BMJ Open

Tuberculosis-associated respiratory impairment and disability in children, adolescents and adults: protocol for a systematic review and individual participant data meta-analysis

Silvia S. Chiang ,[1,2] Kamila Romanowski,[3,4] James C. Johnston,[3,5] Alexandre Petiquan,[6] Mayara Lisboa Bastos,[7] Dick Menzies,[8,9,10] Sierra A. Land,[6] Andrea Benedetti ,[10] Faiz Ahmad Khan,[8,9,11] Marieke M. van der Zalm,[12,13,14] Jonathon R. Campbell [8,9,15]

To cite: Chiang SS, Romanowski K, Johnston JC, *et al*. Tuberculosis-associated respiratory impairment and disability in children, adolescents and adults: protocol for a systematic review and individual participant data meta-analysis. *BMJ Open* 2025;**15**:e094118. doi:10.1136/bmjopen-2024-094118

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (https://doi.org/10.1136/bmjopen-2024-094118).

SSC, KR, MMvdZ and JRC contributed equally.

Received 23 September 2024
Accepted 07 March 2025

For numbered affiliations see end of article.

Correspondence to
Dr Jonathon R. Campbell; jonathon.campbell@mcgill.ca

ABSTRACT

Background Approximately 2% of the global population has survived tuberculosis (TB). Increasing evidence indicates that a significant proportion of pulmonary TB survivors develop TB-associated respiratory impairment and disability—commonly referred to as post-TB lung disease—marked by impaired respiratory function, persistent symptoms and activity limitations. However, the prevalence, risk factors and progression of TB-associated respiratory disability throughout the life course are not well understood. To address these gaps, we will undertake a systematic review and individual participant-level data meta-analysis (IPD-MA) focusing on TB-associated respiratory impairment and disability in children, adolescents and adults successfully treated for pulmonary TB.

Methods and analysis We will systematically search MEDLINE, Embase, CENTRAL, Global Index Medicus and medRxiv for original studies investigating TB-associated respiratory impairment and disability in people of all ages who have completed treatment for microbiologically confirmed or clinically diagnosed pulmonary TB. Authors of eligible studies will be invited to contribute deidentified data and form a collaborative group. Primary outcomes will be (1) abnormal lung function based on spirometry parameters and (2) chronic respiratory symptoms. We will estimate the overall and subgroup-specific prevalence of each outcome through IPD-MA. Next, we will develop clinical prediction tools assessing the risk of future TB-associated respiratory impairment and disability. Finally, we will use stepwise hierarchical modelling to identify epidemiological determinants of respiratory impairment and disability.

Ethics and dissemination This study has been approved by the ethics review boards at the Rhode Island Hospital (2138217-2) and the Research Institute of the McGill University Health Centre (2024-10345). Individual study authors will be required to obtain institutional approval prior to sharing data. Results will be disseminated through open-access, peer-reviewed publications and conference presentations.

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ An individual participant data meta-analysis allows for data harmonisation to help overcome limitations of individual studies and aggregate meta-analysis, including small sample size, heterogeneity and limited reporting of subgroups, such as age and other risk factors.
⇒ We will be able to identify weaknesses in current reporting and recommend standards to support high-quality data collection and facilitate pooling of data moving forward.
⇒ Key limitations include authors' willingness to share data, representativeness of data contributed and missing data.
⇒ We will build an ongoing data collection platform to allow updating of evidence.
⇒ Results will have implications for public health, clinical trial design and clinical practice to support tuberculosis survivors.

PROSPERO registration number CRD42024529906.

INTRODUCTION

Globally, over 10 million people develop tuberculosis (TB) each year,[1] and as of 2020, an estimated 155 million individuals who had TB disease in the preceding 40 years were still alive.[2] Approximately 86% of people who had TB disease had pulmonary TB,[1 3] which even after successful treatment, may lead to respiratory sequelae, including restrictive lung disease, chronic obstructive pulmonary disease and infectious pulmonary complications.[4 5] This spectrum of TB-associated respiratory disorders, commonly known as post-TB lung disease—a misnomer as the pathogenesis occurs during TB disease—encompasses

both respiratory impairment and disability.[6 7] Respiratory impairment is characterised by abnormal lung function, commonly measured through spirometry or structural assessments. Disability relates to functional impacts such as chronic respiratory symptoms and limitations or restrictions in daily activities and participation.[8]

A recent systematic review and aggregate-level meta-analysis, which included 61 original studies and over 40 000 participants, found 59% (95% CI 49%, 69%) of people who survived TB had abnormal spirometry and 25% (95% CI 19%, 32%) reported breathlessness.[9] Additional reviews have been conducted to identify risk factors for TB-associated respiratory impairment and disability, but consistent, robust associations have not been found.[10 11] It remains unclear who is at the highest risk for TB-associated respiratory impairment and disability. Consequently, guidelines recommend screening all pulmonary TB survivors for respiratory impairment and disability,[12–14] a practice that is not feasible in the resource-limited settings where the majority of TB survivors reside.[1 2] An improved understanding of risk factors and epidemiological determinants of TB-associated respiratory impairment and disability is crucial to enhancing the efficiency and feasibility of screening.

Another significant limitation in the existing evidence is the underrepresentation of children and adolescents in studies of TB-associated respiratory impairment and disability, along with a lack of age-disaggregated outcomes.[15] Research across all age groups is essential due to the differences between children, adolescents and adults in the host response to *Mycobacterium tuberculosis*, clinical presentation of pulmonary TB and ongoing lung development in children and adolescents.[6]

Individual participant data meta-analysis (IPD-MA) can overcome many limitations of aggregate methods through harmonisation of variables and outcomes, standardised analysis, analysis of subgroups not reported in primary studies such as age-stratified results and increased sample size.[16] Using IPD-MA, we can identify subgroups at high risk for TB-associated respiratory impairment and disability, supporting risk stratification of individuals for screening. Additionally, this approach allows us to assess the underlying epidemiological determinants of TB-associated respiratory impairment and disability. With a large body of literature now available, an IPD-MA is an attractive approach to improve the targeting and feasibility of screening strategies for TB-associated respiratory impairment and disability.

