 0.73]	0 to 1	0.03 (-2.2, 2.2)	
Bedaquiline used at site (%)	41 (48.8)	NE	NE	0.17 (-1.6, 2.0)	-0.3 (-2.6, 2.0)
Linezolid used at site (%)	58 (69.0)	NE	NE	-0.59 (-2.5, 1.3)	-0.3 (-2.8, 2.2)
Patients with success	79.8 (236.7)	34.5 [11, 76.5]	1 to 2128	-0.03 (-0.12, 0.07)	
Patients treated	131.3 (401.5)	52.5 [15.8, 123.3]	1 to 3626	-0.01 (-0.06, 0.05)	0.0 (-0.1, 0.1)
Treatment duration (Months)	22.8 (4.1)	22.6 [20.3, 24.6]	12 to 36	NE	NE

Note: Extensive disease is defined as: AFB smear positive at baseline. If AFB smear information missing, then if radiographic findings of cavitation or bilateral disease. If value blank in multivariable coefficient column, then the variable was not included in the multivariable model. NE: not estimated; TB: tuberculosis; FQ: fluoroquinolones; SLI: second-line injectable. Note: proportion of patients receiving bedaquiline/linezolid, MDR-TB FQ & SLI sensitive, MDR-TB plus pyrazinamide resistance, and number of patients with success were not included in the multivariable model as they were highly correlated with other relevant variables that were included. For the multivariable model, R^2: 0.24; adjusted R^2: 0.05

Table 6.2A. Description of patient characteristics and their association (adjusted for age and sex) with deviation in treatment duration from site mean.

		Total duration of treatment	Individual Deviation in treatment duration from centre mean	
	n (%) unless specified otherwise	mean (SD) Months	mean (SD) Months	Univariable regression estimate months (95% CI)*
All patients	**n = 6702**	22.0 (4.6)	0.0 (4)	
Clinical characteristics				
Sex = Male$^{\Delta}$	3982 (59.4)	22.1 (4.6)	0.1 (4)	Ref
Female	2719 (40.6)	21.9 (4.6)	-0.1 (4)	-0.18 (-0.37, 0.02)
Age (mean (SD))	37.02 (13)	NE	NE	0.01 (-0.003, 0.01)§
Body mass index (mean (SD))	20.47 (3.84)	NE	NE	-0.04 (-0.07, -0.01)§
Body mass index category				
Normal	2024 (30.2)	22.4 (4.8)	0 (4.2)	Ref
Underweight	1028 (15.3)	22.7 (4.3)	0.4 (3.7)	-0.21 (-0.62, 0.19)
Overweight/Obese	377 (5.6)	22.3 (4.6)	-0.2 (3.9)	0.26 (0.00, 0.52)
Missing	3273 (48.8)	21.5 (4.5)	-0.1 (3.9)	Not estimated
2018 World Bank income category				
Low/Low-middle	1226 (18.3)	22.5 (4.2)	0 (3.5)	Ref
Upper-Middle	3555 (53.0)	22.3 (4)	0 (3.6)	0.01 (-0.26, 0.26)
High	1921 (28.7)	21.3 (5.6)	0 (4.9)	-0.02 (-0.30, 0.27)
Smoking				
Ex-smoker or never smoker	1834 (27.4)	22.5 (5.3)	-0.1 (4.8)	Ref
Current smoker	939 (14.0)	22.5 (5)	0.4 (4.1)	0.17 (-0.09, 0.42)
Unknown	3929 (58.6)	21.7 (4.1)	-0.1 (3.5)	Not estimated
HIV				
Negative	4771 (71.2)	22 (4.8)	0 (4.1)	Ref
Positive	1859 (27.7)	22.1 (3.9)	0.1 (3.5)	0.13 (-0.08, 0.35)
Unknown	72 (1.1)	22.9 (5.4)	-0.1 (5.1)	Not estimated
If HIV positive, on ART	1686 (90.7)	22 (3.8)	0 (3.5)	-0.16 (-0.83, 0.50)
Not on ART	173 (9.3)	23.2 (4.4)	0.5 (4.1)	
Diabetes				
No	3311 (49.4)	22.4 (5)	0 (4.3)	Ref
Yes	466 (7.0)	21.9 (4.4)	0.3 (3.7)	0.21 (-0.18, 0.59)
Unknown	2925 (43.6)	21.7 (4)	0 (3.6)	Not estimated
Cavitation on X-ray				
No	1606 (24.0)	21.7 (4.9)	-0.4 (4.2)	Ref
Yes	2308 (34.4)	22.5 (5.1)	0.3 (4.3)	0.60 (0.35, 0.86)
Unknown	2788 (41.6)	21.8 (3.8)	0 (3.6)	0.37 (0.12, 0.61)
Bilateral disease				
No	1122 (16.7)	21.4 (4.9)	-0.3 (4)	Ref
Yes	1999 (29.8)	22.2 (4.9)	0.2 (4.1)	0.52 (0.22, 0.81)
Unknown	3581 (53.4)	22.1 (4.3)	0 (3.9)	0.35 (0.09, 0.62)
AFB smear result				
Neg	1974 (29.5)	21.4 (4.7)	-0.6 (4)	Ref
Pos	4280 (63.9)	22.4 (4.5)	0.3 (4)	0.91 (0.70, 1.12)
Unknown	448 (6.7)	21.2 (4.2)	0 (3.8)	0.65 (0.24, 1.05)

Table 6.2A. Continued.

		Total treatment duration	Individual Deviation in treatment duration from centre mean	
	n (%) unless specified otherwise	mean (SD) Months	mean (SD) Months	Univariable regression estimate months (95% CI)*
Extensive disease				
No	2147 (32.0)	21.3 (4.6)	-0.6 (3.9)	
Yes	4512 (67.8)	22.4 (4.6)	0.3 (4)	0.90 (0.70, 1.11)
Unknown	43 (0.0)	20.9 (4.7)	-0.1 (4.5)	Not estimated
Treatment history and markers of disease severity				
Past TB treatment				
No	2336 (34.9)	21.2 (4.4)	-0.4 (3.8)	Ref
Yes	4271 (63.7)	22.5 (4.6)	0.2 (4.1)	0.56 (0.36, 0.77)
Unknown	95 (1.4)	21.7 (4.5)	0.1 (3.8)	0.48 (-0.34, 1.29)
Past first-line TB drug use				
No	2336 (34.9)	21.2 (4.4)	-0.4 (3.8)	Ref
Yes	4271 (63.7)	22.5 (4.6)	0.2 (4.1)	0.56 (0.36, 0.76)
Unknown	95 (1.4)	21.7 (4.5)	0.1 (3.8)	Not estimated
Past second-line TB drug used				
No	5048 (75.3)	21.7 (4.1)	-0.2 (3.6)	Ref
Yes	1226 (18.3)	23.3 (5.4)	0.6 (4.6)	0.71 (0.44, 0.98)
Unknown	428 (6.4)	22.4 (6.2)	0.5 (5.6)	Not estimated
Pre-treatment Drug susceptibility results				
DST Performed for FQ	6449 (96.2)			Not estimated
If DST Performed, FQ Resistant = Yes	1172 (18.2)	23.6 (5.8)	0.8 (5)	1.04 (0.77, 1.31)
If DST Performed, FQ Resistant = No	5277 (81.8)	21.6 (4.1)	-0.2 (3.6)	Ref
DST Performed for SLIs	6455 (96.3)			Not estimated
If DST Performed, SLI Resistant = Yes	1629 (25.2)	23 (5.3)	0.5 (4.5)	0.58 (0.36, 0.81)
If DST Performed, SLI Resistant = No	4826 (74.8)	21.7 (4.2)	-0.1 (3.8)	Ref
DST Performed for Linezolid	665 (9.9)			Not estimated
If DST Performed, Linezolid Resistant = Yes	16 (2.4)	21.5 (3.6)	-0.8 (2.8)	-0.76 (-2.74, 1.22)
If DST Performed, Linezolid Resistant = No	649 (97.6)	21.1 (4.4)	0 (3.7)	
DST Performed for Pyrazinamide	3490 (52.1)			Not estimated
If DST Performed, Pyrazinamide Resistant = Yes	1859 (53.3)	22.2 (5.4)	0.3 (4.7)	0.51 (0.30, 0.72)
If DST Performed, Pyrazinamide Resistant = No	1631 (46.7)	21.1 (4.6)	-0.4 (4)	Ref
DST Performed for Clofazimine	252 (3.8)			Not estimated
If DST Performed, Clofazimine Resistant = Yes	9 (3.6)	24.4 (5)	2 (5)	Not estimated †
If DST Performed, Clofazimine Resistant = No	243 (96.4)	21.8 (5.8)	0.1 (4.4)	
DST Performed for Cycloserine‡	2034 (30.3)			Not estimated
If DST Performed, Cycloserine Resistant = Yes	260 (12.8)	23.4 (6.2)	1 (5.2)	1.16 (0.65, 1.68)
If DST Performed, Cycloserine Resistant = No	1774 (87.2)	21.9 (5.3)	-0.1 (4.5)	

Table 6.2A. Continued.

		Total duration of treatment	Individual Deviation in treatment duration from centre mean	
	n (%) unless specified otherwise	mean (SD) Months	mean (SD) Months	Univariable regression estimate months (95% CI)*
MDR category				
MDR-TB FQ &SLI sensitive	4337 (64.7)	21.5 (4)	-0.3 (3.5)	Ref
MDR-TB + FQ resistant & SLI sensitive	929 (13.9)	22.2 (4.5)	0.2 (3.9)	0.48 (0.20, 0.77)
MDR-TB + SLI resistant & FQ sensitive	475 (7.1)	23.2 (5.8)	0.9 (5.1)	1.24 (0.85, 1.63)
MDR-TB + SLI & FQ resistance	688 (10.3)	23.9 (5.9)	0.8 (5)	1.08 (0.76, 1.41)
No DST	273 (4.1)	22.5 (5.7)	0 (5)	Not estimated
MDR-TB + SLI & FQ resistance vs. all others				
No	5741 (85.7)	21.8 (4.3)	-0.1 (3.8)	Ref
Yes	688 (10.3)	23.9 (5.9)	0.8 (5)	0.83 (0.52, 1.14)
Unknown	273 (4.1)	22.5 (5.7)	0 (5)	Not estimated

Table 6.2B. Description of drugs used in treatment and their association (adjusted for age and sex) with deviation in treatment duration from site mean.

		Total duration of treatment	Individual Deviation in treatment duration from centre mean	
All patients **n = 6702**	**n (%) unless specified otherwise**	**mean (SD) Months**	**mean (SD) Months**	**Univariable regression estimate months (95% CI)***
Drugs used in treatment				
Used Bedaquiline Ever During Treatment = Yes	1605 (23.9)	22.4 (3.9)	0.2 (3.5)	0.27 (0.04, 0.49)
No	5097 (76.1)	21.9 (4.8)	-0.1 (4.1)	Ref
Used Ofloxacin Ever During Treatment = Yes	1373 (20.5)	22 (4.2)	-0.1 (3.6)	-0.13 (-0.36, 0.11)
No	5329 (79.5)	22 (4.7)	0 (4.1)	Ref
Used Ciprofloxacin Ever During Treatment = Yes	266 (4.0)	23 (5.8)	0 (5.4)	0.03 (-0.46, 0.52)
No	6436 (96.0)	22 (4.5)	0 (3.9)	Ref
Used Moxifloxacin Ever During Treatment = Yes	3459 (51.6)	22.1 (4.6)	0.2 (4.1)	0.41 (0.22, 0.60)
No	3243 (48.4)	21.9 (4.6)	-0.2 (3.9)	Ref
Used Levofloxacin Ever During Treatment = Yes	1889 (28.2)	21.8 (4.9)	0.1 (4.2)	0.19 (-0.03, 0.40)
No	4813 (71.8)	22.1 (4.4)	-0.1 (3.9)	Ref
Used Linezolid Ever During Treatment = Yes	1594 (23.8)	22.5 (5.1)	0.5 (4.4)	0.63 (0.41, 0.86)
No	5108 (76.2)	21.9 (4.4)	-0.1 (3.8)	Ref
Used Clofazimine Ever During Treatment = Yes	1101 (16.4)	22.5 (4.8)	0.3 (3.8)	0.35 (0.10, 0.61)
No	5601 (83.6)	21.9 (4.5)	-0.1 (4)	Ref
Used Cycloserine/Terizidone Ever During Treatment = Yes	5702 (85.1)	22.1 (4.4)	0 (4.0)	0.19 (-0.08, 0.46)
No	1000 (14.9)	21.7 (5.3)	-0.2 (4.1)	Ref
Used Ethambutol Ever During Treatment = Yes	2895 (43.2)	22 (4.5)	0.1 (3.9)	0.10 (-0.09, 0.29)
No	3807 (56.8)	22.1 (4.7)	0 (4.1)	Ref
Used Pyrazinamide Ever During Treatment = Yes	5175 (77.2)	22 (4.3)	0 (3.8)	-0.13 (-0.36, 0.09)
No	1527 (22.8)	22 (5.5)	0.1 (4.6)	Ref
Used Streptomycin Ever During Treatment = Yes	692 (10.3)	22.5 (5.1)	0.1 (4.5)	0.10 (-0.21, 0.42)
No	6010 (89.7)	22 (4.5)	0 (3.9)	Ref
Used Rifabutin Ever During Treatment = Yes	154 (2.3)	22.8 (6.7)	0.1 (5.8)	0.05 (-0.59, 0.69)
No	6548 (97.7)	22 (4.5)	0 (3.9)	Ref
Used Amikacin Ever During Treatment = Yes	1048 (15.6)	21.8 (4.8)	0 (4.2)	0.06 (-0.21, 0.32)
No	5654 (84.4)	22.1 (4.5)	0 (3.9)	Ref
Used Capreomycin Ever During Treatment = Yes	1446 (21.6)	23.1 (5.6)	0.5 (4.7)	0.66 (0.42, 0.89)
No	5256 (78.4)	21.7 (4.2)	-0.1 (3.8)	Ref
Used Kanamycin Ever During Treatment = Yes	3151 (47.0)	21.9 (3.9)	0.1 (3.5)	0.20 (0.01, 0.39)
No	3551 (53.0)	22.1 (5.1)	-0.1 (4.3)	Ref
Used Ethionamide/Prothionamide Ever During Treatment = Yes	5096 (76.0)	22.1 (4.5)	0 (4)	0..20 (-0.03, 0.42)
No	1606 (24.0)	21.8 (4.9)	-0.2 (4.1)	Ref
Used PAS Ever During Treatment = Yes	2759 (41.2)	22.7 (5.2)	0.3 (4.4)	0.48 (0.29, 0.68)
No	3943 (58.8)	21.6 (4.1)	-0.2 (3.6)	Ref

Table 6.2B. Continued.