The overall aim of this systematic review and IPD-MA is to estimate the burden, risk, and determinants of TB-associated respiratory impairment and disability in children, adolescents and adults treated for pulmonary TB. We will address the following specific objectives. First, we will estimate the prevalence of the two components of TB-associated respiratory impairment and disability: (1) impairment to respiratory function and structure and (2) symptoms and functional limitations. As part of this aim, we will estimate prevalence among subgroups at different times since TB diagnosis and examine risk factors. Second, we will develop and validate risk prediction models to predict future TB-associated respiratory (1) impairment and (2) disability. Third, we will identify the epidemiological and clinical determinants of (1) respiratory impairment and (2) disability, using a conceptual framework and hierarchical approach.

METHODS AND ANALYSIS

The systematic review and IPD-MA will be reported according to the 2020 PRISMA guidelines.[17] This protocol has been reported according to the PRISMA-P guidelines[18] and has been prospectively registered with PROSPERO (CRD42024529906). Any key changes or amendments will be documented there.

Literature search and selection criteria

In preparation for this project, we conducted an initial literature search of MEDLINE, Embase, CENTRAL, Global Index Medicus and medRxiv for original studies published between 1 January 2004 and 26 April 2024, using a comprehensive search strategy (online supplemental appendix) developed in collaboration with a medical librarian experienced in systematic reviews. We selected 2004 as the earliest publication year as there is minimal chance of individual participant-level study data being held for over 20 years. Two reviewers (KR and SAL) screened titles, abstracts, full texts and any studies identified as relevant from reviews and reference lists of eligible articles. Disagreements regarding inclusion or exclusion were resolved by a third reviewer (JRC). Retrieved references were uploaded into Zotero (Center for History and New Media, George Mason University), a reference management software, for deduplication and then subsequently imported into Covidence (Veritas Health Innovation, Australia), a web-based platform designed to streamline the systematic review process. We will update this search through March 2025 to identify new studies published since our initial search.

Our study population will include children (0–9 years old), adolescents (10–19 years old) and adults (≥20 years old) with pulmonary TB that is microbiologically confirmed—through smear microscopy, mycobacterial culture and/or molecular assays including GeneXpert—or clinically diagnosed, and in the case of children, meeting established clinical criteria.[19] Wherever possible, our comparator population will include children, adolescents or adults without a history of pulmonary TB. In the absence of a formal comparison group, the comparator will be assumed (eg, for spirometry and other lung function measurements, international reference values for age, sex and height will be used as the comparator).

Studies will be eligible for inclusion if they meet all of the following criteria: (1) prospective or retrospective cohorts, cross-sectional studies or clinical trials, (2) included ≥10 participants who completed treatment for pulmonary TB, (3) measure and report at least one

Table 1 Outcome measures of tuberculosis-associated respiratory impairment and disability

Domain	Outcome measure(s)
Impairments in respiratory function and structure	**Primary outcome**
	Abnormal lung function, based on prebronchodilator spirometry parameters, defined as FEV1, FVC and FEV1/FVC and classified based on specific disease patterns, including obstructive, restrictive or mixed pattern.
	Secondary outcome(s)
	(1) Postbronchodilator spirometry, (2) prebronchodilator and postbronchodilator oscillometry, (3) measures of lung volume, (4) diffusion capacity measurements, (5) other tidal breathing techniques or (6) presence of bronchiectasis.
Disability	**Primary outcome**
	Chronic respiratory symptoms, defined as experiencing at least one of the following symptoms ≥2 days per week: cough, sputum production, wheeze, dyspnoea and/or chest pain, measured using symptom assessment or validated questionnaire (eg, SGRQ)
	Secondary outcome(s)
	(1) Grades 3–5 on MRC Dyspnoea Scales, (2) grades 2–4 on the modified MRC Dyspnoea Scale, (3) values higher than 'moderate breathlessness' on Borg Dyspnoea Scale or (4) 6MWT result below predicted lower limit of normal, calculated using standard equations based on age, sex, height and weight.

FEV1, forced expiratory volume in 1 s; FVC, functional vital capacity; MRC, Medical Research Council; 6MWT, 6 min Walk Test; SGRQ, St. George's Respiratory Questionnaire.

aspect of TB-associated respiratory disability, as detailed below and (4) evaluate ≥80% of all participants for these outcomes to minimise selection bias associated with selective testing. We will include studies written in any language.

Outcome measures

Our primary outcome will encompass the two aspects of TB-associated respiratory impairment and disability: (1) impairments to respiratory function and structure and (2) chronic respiratory symptoms and/or limitations in daily activities and participation. These outcomes are based on the terminology used by the WHO International Classification of Functioning, Disability and Health, a classification system that provides a standard language and framework for describing health and health-related states.[8]

Our primary outcome for impairments to respiratory function and structure will be abnormal lung function based on prebronchodilator spirometry parameters (table 1). As secondary outcomes, we will also consider other impairments to respiratory function and structure, including (1) postbronchodilator spirometry, (2) prebronchodilator and postbronchodilator oscillometry, (3) measures of lung volume, (4) diffusion capacity measurements, (5) other tidal breathing techniques and (6) the presence of bronchiectasis.

Our primary outcome for disability will be chronic respiratory symptoms (table 1). As secondary outcomes, we will also consider: (1) grades 3–5 on the Medical Research Council (MRC) Dyspnoea Scales,[20] (2) grades 2–4 on the modified MRC Dyspnoea Scale,[21] (3) values higher than 'moderate breathlessness' on the Borg Dyspnoea Scale[22] and (4) 6 min Walk Test results below predicted lower limit of normal, calculated using standard equations based on age, sex, height and weight.[23]

Invitations to authors of eligible studies

Corresponding authors of eligible studies will be contacted via email and invited to join the collaborative group and share their deidentified study data. If there is no response from the author within 4 weeks, we will try a second time. If the author does not respond, we will attempt to contact other authors. If the author(s) do not respond or indicate that the data are unavailable or cannot be shared due to access restrictions, we will note both of these occurrences.

After accepting the invitation to collaborate, signing a data transfer agreement and obtaining institutional approvals, authors will transfer their data securely via a mutually agreed on secure data-sharing platform. Data will be housed on a secure server at The Research Institute of the McGill University Health Centre.

Data collection, processing and management

From each eligible study, we will collect study-level and individual-level variables. Study-level variables will include country, funding source, country-level health characteristics, study design, population, aims, recruitment period and test(s) used to measure outcomes. Individual-level variables will include demographic, clinical, radiographic, microbiologic and outcome data (table 2). All received study data will be reviewed for missing, incomplete, or implausible data and compared against published information; authors will be further consulted for clarifications. As age and outcome data are critical pieces of information for this review, prior to processing, we will exclude all participants with missing information on age or those without an outcome measure of interest.

Table 2 Individual-level variables to be requested from study authors

Category	Variable	Proposed stratification, *if relevant*
Demographic	Age (years) at TB treatment initiation	<5, 5–9, 10–14, 15–19, 20–35, 36–54, 55–75, >75
	Sex	Male, female
	Gender	Male, female, other
	Height	For children, we will consider various measures for height and weight, such as weight-for-age or height-for-age z-scores
	Weight	
	BMI (kg/m^2)	<18.5, 18.5–24.9, 25–29.9, >30
	Active smoking exposure*	Never, former, every day,someday*
	Smoking duration (years), for those who are every day, someday, and former smokers	<1, 1–5, 6–10, 11–19, >20
	Antenatal passive smoking exposure†	Yes, No
	Postnatal passive smoking exposure†	Yes, No
	Environmental biomass exposure, defined as cooking or heating living areas with solid fuel	Yes, No
	Adult education level	Less than high school, high school, university, postgraduate
	Area of residence	Urban, rural
Clinical, biological, radiographic and treatment	HIV status	HIV on ART, HIV without ART, HIV-negative
	CD4 count, for those living with HIV	<350, +350
	Pre-existing asthma	Yes, No
	Pre-existing COPD	Yes, No
	Pre-existing other respiratory comorbidities	Yes, No
	Diabetes	Yes, No
	Previous TB treatment episodes	Yes, No
	Treatment delay	Yes, No
	Diagnostic certainty of TB diagnosis	Clinical diagnosis, microbiological diagnosis
	Drug resistance pattern of *Mycobacterium tuberculosis*	Drug-susceptible, isoniazid-monoresistant, rifampicin/multidrug resistant, extensively drug-resistant
	TB disease severity at diagnosis‡	Severe disease, non-severe disease
	TB treatment duration	<7 months, +7 months
Outcome measures	All outcome measures described in table 1	
	Months between TB treatment end and outcome measure	During treatment, <6, 6–12, 12–24, >24

*An every day smoker is defined as an adult who has smoked 100 cigarettes in his or her lifetime and who currently smokes cigarettes, while someday smoker is defined as an adult who has smoked at least 100 cigarettes in his or her lifetime, who smokes now but does not smoke every day, as defined by the United States Centers for Disease Control and Prevention National Centre for Health Statistics.
†Obtained for paediatric and adolescent populations
‡For children, TB disease severity will be assessed according to World Health Organization criteria, based on bacteriological burden and/or radiographic findings. Non-severe pulmonary disease in children will be defined as intrathoracic lymph node TB without airway obstruction; uncomplicated TB pleural effusion or paucibacillary, non-cavitary disease confined to one lobe of the lungs and without a miliary pattern. For adults, those acid fast bacilli smear positive or with cavitation on chest X-ray will be considered to have severe disease.
ART, antiretroviral treatment; BMI, body mass index; COPD, chronic obstructive pulmonary disease ; TB, tuberculosis.

We will standardise outcomes and covariates between studies using systematic harmonisation methodology.[24] Study-specific data items will be processed into a common format for analysis.

We will quantify individual-level missing data in each study. We will impute missing data using multiple imputations via chained equations (R package mice) while respecting participant clustering by source study.[25] We will generate 20 imputed datasets, each undergoing 25 between-imputation iterations. Analyses will be done in each of the 20 imputed datasets, with estimates pooled according to Rubin's rules.[26] No imputations will be done for outcome measures.