		Total duration of treatment	Individual Deviation in treatment duration from centre mean	
	n (%) unless specified otherwise	mean (SD) Months	mean (SD) Months	n (%) unless specified otherwise
Used Amx-Clv Ever During Treatment = Yes	994 (14.8)	24 (6.2)	1 (5.5)	1.27 (0.99, 1.55)
No	5708 (85.2)	21.7 (4.1)	-0.2 (3.6)	Ref
Used Thioacetazone Ever During Treatment = Yes	68 (1.0)	21 (5.4)	0.2 (4)	0.20 (-0.75, 1.16)
No	6634 (99.0)	22 (4.6)	0 (4)	Ref
Used Clarithromycin Ever During Treatment = Yes	485 (7.2)	24.6 (7)	1.4 (6.2)	1.50 (1.14, 1.87)
No	6217 (92.8)	21.8 (4.3)	-0.1 (3.7)	Ref
Used Imipenem Ever During Treatment = Yes	237 (3.5)	23.4 (4.6)	0.4 (4.1)	0.37 (-0.14, 0.89)
No	6465 (96.5)	22 (4.6)	0 (4)	Ref
Used Meropenem Ever During Treatment = Yes	61 (0.9)	21.1 (5.4)	-0.3 (4.4)	-0.32 (-1.33, 0.68)
No	6641 (99.1)	22 (4.6)	0 (4)	Ref
Used Delamanid Ever During Treatment = Yes	114 (1.7)	21 (4.4)	0 (3.6)	-0.02 (-0.76, 0.72)
No	6588 (98.3)	22 (4.6)	0 (4)	Ref
Number of drugs (median [IQR])	5 [4, 6]	NE	NE	0.37 (0.29, 0.46)§
Number of effective drugs (median [IQR])	4 [4, 5]	NE	NE	0.03 (-0.06, 0.12)§
Number of limited access drugs** (median [IQR])	0 [0, 1]	NE	NE	0.18 (0.09, 0.26)§
Total treatment duration (median [IQR])	22 [19, 24]	NE	NE	Not estimated
Deviation in treatment duration (median [IQR])	-0.15 [-2, 2]	NE	NE	Not estimated

SD: standard deviation; XDR: extensively drug resistant tuberculosis; MDR: multidrug resistant tuberculosis; TB: tuberculosis; AFB: acid-fast bacillus; Amx-Clv: Amoxicillin-Clavulanic Acid;

* Regression coefficients were estimated using imputed data and adjusted for age and sex. † Too few observations to estimate. ‡ Drug susceptibility testing for cycloserine and terizidone combined. § per unit increase.** includes bedaquiline, clofazimine, linezolid, imipenem, and meropenem. Δ One subject missing sex.

Table 6.3. Associations of individual deviation in treatment duration from site mean with patient characteristics, resistance categories, and drugs used, within specified subgroups. Estimates and 95% confidence interval (CI) from multivariable linear mixed models including all variables shown (unless otherwise specified).

Characteristic	Patients with additional SLI & FQ resistance months (95% CI)	Patients without FQ and SLI resistance* months (95% CI)	Patients with extensive disease months (95% CI)	Patients without extensive disease months (95% CI)	Patients with past TB treatment months (95% CI)	Patients without past TB treatment months (95% CI)
Clinical characteristics						
Age (per year increase)	0.02 (-0.01, 0.05)	0 (0, 0.01)	0 (-0.01, 0.01)	0.01 (0, 0.03)	0.01 (0, 0.02)	0 (-0.02, 0.01)
Sex (Female)	0.16 (-0.6, 0.92)	-0.14 (-0.34, 0.06)	-0.26 -0.5, -0.02)	0.12 (-0,21, 0.46)	-0.08 (-0.32, 0.17)	-0.17 (-0.48, 0.14)
Body mass index (per unit increase)	0.01 (-0.1, 0.11)	-0.03 (-0.06, 0.01)	-0.02 (-0.06, 0.02)	-0.03 (-0.09, 0.02)	-0.03 (-0.07, 0.01)	-0.01 (-0.07, 0.05)
HIV infection	0.15 (-1.01, 1.31)	0.34 (0.06, 0.62)	0.33 (0.01, 0.66)	0.33 (-0.1, 0.76)	0.38 (0.02, 0.74)	0.22 (-0.18, 0.62)
AFB smear positive	0.77 (-0.08, 1.62)	0.79 (0.56, 1.01)	Not estimated	Not estimated	0.82 (0.52, 1.12)	0.77 (0.44, 1.1)
Cavitation on X-Ray	1.08 (0.1, 2.05)	0.22 (-0.05, 0.49)	Not estimated	Not estimated	0.23 (-0.09, 0.54)	0.59 (0.12, 1.06)
Bilateral disease on X-ray	0.64 (-0.56, 1.85)	0.16 (-0.16, 0.48)	Not estimated	Not estimated	0.36 (-0.01, 0.74)	-0.04 (-0.58, 0.5)
Treatment history and drug resistance						
Past first-line drug use	-0.22 (-1.56, 1.11)	0.4 (0.16, 0.65)	0.43 (0.13, 0.74)	0.34 (-0.05, 0.73)	Not estimated	Not estimated
Past second-line drug use	0.72 (-0.41, 1.86)	0.15 (-0.24, 0.55)	0.13 (-0.32, 0.57)	0.4 (-0.16, 0.97)	Not estimated	Not estimated
Number of drugs used (per unit increase)	0.35 (-0.09, 0.79)	0.33 (0.19, 0.47)	0.4 (0.24, 0.56)	0.16 (-0.06, 0.38)	0.31 (0.14, 0.47)	0.29 (0.07, 0.51)
MDR-TB + FQ resistant & SLI sensitive	Not estimated	Not estimated	0.59 (0.21, 0.98)	-0.2 (-0.75, 0.34)	0.12 (-0.29, 0.54)	0.62 (0.15, 1.08)
MDR-TB + SLI resistant & FQ sensitive	Not estimated	Not estimated	0.92 (0.38, 1.45)	0.46 (-0.24, 1.15)	0.75 (0.26, 1.24)	0.97 (0.1, 1.83)
MDR-TB + SLI & FQ resistance	Not estimated	Not estimated	0.84 (0.32, 1.36)	0.15 (-0.52, 0.81)	0.55 (0.08, 1.02)	0.86 (0.04, 1.68)
Drugs used in treatment						
Used Bedaquiline Ever During Treatment	-0.89 (-2.19, 0.41)	-0.47 (-0.86, -0.09)	-0.51 (-0.94, -0.09)	-0.61 (-1.25, 0.03)	-0.43 (-0.89, 0.03)	-0.52 (-1.09, 0.05)
Used Moxifloxacin Ever During Treatment	0.63 (-0.2, 1.47)	0.26 (0.02, 0.5)	0.42 (0.14, 0.7)	0.07 (-0.31, 0.45)	0.18 (-0.1, 0.46)	0.7 (0.31, 1.09)
Used Linezolid Ever During Treatment	-0.82 (-2.03, 0.38)	0.54 (0.15, 0.93)	0.22 (-0.19, 0.64)	0.44 (-0.18, 1.06)	0.12 (-0.32, 0.56)	0.69 (0.12, 1.27)
Used Clofazimine Ever During Treatment	0.54 (-0.51, 1.58)	-0.43 (-0.84, -0.02)	-0.37 (-0.8, 0.07)	0.3 (-0.31, 0.91)	0.01 (-0.45, 0.47)	-0.25 (-0.85, 0.34)
Used Capreomycin During Treatment	0.46 (-0.44, 1.36)	0.8 (0.49, 1.11)	0.74 (0.4, 1.09)	0.56 (0.08, 1.04)	0.75 (0.39, 1.1)	0.73 (0.25, 1.22)
Used Kanamycin Ever During Treatment	0.64 (-0.58, 1.86)	0.47 (0.22, 0.71)	0.48 (0.18, 0.77)	0.45 (0.02, 0.88)	0.21 (-0.09, 0.52)	0.9 (0.48, 1.31)
Used Amx-Clv Ever During Treatment	0.7 (-0.18, 1.59)	0.77 (0.39, 1.15)	0.84 (0.43, 1.25)	0.51 (-0.03, 1.05)	0.93 (0.53, 1.33)	0.43 (-0.16, 1.03)
Used Clarithromycin During Treatment	0.64 (-0.46, 1.73)	1.19 (0.73, 1.66)	0.88 (0.38, 1.39)	1.6 (0.9, 2.3)	1.25 (0.77, 1.73)	0.72 (-0.1, 1.55)

*Includes MDR-TB FQ &SLI sensitive, MDR-TB + FQ resistant & SLI sensitive, MDR-TB + SLI resistant & FQ sensitive. Note: for MDR models were also adjusted for resistance to fluoroquinolone (FQ), second line injectables (SLI), pyrazinamide and cycloserine. For MDR-TB + SLI & FQ resistance, models were also adjusted for resistance to pyrazinamide and cycloserine (not shown for consistency with other subgroups). All models also adjusted for use of PAS. MDR: multidrug resistant tuberculosis; TB: tuberculosis; AFB: acid-fast bacillus; Amx-Clv: Amoxicillin-Clavulanic Acid;

Section 6.3 Supplemental material

Supplement 6.1. Eligibility criteria, search strategy, and quality assessment.

Eligibility criteria: Included studies were those reporting original results with end of treatment outcomes (i.e., success, failure or relapse, and death) for 25 or more adults (to avoid small series reporting unusual cases) with bacteriologically confirmed pulmonary multidrug resistant tuberculosis. Studies exclusively in children or of patients treated with short regimens were excluded as these were the topics of two concurrent individual patient data meta-analyses at time of original publication (*Lancet* 2018; **392**(10150): 821-34.)

Search Strategy:

Medline search (through Ovid)
MDR or XDR
1. exp multidrug resistant tuberculosis/ or exp extensively drug resistant tuberculosis/
2. (multidrug resistant tuberculosis or extensive* drug resistant tuberculosis or MDR-TB or XDR-TB).ti,ab,kw.
3. (tuberc* and (MDR or XDR or drug resistan* or multidrug resistan* or multi drug resistan* or poly drug resistan* or extensive*
drug resistan*)).ti,ab,kw.
Drugs
4. exp Fluoroquinolones/ or exp Quinolones/ or exp Levofloxacin/ or (fluoroquinolone* or quinolone* or levofloxacin or Levaquin
or moxifloxacin or Avelox).ti,ab,kw.
5. exp Kanamycin/ or exp Amikacin/ or exp Capreomycin/ or exp Aminoglycosides/ or (Kanamycin or Amikacin or Capreomycin or
(tuberc* and injectable*)).ti,ab,kw.
6. exp Pyrazinamide/ or exp Ethambutol/ or exp Cycloserine/ or exp Ethionamide/ or exp Prothionamide/ or (Pyrazinamide or
Ethambutol or para-aminosalicylic acid or Cycloserine or Ethionamide or Prothionamide).ti,ab,kw.
7. high dose.ti,ab,kw. and ((INH or isoniazid).ti,ab,kw. or exp isoniazid/)
Efficacy
8. exp Treatment Outcome/ or exp Prognosis/ or exp Death/ or exp Mortality/ or exp Treatment Failure/ or exp Survival/ or exp
Recurrence/ or exp Patient Dropouts/ or exp Patient Compliance/
9. (Treatment Outcome* or Prognosis or Death or Mortality or Treatment Failure or drug treatment failure or failure or Survival or

Recurrence or relapse or Patient Dropout* or dropout or non-compliance or compliance or efficacy or effective* or cure or
success* or default or adheren* or conversion* or microbiologic conversion or smear conversion or culture conversion or sputum
conversion).ti,ab,kw.
Toxicity
10. exp Treatment Outcome/ or exp Prognosis/ or exp Death/ or exp Mortality/ or exp Treatment Failure/ or exp Survival/ or exp
Recurrence/ or exp Toxicity Tests/ or exp Drug Tolerance/ or exp "Drug-Related Side Effects and Adverse Reactions"/
11. (Treatment outcome* or Prognosis or Death or Mortality or Treatment Failure or drug treatment failure or failure or Survival
or Recurrence or relapse or Toxicity Test* or toxicity or Drug Tolerance or toler* or intolerance or Side Effect* or Adverse Drug
Reaction* or adverse drug event* or adverse event* or adverse reaction* or safe* or drug safety).ti,ab,kw.
New drugs
12. (Bedaquiline or TMC-207 or delamanid or OPC-67683).ti,ab,kw.
Final steps
13. 1 or 2 or 3
14. 4 or 5 or 6 or 7
15. 8 or 9
16. 10 or 11
17. 13 and 14 and 15
18. 13 and 14 and 16
19. 12 and 13 and 15
20. 12 and 13 and 16
21. limit 17 to (humans and yr="2009 -Current")
22. limit 18 to (humans and yr="2009 -Current")
23. limit 19 to (humans and yr="2012 -Current")
24. limit 20 to (humans and yr="2012 -Current")
25. 21 or 22
26. 23 or 24
27. 25 or 26

(EmBase and the Cochrane Library were searched using the same strategy)

Quality assessment:

This is discussed in detail in Lancet 2018; 392: 821–34.

A checklist of seven indicators was developed (adapted from the Risk of Bias in Non-randomised Studies of Interventions (ROBINS-I) tool) to assess the quality of included studies. Two of these indicators were considered essential: 1) population selection using a census (all) or random selection approach; and 2) availability of drug susceptibility tests results to at least one fluoroquinolone and one second-line injectable (defined as any of amikacin, kanamycin, or capreomycin). For the remaining indicators, quality was judged to be adequate if the participation rate >80%, loss to follow-up <20%, treatment outcomes were defined according to published guidelines (Laserson and WHO 2013), and >90% of patient records had information about HIV infection, previous tuberculosis treatment, and age (as these are all important determinants of outcomes). Participation rates were based on the reported total number of eligible patients and the number enrolled, and we considered participation to be 100% if the investigators stated that all patients with multidrug-resistant tuberculosis were enrolled. Studies of high quality met both essential criteria and at least four of the other six. Studies of moderate quality met one of the two essential parameters and at least five in total. Remaining studies were considered of low quality.

Supplement 6.2. Completeness of information, quality of all studies in the IPD, and their inclusion status in this analysis.

No	Contact person (ref)	Sampling method	Info on SLI sensitivity	Info on FQ sensitivity	Participation rate	Lost to follow-up rate	Outcome definition	Info on age	Info on HIV	Info on TB treatment history	Quality	Included in this analysis
1	Ahmad[1]	Census	100.00%	100.00%	96.80%	1.70%	Laserson	100.00%	100.00%	100.00%	High	No
2	Ahuja[2]	Random	92.40%	92.40%	100.00%	19.00%	Laserson	100.00%	80.00%	100.00%	High	Yes
3	Anderson[3]	Census	100.00%	100.00%	100.00%	12.40%	Neither	100.00%	100.00%	90.50%	High	Yes
4	Bang[4]	Census	96.60%	93.10%	96.70%	17.20%	Laserson	100.00%	100.00%	100.00%	High	Yes
5	Barkane[5]	Census	100.00%	100.00%	100.00%	15.60%	Laserson	100.00%	100.00%	100.00%	High	Yes
6	Barry (Korea)[6,7]	RCT	100.00%	100.00%	92.70%	10.50%	Laserson	100.00%	100.00%	100.00%	High	Yes
7	Barry/Flood (Calif)[8]	Unclear	98.40%	95.20%	100.00%	4.80%	WHO2013	98.40%	100.00%	100.00%	Moderate	Yes
8	Bonnet[9]	Census	93.30%	93.30%	100.00%	41.30%	Laserson	100.00%	11.50%	98.60%	High	Yes
9	Brode[10]	Census	100.00%	100.00%	100.00%	0.00%	Laserson	100.00%	100.00%	100.00%	High	Yes
10	Brust[11]	Census	100.00%	100.00%	100.00%	24.10%	Laserson	99.30%	57.80%	98.50%	Moderate	No
11	Cegielski[12,13]	Census	92.80%	92.20%	60.10%	19.80%	Laserson	100.00%	68.30%	98.20%	High	Yes
12	Chan (Denver) [14]	Census	100.00%	100.00%	100.00%	26.70%	Laserson	100.00%	80.00%	100.00%	High	Yes
13	Dheda[15-17]	Census	100.00%	100.00%	61.50%	4.70%	Laserson	99.10%	100.00%	93.50%	High	Yes
14	Fox[18]	Census	93.10%	96.60%	100.00%	3.40%	WHO2013	100.00%	100.00%	100.00%	High	Yes
15	Gegia[19]	Census	100.00%	100.00%	100.00%	21.80%	Laserson	100.00%	72.90%	100.00%	High	No
16	Guglielmetti[20,21]	Census	100.00%	100.00%	100.00%	11.10%	WHO2013	100.00%	100.00%	100.00%	High	Yes
17	Guglielmetti[22]	Census	100.00%	100.00%	100.00%	10%	WHO2013	100.00%	100.00%	90.00%	High	Yes
18	Hughes[23]	Census	94.90%	94.90%	100.00%	25.40%	Laserson	100.00%	100.00%	100.00%	High	No
19	Isaakidis[23,24]	Census	96.70%	95.40%	100.00%	11.80%	Laserson	100.00%	100.00%	98.00%	High	Yes
20	Jarlsberg[25]	Census	96.40%	96.40%	100.00%	3.60%	Laserson	100.00%	92.90%	100.00%	High	Yes
21	Kempker[26]	Census	100.00%	100.00%	94.90%	32.70%	Laserson	100.00%	94.70%	100.00%	High	Yes
22	Koenig[27]	Census	96.30%	93.30%	100.00%	6.10%	Laserson	99.40%	100.00%	100.00%	High	Yes
23	Koh[28,29]	Census	100.00%	100.00%	100.00%	13.40%	WHO2013	100.00%	100.00%	100.00%	High	Yes
24	Kuksa[30]	Census	100.00%	100.00%	100.00%	15%	Laserson	100.00%	100.00%	100.00%	High	Yes
25	Kvasnovsky[31,32]	Census	100.00%	100.00%	100.00%	11.50%	Laserson	100.00%	96.90%	100.00%	High	Yes
26	Lange[33]	Census	94.00%	96.70%	100.00%	20.10%	Laserson	100.00%	99.50%	98.40%	High	Yes
27	Laniado-Laborin[34]	Census	100.00%	100.00%	100.00%	13.50%	Laserson	100.00%	100.00%	100.00%	High	Yes
28	Leung[35,36]	Census	100.00%	100.00%	100.00%	19.90%	Laserson	100.00%	100.00%	100.00%	High	Yes
29	Marks[37]	Random	92.30%	91.50%	100.00%	12.30%	Neither	100.00%	85.40%	100.00%	High	Yes
30	Migliori[38,39]	Census	96.60%	96.60%	Unclear	10.90%	WHO2013	100.00%	98.10%	99.30%	High	Yes
31	Migliori (BDQ)[40]	Census	97.00%	100.00%	Unclear	3.70%	WHO2013	100.00%	99.30%	100.00%	High	Yes
32	Milanov[41]	Census	94.00%	94.00%	100.00%	2.00%	Laserson	100.00%	100.00%	100.00%	High	Yes
33	Ndjeka[42]	Unclear	78.20%	81.20%	Unclear	21.10%	Laserson	100.00%	95.50%	0.00%	Low	Yes
34	Ndjeka[43]	Census	100.00%	100.00%	100.00%	18.50%	Both	100.00%	100.00%	100.00%	Low	Yes
35	O'Donnell[44]	Census	100.00%	100.00%	100.00%	13.20%	Laserson	100.00%	93.90%	93.90%	High	Yes
36	Palmero[45]	Census	100.00%	100.00%	100.00%	22.20%	WHO2013	100.00%	100.00%	100.00%	High	No
37	Podewils[46]	Census	91.00%	91.20%	100.00%	15.20%	Laserson	100.00%	55.60%	100.00%	High	Yes
38	Riekstina/Leimane[47]	Census	100.00%	100.00%	100.00%	14.70%	Laserson	100.00%	94.00%	100.00%	High	Yes
39	Rodrigues[48]	Census	87.00%	85.00%	100.00%	10.00%	Laserson	100.00%	98.00%	100.00%	High	Yes
40	Seo[49]	Census	100.00%	100.00%	100.00%	16.00%	Laserson	100.00%	100.00%	100.00%	High	Yes
41	Seung[50]	Census	80.20%	80.20%	100.00%	1.40%	Unclear	100.00%	0%	88.70%	High	No
42	Shim[29,51]	Census	100.00%	100.00%	86.40%	8.20%	WHO2013	100.00%	40.00%	100.00%	High	Yes
43	Singla[52]	Census	100.00%	100.00%	100.00%	13.80%	Laserson	100.00%	100.00%	100.00%	High	Yes
44	Skrahina[53]	Census	100.00%	100.00%	100.00%	1.00%	WHO2013	100.00%	99.00%	100.00%	High	Yes
45	Smith[54]	Census	100.00%	100.00%	100.00%	21.50%	Laserson	100.00%	100.00%	98.50%	High	Yes
46	TMC207-C208[55,56]	RCT	84.80%	84.80%	82.50%	28.80%	Laserson	100.00%	100.00%	100.00%	High	Yes
47	TMC207-C209[57]	Census	76.10%	76.10%	93.10%	15.20%	Laserson	100.00%	96.50%	100.00%	Moderate	Yes
48	Udwadia[58]	Census	100.00%	100.00%	100.00%	27.80%	Laserson	100.00%	44.40%	100.00%	High	Yes
49	van der Werf[59]	Census	100.00%	98.20%	100.00%	13.40%	Laserson	100.00%	92.00%	96.40%	High	Yes
50	Vasilyeva[60]	Census	94.40%	94.40%	100.00%	16.00%	WHO2013	100.00%	100.00%	100.00%	High	Yes
51	Viiklepp[61]	Census	100.00%	100.00%	100.00%	11.70%	Laserson	100.00%	99.70%	100.00%	High	Yes
52	Yim/Kwak[62]	Census	100.00%	100.00%	100.00%	4.90%	WHO2013	100.00%	100.00%	100.00%	High	Yes
53	Achar[63]	Census	66.10%	65.30%	100.00%	22.00%	WHO2013	100.00%	80.60%	100.00%	Moderate	Yes
54	Isaakidis[64]	Census	95.00%	96.00%	100.00%	6.00%	Laserson	100.00%	100.00%	100.00%	High	Yes
55	Skrahina[65]	Census	96.40%	95.50%	100.00%	0.00%	WHO2013	100.00%	99.10%	99.10%	High	Yes

SLI: Second-line injectable. FQ: fluoroquinolone. Both: indicates Laserson and WHO 2013 were used. Neither: indicates neither Laserson nor WHO 2013 were used. Quality Assessment Reference: Lancet 2018; 392: 821–34.

Supplement 6.3. Detail on models for ecological level analysis.

To estimate conditional regression coefficients and their 95% CI for associations with site-level treatment duration a multivariable linear regression model was constructed that included the site-level proportion of female sex, HIV infection, extensive disease (defined as yes if AFB smear positive at baseline or if AFB smear status was missing, presence of radiographic findings of cavitation or bilateral disease), past first-line drug use, past second-line drug use, MDR-TB, MDR-TB plus resistance to FQ but SLI sensitive (MDR-FQ), MDR-TB plus resistance to SLI but FQ sensitive (MDR-SLI), MDR-TB plus resistance to both FQ and SLI (MDR-FQ+SLI), resistance to pyrazinamide, and proportion of patients at that site who received bedaquiline or linezolid. We also included the site-level mean number of effective drugs used, mean patient age, number of patients treated, and 2018 World Bank income category of the site.

Supplement 6.4. Detail on descriptions and definitions of baseline patient characteristics, drug susceptibility testing, and treatments used.

Categorical baseline characteristics were described as n (%), and included the following: sex (male or female); body mass index category (underweight: <18.5 kg/m²; normal: ≥18.5 and <25 kg/m²; and overweight or obese: ≥25 kg/m); World Bank 2018 category of country level income (low/lower-middle, upper-middle, and high income); smoking (yes/no); alcohol use disorder (yes/no); HIV status (positive/negative); if HIV positive, on ART (yes/no); diabetes (yes/no); cavitation on chest radiography (yes/no); bilateral disease on chest radiography (yes/no); acid-fast bacilli (AFB) smear positivity at baseline (positive/negative); extensive disease (defined as yes if AFB smear positive at baseline or if AFB smear status was missing, presence of

radiographic findings of cavitation or bilateral disease); past TB treatment, past first-line drug use, and past second-line drug use (all yes/no); resistance on drug susceptibility testing (DST) for FQ, SLI, linezolid, pyrazinamide, clofazimine, and cycloserine/terizidone (all yes/no); drug resistant profile category defined as only MDR-TB, MDR-TB plus resistance to FQ but SLI sensitive (MDR-FQ), MDR-TB plus resistance to SLI but FQ sensitive (MDR-SLI), MDR-TB plus resistance to both FQ and SLI (MDR-FQ+SLI); and each individual drug used during treatment (yes/no). All missing observations were included as a missing category for all categorical variables for descriptive statistics. Age, body mass index, deviation in treatment duration, and total individual duration of treatment were described using mean (SD), while total number of drugs, effective drugs (based on DST results), and new/limited access drugs used (bedaquiline, clofazimine, linezolid, and/or meropenem/imipenem) were described using median [IQR].

Supplement 6.5. Detail on multivariate imputation by chained equations (MICE).

Missing data were imputed using age, sex, body mass index, previous treatment history, radiographic features, World Bank income level, drug susceptibility testing (DST) results for: fluoroquinolones (FQ), second line injectables (SLI), ethambutol, pyrazinamide, prothionamide/ethionamide, and para-aminosalicylic acid (PAS), as well as deviation in treatment duration. The deviation in treatment duration was imputed for those with either only planned or missing deviation in treatment duration for our sensitivity analyses, along with the other variables, however we only included subjects with non-missing duration in our primary analysis.

Supplement 6.6. Detail on inverse probability of selection weights.

We explored the possible effect of selection bias on our population by analyzing our final model adjusted with inverse probability of selection weights (IPSW) calculated using a logistic regression model with binary variable for inclusion (1: indicating inclusion; 0:indicating exclusion from study population) as the outcome and adjusted for age, sex, body mass index, World Bank income category, HIV status, diabetes, past first and second line drug use, radiographic findings, drug resistance to FQ, SLIs, ethambutol, PAS, amikacin/kanamycin/capreomycin/streptomycin, ethionamide/prothionamide and pyrazinamide, as well as number of drugs and use of bedaquiline, pyrazinamide, FQ, linezolid, clofazimine, and cycloserine/terizidone.

Supplement 6.7. Detailed description of E-Values.

With E-values, linear regression coefficients are converted to an approximation of the risk ratio from an approximation of the odds ratio as defined in Chinn[1] and VanderWeele[2], using the following formula: RR = $\sqrt{e^{(\frac{\beta}{SDo}*1.81)}}$ where β is the regression coefficient of an exposure and SD_o is the standard deviation of our treatment duration outcome. The E-values are calculated as follows: for RR>1: E-Value = RR + $\sqrt{RR*(RR-1)}$ with the confidence interval (CI) of the E-value being the lower limit of the RR (LL) + $\sqrt{LL*(LL-1)}$, if the LL>1, while if LL≤ then the CI = 1; for RR<1: E-value = 1/RR + $\sqrt{1/RR*([1/RR]-1)}$ with the CI of the E-value being 1/upper limit of the RR (UL) +$\sqrt{1/UL*([1/UL]-1)}$, if the UL<1, while if UL≥ 1 then CI = 1.

1. Chinn S. A simple method for converting an odds ratio to effect size for use in meta-analysis. *Stat Med* 2000; **19**(22): 3127-31.

2. VanderWeele TJ. On a Square-Root Transformation of the Odds Ratio for a Common Outcome. *Epidemiology* 2017; **28**(6): e58-e60.

Supplement 6.8. Detail on variable selection due to correlation in the multivariable model for the primary analysis.

Pearson coefficients were used to assess correlation between variables to be included. When highly correlated variables were present, we chose the more clinically relevant variable. This occurred between the following: individual DST results and MDR categories; number of drugs used, and number of effective drugs used; and extensive disease and individual measures of disease extent (AFB smear, cavitation/bilateral disease on x-ray). We included MDR categories as they capture the same information as DST results but with more clinically relevant categorization. Number of drugs was included as number of effective drugs was correlated with use of bedaquiline and clofazimine. Finally, the individual measures of disease extent were used as these provided a more granular description of the markers of extensive disease. The final covariate list included: age (continuous), sex, body mass index (continuous), HIV infection, AFB smear, cavitation on x-ray, bilateral disease on x-ray, past first- and second-line drug use, MDR category, number of drugs used in treatment, and use of bedaquiline, clofazimine, PAS, moxifloxacin, linezolid, kanamycin, amoxicillin-clavulanate (Amx-Clv), capreomycin, and clarithromycin at anytime during treatment.

Supplemental Table 6.S1. Comparison of patients in the IPD who were included and excluded from this analysis.