Table 3 Risk of bias criteria

Domain	Question(s)
Risk of bias due to confounding Key confounders: environmental exposures, smoking exposure, age, weight, sex, HIV status and pre-existing respiratory comorbidities.	Did the author(s) consider all important confounders in the design of the study (ie, matching)?
	Did the author(s) consider all important confounders in the analysis (ie, stratified analyses, multivariate models)?
	Did the author(s) collect information on all important confounders?
	Were confounding factors measured validly and reliably (ie, secure record)?
Risk of bias arising from measurement of the exposure	Was TB diagnosis based on standardised definitions for diagnostic classification?
	Were results presented stratified by microbiological confirmation and clinical diagnosis?
	Was TB treatment data ascertained validly and reliably (ie, secure record)?
Risk of bias in selection of participants into the study	Was the TB-exposed cohort selected randomly or using a consecutive sample?
	Was the comparator population selected from the same community as the TB population?
Risk of bias due to postexposure interventions Postexposure interventions: pulmonary rehab, steroids, surgery	Did the study report if participants received any postexposure interventions that may have decreased the sensitivity of the outcome?
	Did the analyses attempt to correct for the effect of any postexposure interventions?
Risk of bias due to missing data	Were complete data on exposure status available for all, or nearly all (>90%), participants?
	Were complete data on outcome status available for all, or nearly all (>90%), participants?
	Were complete data on collected confounders available for all, or nearly all (>90%), participants?
Risk of bias arising from measurement of the outcome	Was TB-associated respiratory disability measured using standardised methods (ie, validated tools, according to guidelines)?
	Did the timing of the outcome vary significantly among participants?

TB, tuberculosis.

Risk of bias assessment

Two study investigators will independently assess the risk of bias in included studies using an adapted version of the ROBINS-E tool (table 3).[27] We will evaluate the following criteria at the study level: (1) selection of participants, (2) measurement of exposure, (3) confounding, (4) post-exposure interventions, (5) measurement of outcome and (6) missing data. Any disagreements will be resolved through discussion or consultation with a third study investigator. For each subdomain, we will assign a risk of bias of low, medium, high, very high or uncertain. As no detailed guidance has been developed on providing an overall risk of bias for each study, we will not give an overall risk of bias but rather discuss the potential impacts of identified sources of bias on the interpretation of our findings.

Grading the strength of existing evidence

We will assess the existing evidence using categories in the Cochrane Handbook, which considers inconsistency, indirectness, imprecision and bias.[28] Bias will be assessed for each individual study using the risk of bias assessment described above, and globally we will assess publication bias visually using Egger plots and forest plots for our two primary outcomes. Inconsistency will be assessed according to the similarity in the magnitude and direction of effects across studies; we will use stratified analyses to evaluate potential sources of heterogeneity. Indirectness will be evaluated based on applicability (tests or evaluations performed, time frame) and participant selection. Imprecision will be measured according to precision/CIs around estimates while considering sources of heterogeneity. Based on the above considerations, the strength of

the existing evidence will be graded as high (no concerns with any of the above considerations), moderate, low or very low (based on the number of concerns). In line with Cochrane, we may adjust certainty based on other factors, such as large effects, observing a dose response or the presence of plausible confounding.[28]

Statistical analysis

Unless otherwise specified, we will conduct analyses on the imputed datasets as our primary analysis and complete case analysis as a confirmatory secondary analysis. We will consider stratified analyses on the variables included in table 2, as possible. All analyses will be done using R (The R Project for Statistical Computing, the latest version available at the time of analysis start).

Objective 1

We will estimate the overall and subgroup-specific prevalence of TB-associated respiratory impairment and disability in two stages,[29] conducting such analysis among participants grouped according to timing of outcome measurement. The timing of outcome measurements analysed will be dictated by the follow-up time points among data contributed to the IPD. Potential subgroups for analysis, described in table 2, have been selected based on evidence from previous reviews, expert opinion of the investigators and biological plausibility. In the first stage, we will use the IPD to estimate the proportion of people with respiratory impairment and disability and the SE within each individual study. In the second stage, we will pool the logit-transformed proportions of impairment and disability across studies using generalised linear mixed models. We will back-transform the pooled estimates and SEs to obtain prevalence estimates and 95% CIs and generate forest plots to compare prevalence estimates across strata.[30 31] We will use I^2, τ^2 and prediction intervals to describe between-study heterogeneity. Subgroups will only be analysed if at least two studies provided information allowing such subgroup analysis.

Next, we will use one-stage IPD-MA to estimate the prevalence ratio for the outcomes of (1) respiratory impairment and (2) disability across strata. We will use generalised linear mixed log-binomial models with the variables to include in these models as covariates or effect modifiers selected from table 2. We will consider variables in two ways: variables that would only be known at TB treatment initiation and variables that would be known at the end of TB treatment. Further, we will conduct an exploratory analysis which includes the time point of outcome measurement to see how outcomes change over time; we will only include time points for which at least two studies have a measurement. In all analyses, we will account for clustering at the study and participant level. Continuous variables will be included in models using flexible forms, such as splines. Heterogeneity will be assessed with I^2 and prediction intervals.[32]

Based on findings from a previous aggregate meta-analysis, we estimated the sample size required to estimate a prevalence of 59% for respiratory impairment and 25% for disability, with an estimated average cluster size of 140.[9] Under the assumption that the intraclass correlation coefficient between studies was 0.02,[33 34] with a power of 80%, type I error rate of 5% and absolute precision of 10%, we will require a sample size of 352 for our outcome of respiratory impairment and 273 for our outcome of disability. This suggests we will likely have sufficient power for our primary outcomes and for several key subgroups. For subgroups where we do not have sufficient power, these will be classified and interpreted as exploratory.

Objective 2

We will develop risk prediction models adhering to Transparent Reporting of a Multivariate Model for Individual Prognosis or Diagnosis[35] guidance to predict the future risk of TB-associated respiratory impairment and disability. Our first model will aim to predict future TB-associated respiratory impairment and disability at the end of treatment based on demographic and clinical information known at the start of TB treatment. If the data permit, we will also look at time points further in the future and develop two other models: (1) predicting future risk from the end of TB treatment among those who do not already have evidence of TB-associated respiratory impairment and/or disability and (2) predicting persistence from the end of TB treatment among those with evidence of TB-associated respiratory impairment and/or disability. These latter two outcomes are important as existing data suggest up to one-quarter of people who go on to develop TB-associated respiratory impairment and disability are asymptomatic at the end of treatment,[36] and it is of interest to determine if there are people with impairment and/or disability at the end of treatment who resolve over time.

For our prediction models, we will use a sequential method to select candidate predictors. First, we will use a priori selection, based on consensus among investigators, published data and expert opinion, to sequentially include the chosen predictors in a layered fashion.[37] Next, we will use elastic net penalised regression models to identify key predictors from among the larger set of variables available to evaluate if our a priori predictors were missing highly predictive variables.[38] All models will include random effects for study, which will be considered when generating predictions. We will evaluate predictive performance using discrimination and calibration measures and validate the models using the internal-external cross-validation framework.[39]

To estimate the required sample size for risk prediction models, we used a conservative predictive performance estimate of a previous model (0.71)[40] and estimated sample size with 10 predictors, intraclass correlation coefficient of 0.02 and 140 average cluster size. If the proportion of survivors with disability is 25%, we require a sample size of 3134 to develop a prediction model; if the prevalence of respiratory impairment is 59%, we require a sample size of 2435.[41]

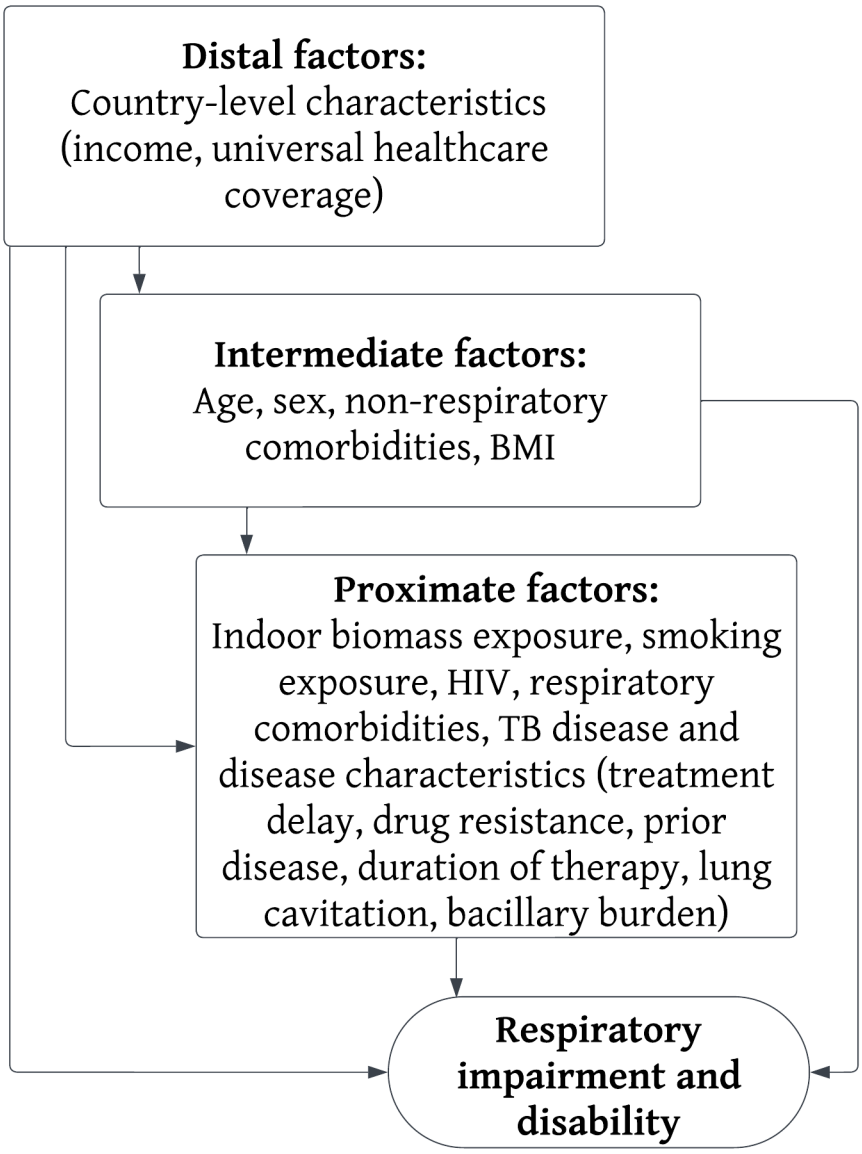

Figure 1 Epidemiological determinants of tuberculosis (TB)-associated respiratory impairment and disability. BMI, body mass index.

Objective 3

To improve our understanding of the contribution of TB to respiratory impairment and disability, as opposed to other sociodemographic, clinical and behavioural determinants, this specific aim will include only studies with a comparator group of people who never had TB disease. We will use a stepwise hierarchical modelling approach, which can systematically delineate confounding and mediating factors from our outcome of respiratory impairment and disability.[42] Our modelling approach will follow the conceptual framework in figure 1, with factors organised according to how proximate each factor is to our outcome of respiratory impairment and disability.

We will use generalised linear log-binomial models with random effects for each study. Models will be structured by first including distal determinants and subsequently including intermediate and proximate determinants in a hierarchical fashion. In this framework, the impact of TB on respiratory impairment and disability is mediated by disease-related factors such as bacillary burden and length of treatment. In our final, fully adjusted model, we will estimate the effect of TB on respiratory disability and impairment while accounting for confounders at each hierarchical level, ensuring that adjustments do not control for variables that may act as mediators. By structuring our analysis this way, we aim to isolate the independent effect of TB but acknowledge that some confounders or mediators

may be underreported or unavailable. To assess the potential impact of unmeasured confounding, we will calculate E-values to quantify the strength of association that an unmeasured confounder would need to have with both the exposure and outcome to explain away the observed effect.