	Excluded patients (unsuccessful treatment)	Exclude patients (successful treatment)	Included patients	Overall
n	**5387**	**1183**	**6702**	**13272**
Sex (Female) (%)	2009 (37.3)	481 (40.7)	2719 (40.6)	5209 (39.3)
Age (mean (SD))	37.2 (12.7)	35.5 (12.7)	37.0 (13)	36.9 (12.8)
Body mass index (mean (SD))	19.5 (3.7)	20.3 (3.5)	20.5 (3.8)	20.1 (3.8)
Body mass index category (%)				
Normal	1183 (22.0)	268 (22.7)	2024 (30.2)	3475 (26.2)
Underweight	982 (18.2)	134 (11.3)	1028 (15.3)	2144 (16.2)
Overweight/Obese	160 (3.0)	39 (3.3)	377 (5.6)	576 (4.3)
Missing	3062 (56.8)	742 (62.7)	3273 (48.8)	7077 (53.3)
2018 World Bank income category				
Low/Low-middle	1043 (19.4)	596 (50.4)	1226 (18.3)	2865 (21.6)
Upper-Middle	3635 (67.5)	562 (47.5)	3555 (53.0)	7752 (58.4)
High	709 (13.2)	25 (2.1)	1921 (28.7)	2655 (20.0)
Smoking (%)				
Ex-smoker or never smoker	1070 (19.9)	241 (20.4)	1834 (27.4)	3145 (23.7)
Current smoker	750 (13.9)	105 (8.9)	939 (14.0)	1794 (13.5)
Unknown	3567 (66.2)	837 (70.8)	3929 (58.6)	8333 (62.8)
HIV (%)				
Negative	3005 (55.8)	821 (69.4)	4771 (71.2)	8597 (64.8)
Positive	1969 (36.6)	152 (12.8)	1859 (27.7)	3980 (30.0)
Unknown	413 (7.7)	210 (17.8)	72 (1.1)	695 (5.2)
If HIV positive, on ART	1375 (69.8)	16 (10.5)	1686 (90.7)	3077 (77.3)
Diabetes (%)				
No	2245 (41.7)	339 (28.7)	3311 (49.4)	5895 (44.4)
Yes	254 (4.7)	34 (2.9)	466 (7.0)	754 (5.7)
Unknown	2888 (53.6)	810 (68.5)	2925 (43.6)	6623 (49.9)
Past TB treatment (%)				
No	1297 (24.1)	175 (14.8)	2336 (34.9)	3808 (28.7)
Yes	3986 (74.0)	969 (81.9)	4271 (63.7)	9226 (69.5)
Unknown	104 (1.9)	39 (3.3)	95 (1.4)	238 (1.8)
Past first-line TB drug use (%)				
No	1297 (24.1)	175 (14.8)	2336 (34.9)	3808 (28.7)
Yes	3986 (74.0)	969 (81.9)	4271 (63.7)	9226 (69.5)
Unknown	104 (1.9)	39 (3.3)	95 (1.4)	238 (1.8)
Past second- line TB drug used (%)				
No	3377 (62.7)	662 (56.0)	5048 (75.3)	9087 (68.5)
Yes	1141 (21.2)	54 (4.6)	1226 (18.3)	2421 (18.2)
Unknown	869 (16.1)	467 (39.5)	428 (6.4)	1764 (13.3)

Supplemental Table 6.S1. Continued.

	Excluded patients (unsuccessful treatment)	Exclude patients (successful treatment)	Included patients	Overall
n	**5387**	**1183**	**6702**	**13272**
Cavitation on X-ray (%)				
No	807 (15.0)	159 (13.4)	1606 (24.0)	2572 (19.4)
Yes	1618 (30.0)	211 (17.8)	2308 (34.4)	4137 (31.2)
Unknown	2962 (55.0)	813 (68.7)	2788 (41.6)	6563 (49.4)
Bilateral disease (%)				
No	488 (9.1)	174 (14.7)	1122 (16.7)	1784 (13.4)
Yes	1526 (28.3)	262 (22.1)	1999 (29.8)	3787 (28.5)
Unknown	3373 (62.6)	747 (63.1)	3581 (53.4)	7701 (58.0)
AFB smear result (%)				
Negative	1049 (19.5)	36 (3.0)	1974 (29.5)	3059 (23.0)
Positive	3028 (56.2)	163 (13.8)	4280 (63.9)	7471 (56.3)
Unknown	1310 (24.3)	984 (83.2)	448 (6.7)	2742 (20.7)
Extensive disease, yes (%)	3360 (62.4)	415 (35.1)	4512 (67.3)	8287 (62.4)
DST Performed for Fluoroquinolone	5205 (96.6)	1131 (95.6)	6449 (96.2)	12785 (96.3)
If DST Performed, Fluoroquinolone Resistant	1293 (24.8)	149 (13.2)	1172 (18.2)	2614 (20.4)
DST Performed for Second Line Injectables	5212 (96.8)	1130 (95.5)	6455 (96.3)	12797 (96.4)
If DST Performed, Second Line Injectable Resistant	1599 (30.7)	131 (11.6)	1629 (25.2)	3359 (26.2)
DST Performed for Linezolid	250 (4.6)	23 (1.9)	665 (9.9)	938 (7.1)
If DST Performed, Linezolid Resistant	12 (4.8)	0 (0.0)	16 (2.4)	28 (3.0)
DST Performed for Pyrazinamide	1760 (32.7)	440 (37.2)	3490 (52.1)	5690 (42.9)
If DST Performed, Pyrazinamide Resistant	1019 (57.9)	237 (53.9)	1859 (53.3)	3115 (54.7)
DST Performed for Clofazimine	104 (1.9)	11 (0.9)	252 (3.8)	367 (2.8)
If DST Performed, Clofazimine Resistant	2 (1.9)	1 (9.1)	9 (3.6)	12 (3.3)
DST Performed for Cycloserine/Terizidone	1863 (34.6)	956 (80.8)	2034 (30.3)	4853 (36.6)
If DST Performed, Cycloserine/Terizidone Resistant	136 (7.3)	34 (3.6)	260 (12.8)	430 (8.9)
MDR category (%)				
MDR-TB FQ &SLI sensitive	3198 (59.4)	893 (75.5)	4337 (64.7)	8428 (63.5)
MDR-TB + FQ resistant & SLI sensitive	710 (13.2)	89 (7.5)	929 (13.9)	1728 (13.0)
MDR-TB + SLI resistant & FQ sensitive	404 (7.5)	106 (9.0)	475 (7.1)	985 (7.4)
MDR-TB + SLI & FQ resistance	888 (16.5)	42 (3.6)	688 (10.3)	1618 (12.2)
No DST	187 (3.5)	53 (4.5)	273 (4.1)	513 (3.9)
MDR-TB + SLI & FQ resistance vs. all others (%)				
No	4312 (80.0)	1088 (92.0)	5741 (85.7)	11141 (83.9)
Yes	888 (16.5)	42 (3.6)	688 (10.3)	1618 (12.2)
U	187 (3.5)	53 (4.5)	273 (4.1)	513 (3.9)
Used Ethambutol Ever During Treatment = Yes (%)	2490 (46.2)	281 (23.8)	2895 (43.2)	5666 (42.7)
Used Pyrazinamide Ever During Treatment = Yes (%)	4686 (87.0)	1156 (97.7)	5175 (77.2)	11017 (83.0)
Used Streptomycin Ever During Treatment = Yes (%)	386 (7.2)	2 (0.2)	692 (10.3)	1080 (8.1)
Used Rifabutin Ever During Treatment = Yes (%)	58 (1.1)	1 (0.1)	154 (2.3)	213 (1.6)

Supplemental Table 6.S1. Continued.

	Excluded patients (unsuccessful treatment)	Exclude patients (successful treatment)	Included patients	Overall
n	**5387**	**1183**	**6702**	**13272**
Used Amikacin Ever During Treatment = Yes (%)	621 (11.5)	168 (14.2)	1048 (15.6)	1837 (13.8)
Used Capreomycin Ever During Treatment = Yes (%)	1569 (29.1)	232 (19.6)	1446 (21.6)	3247 (24.5)
Used Kanamycin Ever During Treatment = Yes (%)	2787 (51.7)	784 (66.3)	3151 (47.0)	6722 (50.6)
Used Ofloxacin Ever During Treatment = Yes (%)	1658 (30.8)	528 (44.6)	1373 (20.5)	3559 (26.8)
Used Ciprofloxacin Ever During Treatment = Yes (%)	204 (3.8)	4 (0.3)	266 (4.0)	474 (3.6)
Used Moxifloxacin Ever During Treatment = Yes (%)	2194 (40.7)	109 (9.2)	3459 (51.6)	5762 (43.4)
Used Levofloxacin Ever During Treatment = Yes (%)	1123 (20.8)	535 (45.2)	1889 (28.2)	3547 (26.7)
Used Ethionamide Ever During Treatment = Yes (%)	3111 (57.8)	676 (57.1)	2859 (42.7)	6646 (50.1)
Used Prothionamide Ever During Treatment = Yes (%)	1453 (27.0)	463 (39.1)	2258 (33.7)	4174 (31.4)
Used Cycloserine Ever During Treatment = Yes (%)	2351 (43.6)	899 (76.0)	2873 (42.9)	6123 (46.1)
Used Terizidone Ever During Treatment = Yes (%)	1963 (36.4)	29 (2.5)	2922 (43.6)	4914 (37.0)
Used PAS Ever During Treatment = Yes (%)	2344 (43.5)	565 (47.8)	2759 (41.2)	5668 (42.7)
Used Linezolid Ever During Treatment = Yes (%)	638 (11.8)	52 (4.4)	1594 (23.8)	2284 (17.2)
Used Clofazimine Ever During Treatment = Yes (%)	651 (12.1)	111 (9.4)	1101 (16.4)	1863 (14.0)
Used Amx-Clv Ever During Treatment = Yes (%)	763 (14.2)	93 (7.9)	994 (14.8)	1850 (13.9)
Used Thioacetazone Ever During Treatment = Yes (%)	30 (0.6)	0 (0.0)	68 (1.0)	98 (0.7)
Used Clarithromycin Ever During Treatment = Yes (%)	507 (9.4)	84 (7.1)	485 (7.2)	1076 (8.1)
Used Imp-Cilastatin Ever During Treatment = Yes (%)	78 (1.4)	2 (0.2)	237 (3.5)	317 (2.4)
Used Meropenem Ever During Treatment = Yes (%)	23 (0.4)	3 (0.3)	61 (0.9)	87 (0.7)
Used Bedaquiline Ever During Treatment = Yes (%)	756 (14.0)	17 (1.4)	1605 (23.9)	2378 (17.9)
Used Delamanid Ever During Treatment = Yes (%)	46 (0.9)	0 (0.0)	114 (1.7)	160 (1.2)
Number of drugs (median [IQR])	5 [4, 5]	4 [4, 5]	5 [4, 6]	5 [4, 5]
Number of effective drugs (median [IQR])	4 [4, 5]	4 [4, 5]	4 [4, 5]	4 [4, 5]
Number of limited access drugs** (median [IQR])	0 [0, 0]	0 [0, 0]	0 [0, 1]	0 [0, 1]
Total Treatment duration (mean (SD))	14.5 (9.8)	NA	22.0 (4.6)	19 (8)
Median [IQR]	14 [6, 23]	NA	22 [19, 24]	21 [16, 24]

SD: standard deviation; XDR: extensively drug resistant tuberculosis; MDR: multidrug resistant tuberculosis; TB: tuberculosis; AFB: acid-fast bacillus; Amx-Clv: Amoxicillin-Clavulanic Acid; Imp: imipenem. ** includes bedaquiline, clofazimine, linezolid, imipenem, and meropenem.

Supplemental Table 6.S2. Mean total treatment duration and total number of patients at each site included in the study population.

Author	Treatment site*	Number of patients	Mean total treatment duration (SD)
Ahuja	USA	38	23.5 (8.9)
Anderson	UK	90	21.4 (7.2)
Fox	Australia	25	21.7 (3.7)
Bang	Denmark	19	18.3 (4.1)
Barry/Flood (Calif)	USA	45	21.1 (3.4)
Barry (Korea)	South Korea	30	25.1 (3.1)
	Brazil	2	22 (3.8)
	India	4	24 (0)
	Latvia	3	21.8 (2)
TMC207-C208	Peru	17	21 (2.6)
	Philippines	1	25.6 (NA)
	South Africa	28	22.5 (3.2)
	Thailand	1	16.3 (NA)
Skrahina	Belarus	94	23.6 (1.7)
Skrahina (2019)	Belarus	106	23.2 (2.7)
Bonnet	Georgia	68	26.7 (5.3)
Rodrigues	Brazil	82	18.4 (2.2)
Brode	Canada	17	24.6 (1.9)
	Estonia	30	21.6 (3.6)
	Latvia	99	20.3 (3.6)
	Peru	109	23.8 (6.8)
	Philippines	320	21.6 (3.3)
Cegielski	Russia	81	21.9 (3.5)
	South Africa	250	24.1 (4)
	South Korea	46	25.8 (6)
	Taiwan	45	21.4 (2.1)
	Thailand	41	20.3 (3.8)
Chan (Denver)	USA	7	28.1 (17.1)
Dheda	South Africa	14	33.9 (12.1)
Guglielmetti	France	35	21.6 (3.2)
Guglielmetti	France	9	25.1 (2)
Isaakidis	India	69	23 (6.5)
	Asia	58	23.1 (2.9)
TMC207-209	Europe	38	20.1 (2.9)
	Peru	10	18.2 (2.5)
	South Africa	39	22.3 (3.3)
Jarlsberg	USA	20	21.6 (4.5)
Kempker	Georgia	84	24.1 (3.2)
Koenig	Haiti	126	24 (0.6)
Koh	South Korea	272	23 (5.4)
Kvasnovsky	South Africa	34	29.7 (7.1)

Supplemental Table 6.S2. Continued

Author	Treatment site*	Number of patients	Mean total treatment duration (SD)
Lange	Germany	91	23.8 (1.5)
Laniado-Laborin	Mexico	37	22.6 (5.3)
Kuksa	Latvia	31	18.4 (3.6)
Barkane	Latvia	26	17 (4)
Leung	Hong Kong	136	16.7 (3)
Marks	USA	92	22.6 (6)
Migliori	Belarus	6	16.8 (2.8)
	Belgium	13	25.3 (5.3)
	Brazil	3	35.3 (11)
	Ecuador	1	12 (NA)
	Greece	11	23 (5.4)
	Italy	74	19.9 (3.6)
	Netherlands	43	17.5 (3.5)
	Peru	13	20 (11.9)
	Slovakia	2	25 (1.4)
	UK	3	21.8 (3.9)
Migliori (BDQ)	Australia	1	24 (NA)
	Belgium	2	21.3 (0.4)
	Greece	2	25 (0)
	India	11	34.3 (4.5)
	Italy	7	22.6 (2.8)
	Netherlands	2	20 (0)
	Peru	1	36 (NA)
	Russia	48	23.4 (5)
	South Africa	24	24.7 (4.7)
	Sweden	3	26.7 (6.4)
Milanov	Bulgaria	23	23.4 (2.1)
Achar	Uzbekistan	63	26.4 (4.4)
Isaakidis	India	63	23 (3.9)
Ndjeka	South Africa	60	20.4 (2.8)
O'Donnell	South Africa	25	24.1 (2.1)
Podewils	Philippines	385	20.7 (3)
Riekstina/Leimane	Latvia	108	19.4 (3.4)
Vasilyeva	Russia	68	20 (3.4)
Shim	South Korea	42	22.3 (4.5)
Singla	India	20	24.7 (0.8)
Smith	Russia	114	26.6 (3.9)
Ndjeka	South Africa	2128	21.7 (3.2)
Seo	South Korea	19	30 (11.1)
Udwadia	India	12	19 (2.3)
van der Werf	Netherlands	75	16.7 (3.4)
Viiklepp	Estonia	205	19.3 (4.3)
Yim/Kwak	South Korea	103	25.3 (7.5)

Treatment sites were identified by country within a given study. SD: standard deviation

Supplemental Table 6.S3. Associations between deviation in treatment duration from site mean and patient characteristics, resistance categories, and drugs used from the full multivariable model with inverse probability of selection weights and without (primary analysis), including all variables listed.