Methodological challenges and limitations

A major challenge for any IPD-MA is mapping and harmonising variables due to inconsistencies in data collection methods, variable definitions and missing information across studies. We will partner with Maelstrom Research, a group with recognised expertise in data harmonisation, for rigorous data preprocessing and harmonisation techniques.[24] Given the heterogeneity in data collection across studies in all fields, our work on data harmonisation will also help us make recommendations to improve the uniformity, accuracy and completeness of data collection for future studies.

Authors' willingness to share data is another barrier, and data for certain regions or subgroups may be challenging to obtain.[43] We have budgeted resources to facilitate data sharing for authors from all settings and have established guidelines in our data-sharing agreements to promote transparency and collaboration. We will also establish a data repository, allowing us to include more data as it becomes available. Authors contributing data likely come from large centres, with adequate resources for monitoring and evaluation post-treatment. This may not be generalisable to lower-resource settings; however, we have designed our risk stratification analysis to provide evidence to programmes of all resource levels. Bias related to participant selection, data collection or reporting may affect the validity of our analyses. We are using a robust bias assessment tool to structure our bias and study quality assessment and have implemented inclusion criteria to minimise participant selection. Finally, paediatric subgroup analyses may be limited by small sample size. The inclusion of children and adolescents in this IPD-MA supports advocacy efforts for more data in this group and allows for the inclusion of currently underway studies.

Our analytical approach also has limitations. When analysing prevalence of respiratory impairment or disability, we are limited by the data available, time points measured and heterogeneity among the included studies. Despite this, we will be evaluating several subgroups which may increase our type I error rate, and so will interpret and report all results with caution. While we propose various approaches to assess and explore characteristics associated with respiratory impairment or disability, it is unlikely all important confounders will be measured at each follow-up, such as smoking, which poses a challenge for time-varying confounders. In addition, while we propose several approaches to explore heterogeneous effects, including subgroup analysis and effect modification, it is unlikely all sources of heterogeneity will be measured as it is unlikely heterogeneity is driven by single, observable characteristics. To address this, we will consider a sensitivity analysis which evaluates heterogeneous treatment effects using Bayesian latent class models to better understand heterogeneity.[44]

Patient and public involvement

No direct patient or public involvement has taken place during the development of this protocol. However, we will work with the Community Advisory Boards of the McGill TB Centre (Canada) and the Desmond Tutu TB Centre (South Africa), as well as TB survivor networks and advocacy groups such as STOP TB USA, STOP TB Canada, TB Proof, and the Child and Adolescent Working Group of the WHO/Stop TB Partnership during the interpretation and dissemination of the study results.

Ethics and dissemination

The IPD-MA was approved by institutional ethics review boards at Rhode Island Hospital, USA (2138217-2) and The Research Institute of the McGill University Health Centre, Canada (2024-10345). Individual studies will share deidentified data and follow institutional and national guidelines for data sharing.

All contributing authors will sign data-sharing agreements with the institution hosting the secure data server (The Research Institute of the McGill University Health Centre), which outlines standards of data security, management and ownership. The initial length of this data-sharing agreement will be five years. All data will be treated as confidential and will remain the property of the contributing institution; data can be withdrawn at any time. All contributing authors will enter a consortium, which collaboratively completes the outlined analyses. JRC will act as the data custodian and will be the primary point of contact between consortium members.

An oversight committee, consisting of seven members, will be established with the data custodian (JRC) acting as an additional member in a non-voting role. For the first four years, this committee will comprise the other three study principal investigators (SSC, JCJ and MvdZ) and four elected members of the consortium; thereafter, all oversight committee members will be elected. As the data custodian, JRC will remain on the oversight committee in an unelected position with a non-voting role. The oversight committee has three specific roles: (1) review requests for individuals or organisations to have access to the IPD for the purposes of agreed on analyses; (2) review and discuss proposals by members of the consortium for projects or analyses not outlined in this protocol and (3) review opportunities and requests for authorship and/or participation in analyses to ensure contribution and opportunity is equitable among the consortium. Issues deemed significant by the oversight committee may be brought forth to all consortium members for input.

Given that approximately 1 in 50 people globally has survived pulmonary TB and the growing demand for evidence to inform post-TB care,[2] this IPD-MA has the potential to significantly impact recommendations surrounding TB-associated respiratory impairment and

disability. We have explicit plans to make the IPD database we establish available to interested research groups and organisations, such as the WHO, to advance research in this field and permit analyses beyond those initially proposed in this protocol. As described in the Patient and public involvement section, we will engage various groups, including TB survivors, advocates, researchers and civil society when planning and disseminating our findings. We plan to publish our results in peer-reviewed academic journals, present our research at conferences and give seminars to stakeholders. As our results can support clinical trial design and recruitment for interventions to prevent or mitigate TB-associated respiratory consequences, we will disseminate our findings within global TB clinical trial networks. To reach TB-affected communities, we will create non-technical summary reports and infographics in multiple languages. Finally, we will create a user-friendly multilingual website to host the risk prediction model(s) we develop.

In summary, we will conduct an IPD-MA of TB-associated respiratory impairment and disability among TB survivors of all ages to determine who is most at risk, help predict those who might benefit from screening and improve our understanding of TB's contribution to respiratory impairment and disability. The insights gained from these analyses may enhance strategies for detecting and preventing TB-associated respiratory impairment and disability and inform the design of clinical trials of interventions.

Author affiliations
[1]Department of Pediatrics, Brown University, Providence, Rhode Island, USA
[2]Center for International Health Research, Rhode Island Hospital, Providence, Rhode Island, USA
[3]BC Centre for Disease Control, Vancouver, British Columbia, Canada
[4]Department of Global and Public Health, McGill University, Montreal, Vancouver, Canada
[5]Faculty of Medicine, University of British Columbia, Vancouver, BC, Canada
[6]University of Ottawa, Ottawa, Ontario, Canada
[7]Department of Family Medicine, University of Manitoba, Winnipeg, Manitoba, Canada
[8]McGill International TB Centre, Montreal, Quebec, Canada
[9]Respiratory Epidemiology and Clinical Research Unit, Centre for Outcomes Research and Evaluation, Research Institute of the McGill University Health Centre, Montreal, Quebec, Canada
[10]Departments of Medicine and of Epidemiology, Biostatistics & Occupational Health, McGill University, Montreal, Quebec, Canada
[11]Department of Medicine, McGill University, Montreal, QC, Canada
[12]Department of Paediatrics and Child Health, Faculty of Medicine and Health Sciences, Stellenbosch University, Desmond Tutu TB Centre, Cape Town, Western Cape province, South Africa
[13]Desmond Tutu TB Centre, Cape Town, South Africa
[14]Department of Paediatrics and Child Health, Stellenbosch University, Stellenbosch, South Africa
[15]Departments of Medicine & Global and Public Health, McGill University, Montreal, QC, Canada

X Silvia S. Chiang @SilviaChiangMD

Contributors Conception and design: SSC, KR, JCJ, MvdZ and JRC. Intellectual contribution and interpretation: all authors. First draft: SSC, MvdZ, KR and JRC. Revisions: all authors. All authors read and approved the final manuscript. JRC is the guarantor.

Funding This project has received funding from the Robert E. Leet & Clara Guthrie Patterson Trust (USA, PI, SSC) and the Canadian Institutes of Health Research (Canada, PI, JRC; PJT-195781). MvdZ is supported by a career development grant from the EDCTP2 programprogramme supported by the European Union (TMA2019SFP-2836 TB lung-FACT2), the Fogarty International Centre of the National Institutes of Health (NIH) under Award Number K43TW011028, and a researcher-initiated grant from the South African Medical Research Council. KR is supported by a Fonds de Recherche du Quebec—Sante postdoctoral fellowship. JRC receives salary support from the McGill University Health Centre Foundation, the McGill University Department of Medicine, and holds a Chercheur-boursier award from the Fonds de recherche du Québec—Santé (#330287).

Disclaimer The funders had no role in developing this protocol.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

ORCID iDs
Silvia S. Chiang http://orcid.org/0000-0002-2318-3802
Andrea Benedetti http://orcid.org/0000-0002-8314-9497
Jonathon R. Campbell http://orcid.org/0000-0003-2341-2166

REFERENCES

1 World Health Organization. Global tuberculosis report 2023. 2024. Available: https://www.who.int/publications-detail-redirect/9789240083851
2 Dodd PJ, Yuen CM, Jayasooriya SM, *et al*. Quantifying the global number of tuberculosis survivors: a modelling study. *Lancet Infect Dis* 2021;21:984–92.
3 Rolo M, González-Blanco B, Reyes CA, *et al*. Epidemiology and factors associated with Extra-pulmonary tuberculosis in a Low-prevalence area. *J Clin Tuberc Other Mycobact Dis* 2023;32:100377.
4 Byrne AL, Marais BJ, Mitnick CD, *et al*. Tuberculosis and chronic respiratory disease: a systematic review. *Int J Infect Dis* 2015;32:138–46.
5 Hsu D, Irfan M, Jabeen K, *et al*. Post tuberculosis treatment infectious complications. *Int J Infect Dis* 2020;92S:S41–5.
6 Allwood BW, Byrne A, Meghji J, *et al*. Post-Tuberculosis Lung Disease: Clinical Review of an Under-Recognised Global Challenge. *Respiration* 2021;100:751–63.
7 Allwood BW, van der Zalm MM, Amaral AFS, *et al*. Post-tuberculosis lung health: perspectives from the First International Symposium. *Int J Tuberc Lung Dis* 2020;24:820–8.
8 World Health Organization. *How to Use the ICF: A Practical Manual for Using the International Classification of Functioning, Disability and Health (ICF)*. 2013.
9 Taylor J, Bastos ML, Lachapelle-Chisholm S, *et al*. Residual respiratory disability after successful treatment of pulmonary tuberculosis: a systematic review and meta-analysis. *eClinicalMedicine* 2023;59:101979.
10 Ivanova O, Hoffmann VS, Lange C, *et al*. Post-tuberculosis lung impairment: systematic review and meta-analysis of spirometry data from 14 621 people. *Eur Respir Rev* 2023;32:220221.
11 Maleche-Obimbo E, Odhiambo MA, Njeri L, *et al*. Magnitude and factors associated with post-tuberculosis lung disease in low- and middle-income countries: A systematic review and meta-analysis. *PLOS Glob Public Health* 2022;2:e0000805.