	With inverse probability of selection weighting months (95% CI)	Without weighting months (95% CI)
Clinical characteristics		
Age (per year increase)	0 (0.0, 0.01)	0 (0.0, 0.01)
Sex (Female)	-0.09 (-0.29, 0.11)	-0.11 (-0.3, 0.08)
Body mass index (per unit increase)	-0.02 (-0.06, 0.01)	-0.02 (-0.05, 0.01)
HIV infection	0.34 (0.08, 0.61)	0.33 (0.06, 0.6)
AFB smear positive	0.77 (0.54, 1.0)	0.82 (0.6, 1.04)
Cavitation on X-Ray	0.35 (0.09, 0.61)	0.3 (0.05, 0.56)
Bilateral disease on X-ray	0.4 (0.1, 0.71)	0.24 (-0.07, 0.54)
Treatment history and drug resistance		
Past first-line drug use	0.36 (0.11, 0.62)	0.38 (0.14, 0.62)
Past second-line drug use	0.29 (-0.06, 0.65)	0.21 (-0.16, 0.58)
Number of drugs used	0.3 (0.16, 0.43)	0.34 (0.21, 0.47)
MDR-TB + FQ resistant & SLI sensitive	0.39 (0.08, 0.71)	0.34 (0.03, 0.65)
MDR-TB + SLI resistant & FQ sensitive	0.72 (0.3, 1.15)	0.78 (0.36, 1.21)
MDR-TB + SLI & FQ resistance	0.63 (0.21, 1.04)	0.61 (0.21, 1.02)
Drugs used in treatment		
Used Bedaquiline Ever During Treatment	-0.39 (-0.79, 0.0)	-0.51 (-0.87, -0.15)
Used Linezolid Ever During Treatment	0.09 (-0.3, 0.48)	0.29 (-0.06, 0.63)
Used Moxifloxacin Ever During Treatment	0.37 (0.13, 0.61)	0.32 (0.09, 0.54)
Used Clofazimine Ever During Treatment	0.16 (-0.24, 0.56)	-0.15 (-0.51, 0.21)
Used Capreomycin Ever During Treatment	0.7 (0.41, 0.99)	0.69 (0.41, 0.97)
Used Kanamycin Ever During Treatment	0.46 (0.22, 0.71)	0.44 (0.2, 0.68)
Used PAS Ever During Treatment	-0.05 (-0.3, 0.2)	-0.03 (-0.27, 0.21)
Used Amx-Clv Ever During Treatment	0.67 (0.32, 1.02)	0.73 (0.4, 1.06)
Used Clarithromycin Ever During Treatment	0.85 (0.43, 1.28)	1.12 (0.71, 1.53)

Supplemental Table 6.S4. E-values for selected characteristics with largest effect estimates for deviation in treatment duration.

	Linear regression estimate (95% CI)	Converted RR (95% CI)	E-Value (limit†)
Used Bedaquiline Ever During Treatment	-0.51 (-0.87, -0.15)	0.89 (0.82, 0.97)	1.50 (NE, 1.23)
Used Clarithromycin Ever During Treatment	1.12 (0.71, 1.53)	1.29 (1.18, 1.42)	1.90 (1.63, NE)
MDR-TB + SLI resistant & FQ sensitive	0.34 (0.03, 0.65)	1.08 (1.01, 1.16)	1.38 (1.09, NE)
MDR-TB + FQ resistant & SLI sensitive	0.78 (0.36, 1.21)	1.20 (1.09, 1.32)	1.68 (1.39, NE)
Past first-line drug use	0.38 (0.14, 0.62)	1.09 (1.03, 1.15)	1.41 (1.21, NE)
HIV infection	0.33 (0.06, 0.60)	1.08 (1.01, 1.15)	1.37 (1.13, NE)
AFB smear positive	0.82 (0.60, 1.04)	1.21 (1.15, 1.27)	1.71 (1.56, NE)
Cavitation on X-Ray	0.30 (0.05, 0.56)	1.07 (1.01, 1.14)	1.35 (1.12, NE)

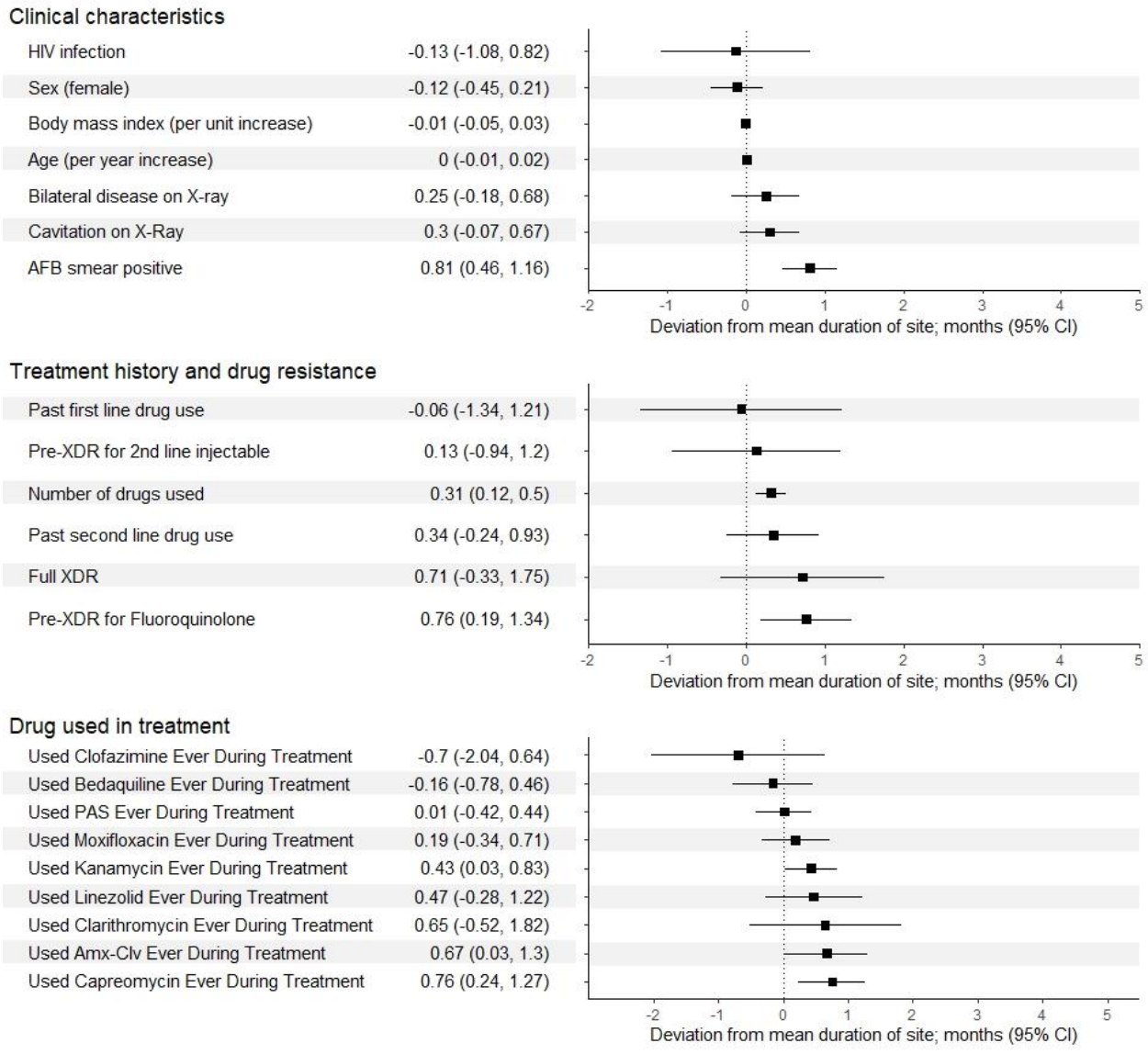
RR: risk ratio, CI: confidence interval. Risk ratios are converted from linear regression coefficient. †E-values are calculated as follows: for RR>1: E-Value = RR + $\sqrt{(RR*(RR-1))}$ with the confidence interval (CI) of the E-value being the lower limit of the RR (LL) + $\sqrt{(LL*(LL-1))}$, if the LL>1, while if LL≤ then the CI = 1; for RR<1: E-value = 1/RR + $\sqrt{(1/RR*([1/RR]-1))}$ with the CI of the E-value being 1/upper limit of the RR (UL) +$\sqrt{(1/UL*([1/UL]-1))}$, if the UL<1, while if UL≥ 1 then CI = 1. E-values interpreted as the minimum strength of association that an unmeasured confounder would need to have with both the treatment and the outcome to fully explain away a specific treatment-outcome association.

Supplemental Table 6.S5. Associations between deviation in treatment duration from site mean and patient characteristics, resistance categories, and drugs used by different combinations of subgroups, and their regression estimates and 95% confidence intervals (CI) from a multivariable linear mixed model including all variables listed (unless specified otherwise).

Characteristic	Those with extensive disease and only MDR-TB months (95% CI)	Those without extensive disease and with resistance in addition to MDR-TB* months (95% CI)	Those with past treatment and only MDR-TB months (95% CI)	Those without past treatment and resistance in addition to MDR-TB* months (95% CI)
Clinical characteristics				
Age (per year increase)	0.01 (0, 0.02)	0 (-0.02, 0.03)	0.02 (0.01, 0.03)	-0.02 (-0.05, 0.01)
Sex (Female)	-0.21 (-0.48, 0.05)	0.16 (-0.5, 0.82)	-0.18 (-0.47, 0.11)	-0.76 (-1.52, 0)
Body mass index (per unit increase)	-0.03 (-0.07, 0.01)	-0.02 (-0.13, 0.09)	-0.04 (-0.08, 0)	0 (-0.14, 0.14)
HIV infection	0.32 (-0.03, 0.67)	0.42 (-0.38, 1.23)	0.39 (-0.01, 0.79)	0.07 (-0.97, 1.11)
AFB smear positive	NE	NE	0.69 (0.34, 1.05)	0.8 (-0.03, 1.63)
Cavitation on X-Ray	NE	NE	0.25 (-0.11, 0.61)	0.83 (-0.12, 1.79)
Bilateral disease on X-ray	NE	NE	0.35 (-0.08, 0.78)	-0.05 (-1.15, 1.04)
Treatment history				
Past first-line drug use	0.56 (0.24, 0.88)	0.16 (-0.7, 1.02)	NE	NE
Past second-line drug use	0.05 (-0.52, 0.63)	0.41 (-0.46, 1.29)	NE	NE
Number of drugs used	0.4 (0.21, 0.6)	0.07 (-0.31, 0.45)	0.29 (0.07, 0.5)	0.3 (-0.17, 0.77)
Drugs used in treatment				
Used Bedaquiline Ever During Treatment	-0.43 (-0.92, 0.05)	-0.43 (-1.42, 0.57)	-0.69 (-1.24, -0.13)	-1.41 (-2.71, -0.1)
Used Moxifloxacin Ever During Treatment	0.52 (0.18, 0.85)	0.06 (-0.6, 0.72)	0.25 (-0.09, 0.6)	0.79 (-0.04, 1.61)
Used Linezolid Ever During Treatment	0.52 (-0.07, 1.12)	0.9 (-0.04, 1.85)	0.6 (-0.1, 1.29)	1.5 (0.37, 2.63)
Used Clofazimine Ever During Treatment	-0.84 (-1.47, -0.22)	0.42 (-0.5, 1.35)	-0.22 (-0.95, 0.5)	0.57 (-0.57, 1.7)
Used Capreomycin Ever During Treatment	0.74 (0.31, 1.16)	0.89 (0.14, 1.65)	0.71 (0.24, 1.17)	0.98 (-0.04, 2)
Used Kanamycin Ever During Treatment	0.49 (0.18, 0.8)	1.12 (0.21, 2.03)	0.24 (-0.09, 0.57)	1.45 (0.41, 2.49)
Used Amx-Clv Ever During Treatment	1.32 (0.73, 1.92)	0.11 (-0.64, 0.85)	1.42 (0.77, 2.06)	0.12 (-0.89, 1.14)
Used Clarithromycin Ever During Treatment	0.07 (-0.62, 0.76)	1.74 (0.74, 2.75)	0.86 (0.15, 1.57)	1.57 (0.11, 3.03)

* Includes MDR-TB + resistance to a fluoroquinolone (but not a second-line injectable), MDR-TB + resistance to a second-line injectable (but not a fluoroquinolone), and MDR-TB with resistance to both a fluoroquinolone and a second-line injectable. All models also adjusted for use of PAS.

Supplemental Figure 6.S1. Forest plot of associations between deviation in treatment duration (in months) from site mean and patient characteristics, resistance categories, and drugs used analyzed using imputed outcomes for subjects with missing or planned duration. Estimates and 95% confidence intervals (CI) from regression using a multivariable linear mixed model, including all variables shown (n = 7885).



* Conditional R^2 for model: 0.15. Note the patients whose outcomes were imputed include those who had either missing duration or only planned duration.

REFERENCES

1. World Health Organization. Global tuberculosis report 2022. Geneva, Switzerland: World Health Organization, 2022.
2. Kwak N, Winters N, Campbell JR, et al. Changes in treatment for multidrug-resistant tuberculosis according to national income. *Eur Respir J* 2020; **56**(5).
3. World Health Organization. Consolidated guidelines on drug-resistant tuberculosis treatment. Geneva, Switzerland, 2019.
4. Nigam S, Sharma RK, Yadav R, et al. Experiences and needs of patients with MDR/XDR-TB: a qualitative study among Saharia tribe in Madhya Pradesh, Central India. *BMJ Open* 2021; **11**(8): e044698.
5. Deshmukh RD, Dhande DJ, Sachdeva KS, et al. Patient and Provider Reported Reasons for Lost to Follow Up in MDRTB Treatment: A Qualitative Study from a Drug Resistant TB Centre in India. *PloS one* 2015; **10**(8): e0135802.
6. Shringarpure KS, Isaakidis P, Sagili KD, Baxi RK, Das M, Daftary A. "When Treatment Is More Challenging than the Disease": A Qualitative Study of MDR-TB Patient Retention. *PloS one* 2016; **11**(3): e0150849.
7. Kohler S, Sitali N, Achar J, Paul N. Programme costs of longer and shorter tuberculosis drug regimens and drug import: a modelling study for Karakalpakstan, Uzbekistan. *ERJ Open Res* 2022; **8**(1).
8. Van Deun A, Maug AK, Salim MA, et al. Short, highly effective, and inexpensive standardized treatment of multidrug-resistant tuberculosis. *Am J Respir Crit Care Med* 2010; **182**(5): 684-92.
9. Nunn AJ, Phillips PPJ, Meredith SK, et al. A Trial of a Shorter Regimen for Rifampin-Resistant Tuberculosis. *N Engl J Med* 2019; **380**(13): 1201-13.
10. Goodall RL, Meredith SK, Nunn AJ, et al. Evaluation of two short standardised regimens for the treatment of rifampicin-resistant tuberculosis (STREAM stage 2): an open-label, multicentre, randomised, non-inferiority trial. *Lancet* 2022; **400**(10366): 1858-68.
11. Conradie F, Bagdasaryan TR, Borisov S, et al. Bedaquiline-Pretomanid-Linezolid Regimens for Drug-Resistant Tuberculosis. *N Engl J Med* 2022; **387**(9): 810-23.
12. Mok J, Lee M, Kim DK, et al. 9 months of delamanid, linezolid, levofloxacin, and pyrazinamide versus conventional therapy for treatment of fluoroquinolone-sensitive multidrug-resistant tuberculosis (MDR-END): a multicentre, randomised, open-label phase 2/3 non-inferiority trial in South Korea. *Lancet* 2022; **400**(10362): 1522-30.
13. Ahmad Khan F, Salim MAH, du Cros P, et al. Effectiveness and safety of standardised shorter regimens for multidrug-resistant tuberculosis: individual patient data and aggregate data meta-analyses. *Eur Respir J* 2017; **50**(1).
14. Collaborative Group for the Meta-Analysis of Individual Patient Data in MDRTBt, Ahmad N, Ahuja SD, et al. Treatment correlates of successful outcomes in pulmonary multidrug-resistant tuberculosis: an individual patient data meta-analysis. *Lancet* 2018; **392**(10150): 821-34.
15. Bastos ML, Lan Z, Menzies D. An updated systematic review and meta-analysis for treatment of multidrug-resistant tuberculosis. *Eur Respir J* 2017; **49**(3).
16. Ahuja SD, Ashkin D, Avendano M, et al. Multidrug resistant pulmonary tuberculosis treatment regimens and patient outcomes: an individual patient data meta-analysis of 9,153 patients. *PLoS medicine* 2012; **9**(8): e1001300.