12 Johnston JC, Cooper R, Menzies D. Chapter 5: Treatment of tuberculosis disease. Can. J. Respir. Crit. Care Sleep Med. 2022;6:66–76.
13 Nightingale R, Carlin F, Meghji J, *et al*. Post-TB health and wellbeing. *Int J Tuberc Lung Dis* 2023;27:248–83.
14 Migliori GB, Marx FM, Ambrosino N, *et al*. Clinical standards for the assessment, management and rehabilitation of post-TB lung disease. *int j tuberc lung dis* 2021;25:797–813.
15 Nkereuwem E, van der Zalm MM, Kampmann B, *et al*. 'Yes! We can end TB,' but remember the sequelae in children. *Lancet Respir Med* 2024;12:348–50.
16 Nevitt SJ, Tudur Smith C. *Practical Considerations and Challenges When Conducting an Individual Participant Data (IPD)Individual Participant Data (IPD) Meta-Analysis*. New York, NY: Springer US, 2022:263–78.
17 Page MJ, McKenzie JE, Bossuyt PM, *et al*. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021;372:n71n71.
18 Moher D, Shamseer L, Clarke M, *et al*. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev* 2015;4:1.
19 Graham SM, Cuevas LE, Jean-Philippe P, *et al*. Clinical Case Definitions for Classification of Intrathoracic Tuberculosis in Children: An Update. *Clin Infect Dis* 2015;61Suppl 3:S179–87.
20 Bestall JC, Paul EA, Garrod R, *et al*. Usefulness of the Medical Research Council (MRC) dyspnoea scale as a measure of disability in patients with chronic obstructive pulmonary disease. *Thorax* 1999;54:581–6.
21 Munari AB, Gulart AA, Dos Santos K, *et al*. Modified Medical Research Council Dyspnea Scale in GOLD Classification Better Reflects Physical Activities of Daily Living. *Respir Care* 2018;63:77–85.
22 Borg G. Perceived exertion as an indicator of somatic stress. *Scand J Rehabil Med* 1970;2:92–8.
23 Enright PL, Sherrill DL. Reference equations for the six-minute walk in healthy adults. *Am J Respir Crit Care Med* 1998;158:1384–7.
24 Fortier I, Raina P, Van den Heuvel ER, *et al*. Maelstrom Research guidelines for rigorous retrospective data harmonization. *Int J Epidemiol* 2017;46:103–5.
25 Audigier V, White IR, Jolani S, *et al*. Multiple Imputation for Multilevel Data with Continuous and Binary Variables. *Statist Sci* 2018;33:160–83.
26 Royston P. Multiple Imputation of Missing Values. *The Stata Journal: Promoting communications on statistics and Stata* 2004;4:227–41.
27 Bero L, Chartres N, Diong J, *et al*. The risk of bias in observational studies of exposures (ROBINS-E) tool: concerns arising from application to observational studies of exposures. *Syst Rev* 2018;7:242.
28 Schünemann HJ, Higgins JP, Vist GE, *et al*. Completing 'summary of findings' tables and grading the certainty of the evidence. In: *Cochrane Handbook for Systematic Reviews of Interventions. John Wiley & Sons*. John Wiley & Sons, Ltd, 2019: 375–402.
29 Burke DL, Ensor J, Riley RD. Meta-analysis using individual participant data: one-stage and two-stage approaches, and why they may differ. *Stat Med* 2017;36:855–75.
30 Veroniki AA, Jackson D, Bender R, *et al*. Methods to calculate uncertainty in the estimated overall effect size from a random-effects meta-analysis. *Res Synth Methods* 2019;10:23–43.
31 Veroniki AA, Jackson D, Viechtbauer W, *et al*. Methods to estimate the between-study variance and its uncertainty in meta-analysis. *Res Synth Methods* 2016;7:55–79.
32 Chen B, Benedetti A. Quantifying heterogeneity in individual participant data meta-analysis with binary outcomes. *Syst Rev* 2017;6:243.
33 Killip S, Mahfoud Z, Pearce K. What is an intracluster correlation coefficient? Crucial concepts for primary care researchers. *Ann Fam Med* 2004;2:204–8.
34 Kerry SM, Bland JM. The intracluster correlation coefficient in cluster randomisation. *BMJ* 1998;316:1455.
35 Collins GS, Reitsma JB, Altman DG, *et al*. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. The TRIPOD Group. *Circulation* 2015;131:211–9.
36 Meghji J, Lesosky M, Joekes E, *et al*. Patient outcomes associated with post-tuberculosis lung damage in Malawi: a prospective cohort study. *Thorax* 2020;75:269–78.
37 Austin PC, Tu JV. Automated variable selection methods for logistic regression produced unstable models for predicting acute myocardial infarction mortality. *J Clin Epidemiol* 2004;57:1138–46.
38 Pavlou M, Ambler G, Seaman S, *et al*. Review and evaluation of penalised regression methods for risk prediction in low-dimensional data with few events. *Stat Med* 2016;35:1159–77.
39 Debray TPA, Moons KGM, Ahmed I, *et al*. A framework for developing, implementing, and evaluating clinical prediction models in an individual participant data meta-analysis. *Stat Med* 2013;32:3158–80.
40 Meghji J, Gunsaru V, Chinoko B, *et al*. Screening for post-TB lung disease at TB treatment completion: Are symptoms sufficient? *PLOS Glob Public Health* 2024;4:e0002659.
41 Riley RD, Ensor J, Snell KIE, *et al*. Calculating the sample size required for developing a clinical prediction model. *BMJ* 2020;368:m441.
42 Victora CG, Huttly SR, Fuchs SC, *et al*. The role of conceptual frameworks in epidemiological analysis: a hierarchical approach. *Int J Epidemiol* 1997;26:224–7.
43 Villain B, Dechartres A, Boyer P, *et al*. Feasibility of individual patient data meta-analyses in orthopaedic surgery. *BMC Med* 2015;13:131.
44 Lyu W, Kim J-S, Suk Y. Estimating Heterogeneous Treatment Effects Within Latent Class Multilevel Models: A Bayesian Approach. *J Educ Behav Stat* 2023;48:3–36.