17. World Health Organization. Public call for individual patient data on treatment of rifampicin and multidrug-resistant (MDR/RR-TB) tuberculosis. . 2018. https://www.who.int/news-room/articles-detail/public-call-for-individual-patient-data-on-treatment-of-rifampicin-and-multidrug-resistant-(mdr-rr-tb)-tuberculosis (accessed 10 Nov 2022.
18. World Health Organization. Public call for individual patient data on treatment of drug-resistant tuberculosis. 2019. https://www.who.int/news/item/01-08-2019-public-call-for-individual-patient-data-on-treatment-of-drug-resistant-tuberculosis (accessed 10 Nov 2022.
19. Bisson GP, Bastos M, Campbell JR, et al. Mortality in adults with multidrug-resistant tuberculosis and HIV by antiretroviral therapy and tuberculosis drug use: an individual patient data meta-analysis. *Lancet* 2020; **396**(10248): 402-11.
20. Campbell JR, Chan ED, Falzon D, et al. Low Body Mass Index at Treatment Initiation and Rifampicin-Resistant Tuberculosis Treatment Outcomes: An Individual Participant Data Meta-Analysis. *Clin Infect Dis* 2022.
21. Laserson KF, Thorpe LE, Leimane V, et al. Speaking the same language: treatment outcome definitions for multidrug-resistant tuberculosis. *The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease* 2005; **9**(6): 640-5.
22. World Health Organization. WHO operational handbook on tuberculosis. Module 4: treatment - drug-resistant tuberculosis treatment. Geneva: World Health Organization, 2022.
23. van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software* 2011; **45**(3): 1 - 67.
24. Zou H. The Adaptive Lasso and Its Oracle Properties. *Journal of the American Statistical Association* 2006; **101**(476): 1418-29.
25. VanderWeele TJ, Ding P. Sensitivity Analysis in Observational Research: Introducing the E-Value. *Annals of internal medicine* 2017; **167**(4): 268-74.
26. R Core Team (2020). R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria, URL https://wwwR-projectorg/*.
27. Ndjeka N, Campbell JR, Meintjes G, et al. Treatment outcomes 24 months after initiating short, all-oral bedaquiline-containing or injectable-containing rifampicin-resistant tuberculosis treatment regimens in South Africa: a retrospective cohort study. *Lancet Infect Dis* 2022; **22**(7): 1042-51.
28. Conradie F, Diacon AH, Ngubane N, et al. Treatment of Highly Drug-Resistant Pulmonary Tuberculosis. *N Engl J Med* 2020; **382**(10): 893-902.
29. Yoshida K, Solomon DH, Kim SC. Active-comparator design and new-user design in observational studies. *Nat Rev Rheumatol* 2015; **11**(7): 437-41.
30. Leeb H, Pötscher BM. Model Selection and Inference: Facts and Fiction. *Econometric Theory* 2005; **21**(1): 21-59.

Chapter 7 – Summary & conclusions

7.1 Summary

The aim of this doctoral thesis was to generate evidence that could advance the scientific knowledge for treatment of tuberculosis infection and MDR-TB, and primarily to provide evidence for more effective, safer, and shorter treatments. In my first objective, I constructed a data set of individual patient data from randomized trials and conducted a network meta-analysis to compare the completion, safety, and efficacy between two tuberculosis preventive treatments, 3HP and 4R, that have previously not been directly compared in head-to-head trials. My second objective involved both determining the efficacy of adding clofazimine to the WHO group A drugs (bedaquiline, linezolid, and a fluoroquinolone) for the treatment of MDR-TB and whether time-varying confounding affected the average treatment effect using observational data to emulate a target trial. In my third objective, I used individual patient data from observational studies of MDR-TB treatment to assess associations between the treatment duration of an individual and patient characteristics, including clinical factors (such as HIV infection, cavitation on x-ray, and additional drug resistance), as well as drugs used in their treatment to identify patients who may benefit from shorter treatment regimens.

With my first manuscript, I demonstrated that in 17,572 patients given tuberculosis preventive treatment in six randomized trials, 3HP had slightly higher treatment completion compared to 4R with an adjusted risk difference of 0.05 (95% CI 0.02 to 0.07). For treatment efficacy we found no difference in the incidence of tuberculosis disease between 3HP and 4R. However, in this study 3HP had higher risk of any adverse events that led to treatment

discontinuation and notably a higher risk of grade 3 to 4 adverse events. These increased risks of adverse events are important, as safety is one of the main priorities of a preventive treatment.[31,54-56] Although there was an increase in treatment completion for 3HP in the overall population, this result was not robust once the study with comparator arm of 6H was removed.[49] Additionally, 3HP was administered under direct observation and the results of the iAdhere study indicated that there were higher rates of completion for those on directly observed therapy (87.2%) than those on self-administered treatment (74%).[51] Therefore, in programmatic settings, if directly observed therapy is not be available, it is possible that 4R would have similar or better treatment completion than 3HP with substantially less risk of adverse events to the patient, and in these situations 4R should be prioritized.

In my second manuscript I found some evidence in the ITT analysis and the second per-protocol analysis that adding clofazimine to three WHO group A drugs provided some benefit for treatment success, with about a 7% increase in treatment success, although confidence intervals included the null. However, the use of censoring weights with adjustment for time-varying confounders in the both the first and third per-protocol analyses, where patients in the control group were censored for starting clofazimine, substantially attenuated the estimate of the average treatment effect compared to their respective unadjusted and baseline IPTW estimates of the same patients as well as compared to estimates in the ITT analysis. Such a substantial attenuation of the unadjusted estimates indicate that there is an important impact of IPCW and time-varying confounders under these types of censoring.

In the third manuscript of this thesis, using 6,702 patients with successful treatment outcomes for MDR-TB I demonstrated that some patients who were previously excluded

from randomized trials of shorter treatment and from WHO treatment recommendations may benefit from shorter treatment. In this individual patient data meta-analysis the use of bedaquiline was associated with a 0.51 (95% CI: 0.15, 0.87) month decrease in duration of treatment, which was consistently observed across the subgroups of populations previously excluded from trials and WHO guideline recommendations (such as those with extensive disease and those with past tuberculosis treatment). In this analysis, clinical factors known in the literature to be associated with poor treatment outcomes, such as extensive disease and HIV, as well as drugs known to have poor efficacy in treating MDR-TB were also associated with longer treatment durations.

7.2 Limitations & strengths

In my first manuscript, there were some limitations to consider. An overall limitation is the indirect nature of the comparison, although randomization is preserved for the individual trials[132] the groups being indirectly compared have not been randomized. To account for this, adjustments for patient- and study-level variables were made to reduce confounding between studies, which is one of the important advantages of using individual patient data. There were also limitations with the ability to account for variance, as treatment site and country could not be included as random intercepts due to missing data for a large portion of the population, leading to an underestimation of variance. However, we assessed both model fit (using Akaike information criterion[133] and Bayesian information criterion[134]) and changes in variance between a model fit using a country variable with a missing category specified as a random intercept and our model fit with just a random intercept for study, and observed no substantial differences. In absence of trials directly comparing 6H to 9H, an

assumption of equivalence was made between the two longer isoniazid regimens, and potential bias may exist due to their pooling (a sensitivity analysis where studies of 6H were removed indicated an attenuation in treatment completion for 3HP). Due to low numbers of adverse events, analyses within important clinical subgroups (people living with HIV or diabetes) were not possible. Additionally, calendar dates were unavailable as the data received were deidentified, precluding assessment of temporal trends within trials. Many sensitivity analyses were conducted to address most limitations, as outlined in the manuscript, and it is likely these limitations would not alter the overall results of the analysis in a substantial way.

Nonetheless, my first objective had several strengths. This was the first study to compare completion, safety, and efficacy between 3HP and 4R. The use of individual patient data allowed for the harmonization of outcomes and adjustments for confounders so that estimates would be more valid than traditional aggregate data meta-analysis and network meta-analysis. The large sample size allowed for the assessment of treatment completion in important subgroups by age and HIV status. Furthermore, with the use of individual patient data many definitions of adverse events could be created and harmonized that allowed for comparisons of several types of adverse events between regimens.

In my second manuscript the analysis was limited by available sample size of those receiving the two regimens being compared, which affected statistical power. The sample size also limited subgroup analyses for important clinical populations (such as DST results for fluoroquinolones, people living with HIV, or those with extensive disease). Additionally, we were not able to restrict our analysis to those who were susceptible to all drugs in each treatment group, as drug susceptibility testing for each drug in the intervention regimens

were not performed for most patients, and sample size would have been inadequate for analysis. Attempts were made to account for this, including adjustment for baseline resistance for all drugs and the third and fourth per-protocol analyses censored patients if they stopped a drug due to acquired drug resistance. We also did no have reasons in the data for why the patients in the control group started clofazimine (although likely to replace drugs outside the core regimen that were stopped due to an adverse event or prescribed to patients who were failing treatment) and the ability to adequately explain the large attenuation observed for the censoring weighted average treatment effects in the first and third per-protocol analyses is limited. Thus, the estimates should be interpreted cautiously, and future studies should consider recording data on reasons for starting drugs after baseline.

This analysis had many strengths, however. The data used had detailed records of drugs given during treatment and their start and stop dates (including reason for why drugs were stopped, and the data were collected with the purpose of accounting for time-varying confounders.[135] Additionally, the use IPCW methods to control for time-varying confounding helped to minimize the biases that may affect analyses for efficacy of drugs given at treatment initiation alone.[87,88,126,127] By using the target trial framework and specifically outlining our hypothetical protocol, we aimed to reduce other biases that are common in observational research[123,124,126] so that focus could be on controlling for measured confounding. We also had explicit eligibility criteria including timing of treatment initiation (including only patients who received their treatment intervention at baseline) to reduce potential immortal time bias[129] and we also accounted for adherence to treatment strategies

with use of time-varying censoring weights, which is especially important when studying the sustained treatment strategies inherent in MDR-TB treatment.[128,129]

My third manuscript also had limitations. The primary limitation is the inability to draw any causal relationships between shorter treatment duration and the factors that were assessed. This is due to the hypothesis generating approach of the analysis, and as such these associations should be confirmed in future studies. The analysis was also conducted on a study population treated between 1993 and 2019 and treatment practices have since changed, including greater use of fluoroquinolones and broader uptake of anti-retroviral therapies. Additionally, there is also the potential for unmeasured confounding and indication bias (due to site and physician level practices) affecting the associations found. However, the outcome we used, the individual deviation from the site-specific mean treatment duration, helped to reduce the site-level differences that may bias results.

However, the goal of this objective was not to inform clinical practice but for hypothesis generation with the aim of informing inclusion criteria into future trials, and for that purpose it had several strengths. This analysis was conducted using a large sample size that allowed for many subgroup analyses of populations that were excluded from trials[10-13] and who are currently not recommended for shorter treatments by the WHO.[9] The analysis was also conducted using observational data from 34 countries, which should provide evidence that is more generalizable to programmatic settings than data from highly restricted and monitored clinical trials. Finally, this was one of the first studies to provide descriptions of site-level factors affecting treatment duration, which had previously not been described.

7.3 Implications and directions for future research

The results from my first manuscript may be used to inform guidelines for tuberculosis preventive treatment, in absence of trials directly comparing 4R and 3HP. Clinicians now have evidence to help decide between available treatments. For future research, it would be ideal for these results to be confirmed in a randomized trial of 4R compared to 3HP, however the likelihood of such a trial is low as cost would be high and these drugs are already used in programmatic settings. This data set is also the first of its kind and can be expanded with new data as more trials finish, allowing for new network meta-analyses or direct comparisons to be conducted. Other shorter treatments that could be added include one month of isoniazid plus rifapentine (1HP)[136], for which data are already available. Additionally, a randomized controlled trial (NCT03988933) is currently underway to compare 4R to shorter durations of higher dose rifampicin (2 months of rifampicin at 20mg/kg and 2 months of rifampicin at 30mg/kg), the results of which will be available soon.

In my second manuscript we found some evidence in the ITT analysis and second per-protocol analysis that in patients receiving other concurrent MDR-TB drugs, adding clofazimine to the three group A drugs (bedaquiline, linezolid, and a fluoroquinolone) may improve treatment success compared to the same regimen without clofazimine. However, applying censoring weights that accounted for censoring of patients in the control group who started clofazimine in the first and third per-protocol analyses resulted in a substantial reduction of the unadjusted and IPTW estimates of the ATE in the same populations, indicating an important effect of censoring weighting and time-varying confounding under these types of censoring. Hopefully more researchers will consider incorporating target trial

framework and time-varying IPCW methodology into future observational studies of MDR-TB treatment to better account for potential biases that may arise throughout the course of treatment.[87,137] This may be especially important when studying drugs with poor safety profiles that may have a larger effect of time-varying confounding, as such drugs have a higher likelihood for causing changes to treatment regimens. This also requires that observational data be collected with this purpose in mind, including detailed information on timing and reasons for not only drug stoppages but also for drugs started later in treatment. Additionally, when trials cannot be conducted due to cost, time, or feasibility, an emulation of a target trial is a useful alternative to obtain a causal effect estimate for a research question using observational data.[124]

Finally, the results of my third manuscript may be used to inform inclusion criteria for future trials on shorter treatments for MDR-TB, as some evidence was found that indicates patients with more advanced disease or past tuberculosis drug use may benefit from shorter treatment. Investigators should consider including these patients in future trials assessing shorter MDR-TB regimens. Further, both this third objective and my second objective indicated the potency that regimens containing bedaquiline have on treatment success regardless of the addition of other drugs or comorbidities of the patients. If future trials can be used to show that MDR-TB patients with advanced disease and extensive resistance can also benefit from shorter treatment if they receive bedaquiline containing regimens, this will help validate and support the application of this analytic approach for use in future research on treatment duration. Importantly, trial results indicating efficacy of shorter regimens in these previously excluded populations will help improve the currently arduous treatment experience for a broader range of patients.

7.4 Conclusion

Treatment for all forms of tuberculosis is burdensome for those affected. Shorter, safer, and more effective treatments are an integral part of improving patient experience and outcomes. The work outlined in this doctoral thesis has generated evidence that will help improve treatments for both tuberculosis infection and MDR-TB.

The results from my individual patient data network meta-analysis have provided evidence for 4R and 3HP in terms of treatment completion, safety, and efficacy in TB prevention, which was previous lacking. Now, clinicians have something to draw on when deciding between these two tuberculosis preventive treatments.

Considering very few new drugs are available to treat MDR-TB, evidence for the efficacy of regimens that include currently available drugs is desperately needed. Additionally, conducting randomized clinical trials for MDR-TB treatment is complicated due to the highly individualized treatments and such trials would be expensive and time consuming, therefore observational data will continue to be used to inform clinical practice. Use of causal methods, implementation of a target trial framework, and accounting for time-varying confounding may improve the validity of evidence that can be generated from use of observational data, and help identify more effective regimens for MDR-TB. Additionally, clinical trials of shorter MDR-TB treatment should consider broadening their inclusion criteria to include patients with more advanced disease and extensive resistance, as the results outlined in this thesis indicate these persons may also benefit from shorter treatment. These analyses also provide methodological contributions to tuberculosis research: the first being the application of causal methods for assessing treatment efficacy and the second being a novel approach to assessing treatment duration.

Although my thesis work has contributed to the scientific knowledge on treatment of tuberculosis infection and MDR-TB, there is still much work needed to improve treatment outcomes and help reduce the devastating effects of tuberculosis.

Back Matter

References

1. World Health Organization. Global tuberculosis report 2022. Geneva, Switzerland: World Health Organization, 2022.
2. Cohen A, Mathiasen VD, Schon T, Wejse C. The global prevalence of latent tuberculosis: a systematic review and meta-analysis. *Eur Respir J* 2019; **54**(3).
3. Houben RM, Dodd PJ. The Global Burden of Latent Tuberculosis Infection: A Re-estimation Using Mathematical Modelling. *PLoS medicine* 2016; **13**(10): e1002152.
4. World Health Organization. Latent tuberculosis infection: updated and consolidated guidelines for programmatic management. Geneva: World Health Organization, 2018.
5. Reid A, Grant AD, White RG, et al. Accelerating progress towards tuberculosis elimination: the need for combination treatment and prevention. *The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease* 2015; **19**(1): 5-9.
6. Uplekar M, Weil D, Lonnroth K, et al. WHO's new end TB strategy. *Lancet* 2015; **385**(9979): 1799-801.
7. Kwak N, Winters N, Campbell JR, et al. Changes in treatment for multidrug-resistant tuberculosis according to national income. *Eur Respir J* 2020; **56**(5).
8. Dheda K, Gumbo T, Maartens G, et al. The Lancet Respiratory Medicine Commission: 2019 update: epidemiology, pathogenesis, transmission, diagnosis, and management of multidrug-resistant and incurable tuberculosis. *Lancet Respir Med* 2019; **7**(9): 820-6.
9. World Health Organization. WHO operational handbook on tuberculosis. Module 4: treatment - drug-resistant tuberculosis treatment. Geneva: World Health Organization, 2022.
10. Conradie F, Bagdasaryan TR, Borisov S, et al. Bedaquiline-Pretomanid-Linezolid Regimens for Drug-Resistant Tuberculosis. *N Engl J Med* 2022; **387**(9): 810-23.
11. Conradie F, Diacon AH, Ngubane N, et al. Treatment of Highly Drug-Resistant Pulmonary Tuberculosis. *N Engl J Med* 2020; **382**(10): 893-902.
12. Goodall RL, Meredith SK, Nunn AJ, et al. Evaluation of two short standardised regimens for the treatment of rifampicin-resistant tuberculosis (STREAM stage 2): an open-label, multicentre, randomised, non-inferiority trial. *Lancet* 2022; **400**(10366): 1858-68.
13. Nunn AJ, Phillips PPJ, Meredith SK, et al. A Trial of a Shorter Regimen for Rifampin-Resistant Tuberculosis. *N Engl J Med* 2019; **380**(13): 1201-13.
14. Barbier M, Wirth T. The Evolutionary History, Demography, and Spread of the Mycobacterium tuberculosis Complex. *Microbiol Spectr* 2016; **4**(4).
15. Daniel TM. The history of tuberculosis. *Respir Med* 2006; **100**(11): 1862-70.
16. Pai M, Behr MA, Dowdy D, et al. Tuberculosis. *Nat Rev Dis Primers* 2016; **2**: 16076.
17. Long R, Divangahi M, Schwartzman K. Chapter 2: Transmission and pathogenesis of tuberculosis. *Canadian Journal of Respiratory, Critical Care, and Sleep Medicine* 2022; **6**(sup1): 22-32.
18. Wingfield T, Tovar MA, Datta S, Saunders MJ, Evans CA. Addressing social determinants to end tuberculosis. *Lancet* 2018; **391**(10126): 1129-32.

19. Lonnroth K, Jaramillo E, Williams BG, Dye C, Raviglione M. Drivers of tuberculosis epidemics: the role of risk factors and social determinants. *Soc Sci Med* 2009; **68**(12): 2240-6.
20. Saunders MJ, Evans CA. Fighting poverty to prevent tuberculosis. *The Lancet Infectious diseases* 2016; **16**(4): 395-6.
21. Mounchili A, Perera R, Lee RS, Njoo H, Brooks J. Chapter 1: Epidemiology of tuberculosis in Canada. *Canadian Journal of Respiratory, Critical Care, and Sleep Medicine* 2022; **6**(sup1): 8-21.
22. Centers for Disease Control and Prevention. Tuberculosis: General Information. 2011. https://www.cdc.gov/tb/publications/factsheets/general/tb.pdf.
23. Behr MA, Lapierre SG, Kunimoto DY, et al. Chapter 3: Diagnosis of tuberculosis disease and drug-resistant tuberculosis. *Canadian Journal of Respiratory, Critical Care, and Sleep Medicine* 2022; **6**(sup1): 33-48.
24. World Health Organization. WHO consolidated guidelines on tuberculosis. Module 4: treatment - drug-susceptible tuberculosis treatment. Geneva: World Health Organization,, 2022.
25. Rustage K, Lobe J, Hayward SE, et al. Initiation and completion of treatment for latent tuberculosis infection in migrants globally: a systematic review and meta-analysis. *The Lancet Infectious diseases* 2021; **21**(12): 1701-12.
26. Alsdurf H, Hill PC, Matteelli A, Getahun H, Menzies D. The cascade of care in diagnosis and treatment of latent tuberculosis infection: a systematic review and meta-analysis. *The Lancet Infectious diseases* 2016; **16**(11): 1269-78.
27. Tiberi S, du Plessis N, Walzl G, et al. Tuberculosis: progress and advances in development of new drugs, treatment regimens, and host-directed therapies. *The Lancet Infectious diseases* 2018.
28. Borisov AS, Bamrah Morris S, Njie GJ, et al. Update of Recommendations for Use of Once-Weekly Isoniazid-Rifapentine Regimen to Treat Latent Mycobacterium tuberculosis Infection. *MMWR Morb Mortal Wkly Rep* 2018; **67**(25): 723-6.
29. Aibana O, Dauria E, Kiriazova T, et al. Patients' perspectives of tuberculosis treatment challenges and barriers to treatment adherence in Ukraine: a qualitative study. *BMJ Open* 2020; **10**(1): e032027.
30. Shringarpure KS, Isaakidis P, Sagili KD, Baxi RK, Das M, Daftary A. "When Treatment Is More Challenging than the Disease": A Qualitative Study of MDR-TB Patient Retention. *PloS one* 2016; **11**(3): e0150849.
31. Pradipta IS, Idrus LR, Probandari A, et al. Barriers and strategies to successful tuberculosis treatment in a high-burden tuberculosis setting: a qualitative study from the patient's perspective. *BMC Public Health* 2021; **21**(1): 1903.
32. Nigam S, Sharma RK, Yadav R, et al. Experiences and needs of patients with MDR/XDR-TB: a qualitative study among Saharia tribe in Madhya Pradesh, Central India. *BMJ Open* 2021; **11**(8): e044698.
33. Deshmukh RD, Dhande DJ, Sachdeva KS, et al. Patient and Provider Reported Reasons for Lost to Follow Up in MDRTB Treatment: A Qualitative Study from a Drug Resistant TB Centre in India. *PloS one* 2015; **10**(8): e0135802.
34. World Health Organization. WHO consolidated guidelines on tuberculosis. Module 5: management of tuberculosis in children and adolescents. Geneva: World Health Organization, 2022.

35. World Health Organization. WHO operational handbook on tuberculosis. Module 1: prevention - tuberculosis preventive treatment. Geneva, Swtizerland: World Health Organization, 2020.
36. Campbell JR, Pease C, Daley PK, Pai M, Menzies D. Chapter 4: Diagnosis of tuberculosis infection. *Canadian Journal of Respiratory, Critical Care, Sleep Medicine*

2022; **6**: 49 - 65.
37. Campbell JR, Winters N, Menzies D. Absolute risk of tuberculosis among untreated populations with a positive tuberculin skin test or interferon-gamma release assay result: systematic review and meta-analysis. *BMJ* 2020; **368**: m549.
38. Shah M, Dorman SE. Latent Tuberculosis Infection. *N Engl J Med* 2021; **385**(24): 2271-80.
39. Huebner RE, Schein MF, Bass JB, Jr. The tuberculin skin test. *Clin Infect Dis* 1993; **17**(6): 968-75.
40. Pai M, Denkinger CM, Kik SV, et al. Gamma interferon release assays for detection of Mycobacterium tuberculosis infection. *Clin Microbiol Rev* 2014; **27**(1): 3-20.
41. Andersen P, Munk ME, Pollock JM, Doherty TM. Specific immune-based diagnosis of tuberculosis. *Lancet* 2000; **356**(9235): 1099-104.
42. Alvarez GG, Pease C, Menzies D. Chapter 6: Tuberculosis preventive treatment in adults. *Canadian Journal of Respiratory, Critical Care, and Sleep Medicine* 2022; **6**(sup1): 77-86.
43. Hirsch-Moverman Y, Daftary A, Franks J, Colson PW. Adherence to treatment for latent tuberculosis infection: systematic review of studies in the US and Canada. *The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease* 2008; **12**(11): 1235-54.
44. Njie GJ, Morris SB, Woodruff RY, Moro RN, Vernon AA, Borisov AS. Isoniazid-Rifapentine for Latent Tuberculosis Infection: A Systematic Review and Meta-analysis. *American journal of preventive medicine* 2018.
45. Menzies D, Long R, Trajman A, et al. Adverse events with 4 months of rifampin therapy or 9 months of isoniazid therapy for latent tuberculosis infection: a randomized trial. *Annals of internal medicine* 2008; **149**(10): 689-97.
46. Sterling TR, Njie G, Zenner D, et al. Guidelines for the Treatment of Latent Tuberculosis Infection: Recommendations from the National Tuberculosis Controllers Association and CDC, 2020. *MMWR Recomm Rep* 2020; **69**(1): 1-11.
47. Sterling TR, Villarino ME, Borisov AS, et al. Three months of rifapentine and isoniazid for latent tuberculosis infection. *N Engl J Med* 2011; **365**(23): 2155-66.
48. Menzies D, Adjobimey M, Ruslami R, et al. Four Months of Rifampin or Nine Months of Isoniazid for Latent Tuberculosis in Adults. *N Engl J Med* 2018; **379**(5): 440-53.
49. Martinson NA, Barnes GL, Moulton LH, et al. New regimens to prevent tuberculosis in adults with HIV infection. *N Engl J Med* 2011; **365**(1): 11-20.
50. Diallo T, Adjobimey M, Ruslami R, et al. Safety and Side Effects of Rifampin versus Isoniazid in Children. *N Engl J Med* 2018; **379**(5): 454-63.
51. Belknap R, Holland D, Feng PJ, et al. Self-administered Versus Directly Observed Once-Weekly Isoniazid and Rifapentine Treatment of Latent Tuberculosis Infection: A Randomized Trial. *Annals of internal medicine* 2017; **167**(10): 689-97.

52. Villarino ME, Scott NA, Weis SE, et al. Treatment for preventing tuberculosis in children and adolescents: a randomized clinical trial of a 3-month, 12-dose regimen of a combination of rifapentine and isoniazid. *JAMA pediatrics*, 2015. (accessed.
53. Sterling TR, Scott NA, Miro JM, et al. Three months of weekly rifapentine and isoniazid for treatment of Mycobacterium tuberculosis infection in HIV-coinfected persons. *AIDS* 2016; **30**(10): 1607-15.
54. Alshammari TM. Drug safety: The concept, inception and its importance in patients' health. *Saudi Pharm J* 2016; **24**(4): 405-12.
55. Heyd A, Heffernan C, Storey K, Wild TC, Long R. Treating latent tuberculosis infection (LTBI) with isoniazid and rifapentine (3HP) in an inner-city population with psychosocial barriers to treatment adherence: A qualitative descriptive study. *PLOS Glob Public Health* 2021; **1**(12): e0000017.
56. Spence BC, Bruxvoort K, Munoz-Plaza C, et al. Patient-Reported Barriers to Treatment Initiation and Completion for Latent Tuberculosis Infection Among Patients Within a Large Integrated Health Care System in Southern California. *J Public Health Manag Pract* 2023; **29**(3): 345-52.
57. Sterling TR, Moro RN, Borisov AS, et al. Flu-like and Other Systemic Drug Reactions Among Persons Receiving Weekly Rifapentine Plus Isoniazid or Daily Isoniazid for Treatment of Latent Tuberculosis Infection in the PREVENT Tuberculosis Study. *Clin Infect Dis* 2015; **61**(4): 527-35.
58. Zenner D, Beer N, Harris RJ, Lipman MC, Stagg HR, van der Werf MJ. Treatment of Latent Tuberculosis Infection: An Updated Network Meta-analysis. *Annals of internal medicine* 2017; **167**(4): 248-55.
59. Kohler S, Sitali N, Achar J, Paul N. Programme costs of longer and shorter tuberculosis drug regimens and drug import: a modelling study for Karakalpakstan, Uzbekistan. *ERJ Open Res* 2022; **8**(1).
60. Moro RN, Borisov AS, Saukkonen J, et al. Factors Associated With Noncompletion of Latent Tuberculosis Infection Treatment: Experience From the PREVENT TB Trial in the United States and Canada. *Clin Infect Dis* 2016; **62**(11): 1390-400. doi: 10.093/cid/ciw126. Epub 2016 Mar 6.
61. Centers for Disease Control and Prevention. TB Elimination Multidrug-Resistant Tuberculosis (MDR TB). In: Centers for Disease Control and Prevention, editor.; 2012.
62. Dheda K, Gumbo T, Maartens G, et al. The epidemiology, pathogenesis, transmission, diagnosis, and management of multidrug-resistant, extensively drug-resistant, and incurable tuberculosis. *Lancet Respir Med* 2017.
63. World Health Organization. Global tuberculosis report 2021. Geneva, Switzerland: World Health Organization, 2021.
64. Bisson GP, Bastos M, Campbell JR, et al. Mortality in adults with multidrug-resistant tuberculosis and HIV by antiretroviral therapy and tuberculosis drug use: an individual patient data meta-analysis. *Lancet* 2020; **396**(10248): 402-11.
65. Campbell JR, Chan ED, Falzon D, et al. Low Body Mass Index at Treatment Initiation and Rifampicin-Resistant Tuberculosis Treatment Outcomes: An Individual Participant Data Meta-Analysis. *Clin Infect Dis* 2022.
66. World Health Organization. Maps on the diagnosis and notification of rifampicin-resistant TB. https://www.who.int/teams/global-tuberculosis-programme/data.

67. Gygli SM, Borrell S, Trauner A, Gagneux S. Antimicrobial resistance in Mycobacterium tuberculosis: mechanistic and evolutionary perspectives. *FEMS Microbiol Rev* 2017; **41**(3): 354-73.
68. Lange C, Chesov D, Heyckendorf J, Leung CC, Udwadia Z, Dheda K. Drug-resistant tuberculosis: An update on disease burden, diagnosis and treatment. *Respirology* 2018; **23**(7): 656-73.
69. Brode SK, Dwilow R, Kunimoto D, Menzies D, Khan FA. Chapter 8: Drug-resistant tuberculosis. *Canadian Journal of Respiratory, Critical Care, and Sleep Medicine* 2022; **6**(sup1): 109-28.
70. Gangadharam PR. Drug resistance in tuberculosis: it's not always the patient's fault! *Tuber Lung Dis* 1993; **74**(1): 65-7.
71. Liebenberg D, Gordhan BG, Kana BD. Drug resistant tuberculosis: Implications for transmission, diagnosis, and disease management. *Front Cell Infect Microbiol* 2022; **12**: 943545.
72. Kim SJ. Drug-susceptibility testing in tuberculosis: methods and reliability of results. *Eur Respir J* 2005; **25**(3): 564-9.
73. Yusoof KA, Garcia JI, Schami A, et al. Tuberculosis Phenotypic and Genotypic Drug Susceptibility Testing and Immunodiagnostics: A Review. *Front Immunol* 2022; **13**: 870768.
74. Almeida Da Silva PE, Palomino JC. Molecular basis and mechanisms of drug resistance in Mycobacterium tuberculosis: classical and new drugs. *J Antimicrob Chemother* 2011; **66**(7): 1417-30.
75. Lange C, Dheda K, Chesov D, Mandalakas AM, Udwadia Z, Horsburgh CR, Jr. Management of drug-resistant tuberculosis. *Lancet* 2019; **394**(10202): 953-66.
76. Lan Z, Ahmad N, Baghaei P, et al. Drug-associated adverse events in the treatment of multidrug-resistant tuberculosis: an individual patient data meta-analysis. *Lancet Respir Med* 2020; **8**(4): 383-94.
77. World Health Organization. WHO consolidated guidelines on tuberculosis. Module 4: treatment - drug-resistant tuberculosis treatment. Geneva: World Health Organization, 2020.
78. Li Y, Sun F, Zhang W. Bedaquiline and delamanid in the treatment of multidrug-resistant tuberculosis: Promising but challenging. *Drug Dev Res* 2019; **80**(1): 98-105.
79. Collaborative Group for the Meta-Analysis of Individual Patient Data in MDR-TB treatment-2017, Ahmad N, Ahuja SD, et al. Treatment correlates of successful outcomes in pulmonary multidrug-resistant tuberculosis: an individual patient data meta-analysis. *Lancet* 2018; **392**(10150): 821-34.
80. World Health Organization. Definitions and reporting framework for tuberculosis—2013 revision. Geneva: World Health Organization, 2013.
81. Laserson KF, Thorpe LE, Leimane V, et al. Speaking the same language: treatment outcome definitions for multidrug-resistant tuberculosis. *The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease* 2005; **9**(6): 640-5.
82. Ramachandran G, Swaminathan S. Safety and tolerability profile of second-line anti-tuberculosis medications. *Drug Saf* 2015; **38**(3): 253-69.
83. Stadler JAM, Maartens G, Meintjes G, Wasserman S. Clofazimine for the treatment of tuberculosis. *Front Pharmacol* 2023; **14**: 1100488.

84. Ramu G, Iyer GG. Side effects of clofazimine therapy. *Lepr India* 1976; **48**(4 Suppl): 722-31.
85. Wang MG, Liu XM, Wu SQ, He JQ. Impacts of clofazimine on the treatment outcomes of drug-resistant tuberculosis. *Microbes Infect* 2023; **25**(1-2): 105020.
86. Siddique AA, Schnitzer ME, Bahamyirou A, et al. Causal inference with multiple concurrent medications: A comparison of methods and an application in multidrug-resistant tuberculosis. *Stat Methods Med Res* 2019; **28**(12): 3534-49.
87. Rodriguez CA, Sy KTL, Mitnick CD, Franke MF. Time-Dependent Confounding in Tuberculosis Treatment Outcome Analyses: A Review of a Source of Bias. *Am J Respir Crit Care Med* 2020; **202**(9): 1311-4.
88. Franke MF, Mitnick CD. Time for a change: considering regimen changes in analyses of observational drug-resistant TB treatment cohort data. *The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease* 2020; **24**(11): 1151-5.
89. Ahmad Khan F, Salim MAH, du Cros P, et al. Effectiveness and safety of standardised shorter regimens for multidrug-resistant tuberculosis: individual patient data and aggregate data meta-analyses. *Eur Respir J* 2017; **50**(1).
90. Clare PJ, Dobbins TA, Mattick RP. Causal models adjusting for time-varying confounding-a systematic review of the literature. *Int J Epidemiol* 2019; **48**(1): 254-65.
91. Daniel RM, Cousens SN, De Stavola BL, Kenward MG, Sterne JA. Methods for dealing with time-dependent confounding. *Stat Med* 2013; **32**(9): 1584-618.
92. Mansournia MA, Etminan M, Danaei G, Kaufman JS, Collins G. Handling time varying confounding in observational research. *BMJ* 2017; **359**: j4587.
93. Shinozaki T, Suzuki E. Understanding Marginal Structural Models for Time-Varying Exposures: Pitfalls and Tips. *J Epidemiol* 2020; **30**(9): 377-89.
94. Mok J, Lee M, Kim DK, et al. 9 months of delamanid, linezolid, levofloxacin, and pyrazinamide versus conventional therapy for treatment of fluoroquinolone-sensitive multidrug-resistant tuberculosis (MDR-END): a multicentre, randomised, open-label phase 2/3 non-inferiority trial in South Korea. *Lancet* 2022; **400**(10362): 1522-30.
95. Van Deun A, Maug AK, Salim MA, et al. Short, highly effective, and inexpensive standardized treatment of multidrug-resistant tuberculosis. *Am J Respir Crit Care Med* 2010; **182**(5): 684-92.
96. Nyang'wa BT, Berry C, Kazounis E, et al. A 24-Week, All-Oral Regimen for Rifampin-Resistant Tuberculosis. *N Engl J Med* 2022; **387**(25): 2331-43.
97. Ndjeka N, Campbell JR, Meintjes G, et al. Treatment outcomes 24 months after initiating short, all-oral bedaquiline-containing or injectable-containing rifampicin-resistant tuberculosis treatment regimens in South Africa: a retrospective cohort study. *The Lancet Infectious diseases* 2022; **22**(7): 1042-51.
98. World Health Organization. Rapid communication: key changes to treatment of drug-resistant tuberculosis. . Geneva, Swizterland, 2019.
99. Keam SJ. Pretomanid: First Approval. *Drugs* 2019; **79**(16): 1797-803.
100. Tonin FS, Rotta I, Mendes AM, Pontarolo R. Network meta-analysis: a technique to gather evidence from direct and indirect comparisons. *Pharm Pract (Granada)* 2017; **15**(1): 943.

101. Watt J, Del Giovane C. Network Meta-Analysis. *Methods Mol Biol* 2022; **2345**: 187-201.
102. Jansen JP. Network meta-analysis of individual and aggregate level data. *Res Synth Methods* 2012; **3**(2): 177-90.
103. Debray TP, Schuit E, Efthimiou O, et al. An overview of methods for network meta-analysis using individual participant data: when do benefits arise? *Stat Methods Med Res* 2018; **27**(5): 1351-64.
104. Fox GJ, Benedetti A, Mitnick CD, Pai M, Menzies D, Collaborative Group for Meta-Analysis of Individual Patient Data in M-T. Propensity Score-Based Approaches to Confounding by Indication in Individual Patient Data Meta-Analysis: Non-Standardized Treatment for Multidrug Resistant Tuberculosis. *PloS one* 2016; **11**(3): e0151724.
105. Riley RD, Lambert PC, Abo-Zaid G. Meta-analysis of individual participant data: rationale, conduct, and reporting. *BMJ* 2010; **340**: c221.
106. Stewart LA, Tierney JF. To IPD or not to IPD? Advantages and disadvantages of systematic reviews using individual patient data. *Eval Health Prof* 2002; **25**(1): 76-97.
107. Leahy J, O'Leary A, Afdhal N, et al. The impact of individual patient data in a network meta-analysis: An investigation into parameter estimation and model selection. *Res Synth Methods* 2018; **9**(3): 441-69.
108. Kanters S, Karim ME, Thorlund K, Anis A, Bansback N. When does the use of individual patient data in network meta-analysis make a difference? A simulation study. *BMC Med Res Methodol* 2021; **21**(1): 21.
109. Hernan MA. A definition of causal effect for epidemiological research. *J Epidemiol Community Health* 2004; **58**(4): 265-71.
110. Rubin D. ESTIMATING CAUSAL EFFECTS OF TREATMENTS IN EXPERIMENTAL AND OBSERVATIONAL STUDIES. 1972; **1972**(2): i-31.
111. Rubin DB. Causal Inference Using Potential Outcomes. *Journal of the American Statistical Association* 2005; **100**(469): 322-31.
112. Austin PC, Stuart EA. Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. *Stat Med* 2015; **34**(28): 3661-79.
113. Hernan MA, Robins JM. Estimating causal effects from epidemiological data. *J Epidemiol Community Health* 2006; **60**(7): 578-86.
114. Lousdal ML. An introduction to instrumental variable assumptions, validation and estimation. *Emerg Themes Epidemiol* 2018; **15**: 1.
115. ROSENBAUM PR, RUBIN DB. The central role of the propensity score in observational studies for causal effects. *Biometrika* 1983; **70**(1): 41-55.
116. Desai RJ, Franklin JM. Alternative approaches for confounding adjustment in observational studies using weighting based on the propensity score: a primer for practitioners. *BMJ* 2019; **367**: l5657.
117. Gruber S, Phillips RV, Lee H, van der Laan MJ. Data-Adaptive Selection of the Propensity Score Truncation Level for Inverse-Probability-Weighted and Targeted Maximum Likelihood Estimators of Marginal Point Treatment Effects. *Am J Epidemiol* 2022; **191**(9): 1640-51.
118. Hernán MA, Robins JM. Causal Inference: What If: Boca Raton: Chapman & Hall/CRC; 2020.

119. Smith LH, Garcia-Albeniz X, Chan JM, et al. Emulation of a target trial with sustained treatment strategies: an application to prostate cancer using both inverse probability weighting and the g-formula. *Eur J Epidemiol* 2022; **37**(12): 1205-13.
120. Petersen M, Schwab J, Gruber S, Blaser N, Schomaker M, van der Laan M. Targeted Maximum Likelihood Estimation for Dynamic and Static Longitudinal Marginal Structural Working Models. *J Causal Inference* 2014; **2**(2): 147-85.
121. Schnitzer ME, Lok JJ, Bosch RJ. Double robust and efficient estimation of a prognostic model for events in the presence of dependent censoring. *Biostatistics* 2016; **17**(1): 165-77.
122. Hernan MA, Lanoy E, Costagliola D, Robins JM. Comparison of dynamic treatment regimes via inverse probability weighting. *Basic Clin Pharmacol Toxicol* 2006; **98**(3): 237-42.
123. Hernan MA, Wang W, Leaf DE. Target Trial Emulation: A Framework for Causal Inference From Observational Data. *JAMA* 2022; **328**(24): 2446-7.
124. Hernan MA, Robins JM. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. *Am J Epidemiol* 2016; **183**(8): 758-64.
125. Matthews AA, Danaei G, Islam N, Kurth T. Target trial emulation: applying principles of randomised trials to observational studies. *BMJ* 2022; **378**: e071108.
126. Hernan MA, Sauer BC, Hernandez-Diaz S, Platt R, Shrier I. Specifying a target trial prevents immortal time bias and other self-inflicted injuries in observational analyses. *J Clin Epidemiol* 2016; **79**: 70-5.
127. Moreno-Betancur M. The Target Trial: A Powerful Device Beyond Well-defined Interventions. *Epidemiology* 2021; **32**(2): 291-4.
128. Danaei G, Garcia Rodriguez LA, Cantero OF, Logan RW, Hernan MA. Electronic medical records can be used to emulate target trials of sustained treatment strategies. *J Clin Epidemiol* 2018; **96**: 12-22.
129. Rudolph JE, Benkeser D, Kennedy EH, Schisterman EF, Naimi AI. Estimation of the Average Causal Effect in Longitudinal Data With Time-Varying Exposures: The Challenge of Nonpositivity and the Impact of Model Flexibility. *Am J Epidemiol* 2022; **191**(11): 1962-9.
130. Winters N, Belknap R, Benedetti A, et al. Completion, safety, and efficacy of tuberculosis preventive treatment regimens containing rifampicin or rifapentine: an individual patient data network meta-analysis. *Lancet Respir Med* 2023; **11**(9): 782-90.
131. Winters N, Schnitzer ME, Campbell JR, et al. Identifying patients with multidrug-resistant tuberculosis who may benefit from shorter durations of treatment. *PloS one* 2023; **18**(10): e0292106.
132. Cipriani A, Higgins JP, Geddes JR, Salanti G. Conceptual and technical challenges in network meta-analysis. *Annals of internal medicine* 2013; **159**(2): 130-7.
133. Akaike H. Information Theory and an Extension of the Maximum Likelihood Principle. In: Parzen E, Tanabe K, Kitagawa G, eds. Selected Papers of Hirotugu Akaike. New York, NY: Springer New York; 1998: 199-213.
134. Schwarz G. Estimating the Dimension of a Model. 1978; **6 %J The Annals of Statistics**(2): 461-4, 4.
135. Khan U, Huerga H, Khan AJ, et al. The endTB observational study protocol: treatment of MDR-TB with bedaquiline or delamanid containing regimens. *BMC Infect Dis* 2019; **19**(1): 733.

136. Swindells S, Ramchandani R, Gupta A, et al. One Month of Rifapentine plus Isoniazid to Prevent HIV-Related Tuberculosis. *N Engl J Med* 2019; **380**(11): 1001-11.
137. Pazzagli L, Linder M, Zhang M, et al. Methods for time-varying exposure related problems in pharmacoepidemiology: An overview. *Pharmacoepidemiol Drug Saf* 2018; **27**(2): 148-60.